Vector spaces and Fourier Theory

Syllabus

Syllabus

1 **Vector spaces and linear maps:** Definitions and examples.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.
6. **Theorems about bases:** Invariance of dimension, rank-nullity formula, $\dim(U+V) + \dim(U \cap V) = \dim(U) + \dim(V)$.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.
6. **Theorems about bases:** Invariance of dimension, rank-nullity formula, $\dim(U+V)+\dim(U\cap V)=\dim(U)+\dim(V)$.
7. **Eigenvalues and eigenvectors:** for abstract vector spaces.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.
6. **Theorems about bases:** Invariance of dimension, rank-nullity formula, $\dim(U+V)+\dim(U\cap V)=\dim(U)+\dim(V)$.
7. **Eigenvalues and eigenvectors:** for abstract vector spaces.
8. **Inner products:** Definitions and examples. The Cauchy-Schwartz inequality, projections, the Gram-Schmidt procedure.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.
6. **Theorems about bases:** Invariance of dimension, rank-nullity formula, $\dim(U+V)+\dim(U\cap V)=\dim(U)+\dim(V)$.
7. **Eigenvalues and eigenvectors:** for abstract vector spaces.
8. **Inner products:** Definitions and examples. The Cauchy-Schwartz inequality, projections, the Gram-Schmidt procedure.
9. **Adjoints of linear maps:** Definition, proof of existence and uniqueness.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.
6. **Theorems about bases:** Invariance of dimension, rank-nullity formula, $\dim(U+V)+\dim(U\cap V)=\dim(U)+\dim(V)$.
7. **Eigenvalues and eigenvectors:** for abstract vector spaces.
8. **Inner products:** Definitions and examples. The Cauchy-Schwartz inequality, projections, the Gram-Schmidt procedure.
9. **Adjoints of linear maps:** Definition, proof of existence and uniqueness.
10. **Diagonalisation of self-adjoint operators:** Self-adjoint operators have real eigenvalues, and admit an orthonormal basis of eigenvectors.

Syllabus

1. **Vector spaces and linear maps:** Definitions and examples.
2. **Subspaces:** Definitions and examples, (direct) sums and intersections.
3. **Independence and spanning sets:** Definitions and examples. Bases.
4. **Linear maps out of $\mathbb{R}^n$:** A linear map $\mathbb{R}^n \to V$ is the same as a list of n elements of V.
5. **Matrices for linear maps:** Definitions, properties, change of basis.
6. **Theorems about bases:** Invariance of dimension, rank-nullity formula, $\dim(U + V) + \dim(U \cap V) = \dim(U) + \dim(V)$.
7. **Eigenvalues and eigenvectors:** for abstract vector spaces.
8. **Inner products:** Definitions and examples. The Cauchy-Schwartz inequality, projections, the Gram-Schmidt procedure.
9. **Adjoints of linear maps:** Definition, proof of existence and uniqueness.
10. **Diagonalisation of self-adjoint operators:** Self-adjoint operators have real eigenvalues, and admit an orthonormal basis of eigenvectors.
11. **Fourier theory:** in terms of inner product spaces.

Predefinition ??: A *vector space* (over $\mathbb{R}$) is a nonempty set V of things such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If u is an element of V and t is a real number, then tu is an element of V.

Predefinition ??: A *vector space* (over $\mathbb{R}$) is a nonempty set V of things such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If u is an element of V and t is a real number, then tu is an element of V.

This definition is not strictly meaningful or rigorous; we will pick holes in it later. But it will do for the moment.

Predefinition ??: A *vector space* (over $\mathbb{R}$) is a nonempty set V of things such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If u is an element of V and t is a real number, then tu is an element of V.

This definition is not strictly meaningful or rigorous; we will pick holes in it later. But it will do for the moment.

Example ??: The set $\mathbb{R}^3$ of all three-dimensional vectors is a vector space, because the sum of two vectors is a vector (eg $\begin{bmatrix}1\\2\\3\end{bmatrix} + \begin{bmatrix}3\\2\\1\end{bmatrix} = \begin{bmatrix}4\\4\\4\end{bmatrix}$) and the product of a real number and a vector is a vector (eg $3\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}3\\6\\9\end{bmatrix}$). In the same way, the set $\mathbb{R}^2$ of two-dimensional vectors is also a vector space.

We generally use column vectors (rather than row vectors), as this makes formulae with matrix multiplication work better.

However, column vectors often fit awkwardly on the page, so we use the following notational device:

$$[1,\ 2,\ 3,\ 4]^T \qquad \text{means} \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$[a+b,\ b+c,\ c+d,\ d+e,\ e+a]^T \qquad \text{means} \qquad \begin{bmatrix} a+b \\ b+c \\ c+d \\ d+e \\ e+a \end{bmatrix}$$

and so on.

The space $\mathbb{R}^n$

The space $\mathbb{R}^n$

Example ??: For any natural number n the set $\mathbb{R}^n$ of vectors of length n is a vector space. For example, the vectors $u = \begin{bmatrix} 1 & 2 & 4 & 8 & 16 \end{bmatrix}^T$ and $v = \begin{bmatrix} 1 & -2 & 4 & -8 & 16 \end{bmatrix}^T$ are elements of $\mathbb{R}^5$, with $u + v = \begin{bmatrix} 2 & 0 & 8 & 0 & 32 \end{bmatrix}^T$. We can even consider the set $\mathbb{R}^\infty$ of all infinite sequences of real numbers, which is again a vector space.

The space $\mathbb{R}^n$

Example ??: For any natural number n the set $\mathbb{R}^n$ of vectors of length n is a vector space. For example, the vectors $u = \begin{bmatrix} 1 & 2 & 4 & 8 & 16 \end{bmatrix}^T$ and $v = \begin{bmatrix} 1 & -2 & 4 & -8 & 16 \end{bmatrix}^T$ are elements of $\mathbb{R}^5$, with $u + v = \begin{bmatrix} 2 & 0 & 8 & 0 & 32 \end{bmatrix}^T$. We can even consider the set $\mathbb{R}^\infty$ of all infinite sequences of real numbers, which is again a vector space.

Example ??: The set $\{0\}$ is a trivial example of a vector space (but it is important in the same way that the number zero is important). This space can also be thought of as $\mathbb{R}^0$.

Example ??: For any natural number n the set $\mathbb{R}^n$ of vectors of length n is a vector space. For example, the vectors $u = \begin{bmatrix} 1 & 2 & 4 & 8 & 16 \end{bmatrix}^T$ and $v = \begin{bmatrix} 1 & -2 & 4 & -8 & 16 \end{bmatrix}^T$ are elements of $\mathbb{R}^5$, with $u + v = \begin{bmatrix} 2 & 0 & 8 & 0 & 32 \end{bmatrix}^T$. We can even consider the set $\mathbb{R}^\infty$ of all infinite sequences of real numbers, which is again a vector space.

Example ??: The set $\{0\}$ is a trivial example of a vector space (but it is important in the same way that the number zero is important). This space can also be thought of as $\mathbb{R}^0$. We often write it as 0 rather than $\{0\}$.

Example ??: For any natural number n the set $\mathbb{R}^n$ of vectors of length n is a vector space. For example, the vectors $u = \begin{bmatrix} 1 & 2 & 4 & 8 & 16 \end{bmatrix}^T$ and $v = \begin{bmatrix} 1 & -2 & 4 & -8 & 16 \end{bmatrix}^T$ are elements of $\mathbb{R}^5$, with $u + v = \begin{bmatrix} 2 & 0 & 8 & 0 & 32 \end{bmatrix}^T$. We can even consider the set $\mathbb{R}^\infty$ of all infinite sequences of real numbers, which is again a vector space.

Example ??: The set $\{0\}$ is a trivial example of a vector space (but it is important in the same way that the number zero is important). This space can also be thought of as $\mathbb{R}^0$. We often write it as 0 rather than $\{0\}$.

Another trivial example is that $\mathbb{R}$ itself is a vector space (which can be thought of as $\mathbb{R}^1$).

Physical vectors

Example ??: The set U of physical vectors is a vector space.

Example ??: The set U of physical vectors is a vector space.
We can define some elements of U by

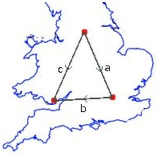
- $\mathbf{a}$ is the vector from Sheffield to London
- $\mathbf{b}$ is the vector from London to Cardiff
- $\mathbf{c}$ is the vector from Sheffield to Cardiff
- $\mathbf{d}$ is the vector from the centre of the earth to the north pole
- $\mathbf{e}$ is the vector from the south pole to the north pole.

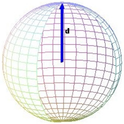
Physical vectors

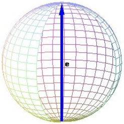
Example ??: The set U of physical vectors is a vector space.
We can define some elements of U by

- $\mathbf{a}$ is the vector from Sheffield to London
- $\mathbf{b}$ is the vector from London to Cardiff
- $\mathbf{c}$ is the vector from Sheffield to Cardiff
- $\mathbf{d}$ is the vector from the centre of the earth to the north pole
- $\mathbf{e}$ is the vector from the south pole to the north pole.

We then have $\mathbf{a}+\mathbf{b}=\mathbf{c}$ and $2\mathbf{d}=\mathbf{e}$.





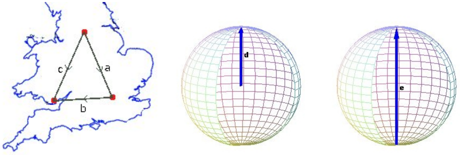


Physical vectors

Example ??: The set U of physical vectors is a vector space.
We can define some elements of U by

- $\mathbf{a}$ is the vector from Sheffield to London
- $\mathbf{b}$ is the vector from London to Cardiff
- $\mathbf{c}$ is the vector from Sheffield to Cardiff
- $\mathbf{d}$ is the vector from the centre of the earth to the north pole
- $\mathbf{e}$ is the vector from the south pole to the north pole.

We then have $\mathbf{a} + \mathbf{b} = \mathbf{c}$ and $2\mathbf{d} = \mathbf{e}$.



Once we have agreed on where our axes should point, and what units of length we should use, we can identify U with $\mathbb{R}^3$.

The space of functions

Example ??: The set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space, because we can add any two functions to get a new function, and we can multiply a function by a number to get a new function.

Example ??: The set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space, because we can add any two functions to get a new function, and we can multiply a function by a number to get a new function.

For example, we can define functions $f, g, h\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = e^x \qquad g(x) = e^{-x} \qquad h(x) = \cosh(x) = \frac{e^x + e^{-x}}{2},$$

so f, g and h are elements of $F(\mathbb{R})$.

Example ??: The set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space, because we can add any two functions to get a new function, and we can multiply a function by a number to get a new function.

For example, we can define functions $f, g, h\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = e^x \qquad g(x) = e^{-x} \qquad h(x) = \cosh(x) = \frac{e^x + e^{-x}}{2},$$

so f, g and h are elements of $F(\mathbb{R})$.

Then $f + g$ and $2h$ are again functions, in other words $f + g \in F(\mathbb{R})$ and $2h \in F(\mathbb{R})$. Of course we actually have $f + g = 2h$ in this example.

Example ??: The set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space, because we can add any two functions to get a new function, and we can multiply a function by a number to get a new function.

For example, we can define functions $f, g, h\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = e^x \qquad g(x) = e^{-x} \qquad h(x) = \cosh(x) = \frac{e^x + e^{-x}}{2},$$

so f, g and h are elements of $F(\mathbb{R})$.

Then $f + g$ and $2h$ are again functions, in other words $f + g \in F(\mathbb{R})$ and $2h \in F(\mathbb{R})$. Of course we actually have $f + g = 2h$ in this example.

For this to work properly, we must insist that $f(x)$ is defined for all x, and is a real number for all x; it cannot be infinite or imaginary. Thus the rules $p(x) = 1/x$ and $q(x) = \sqrt{x}$ do not define elements $p, q \in F(\mathbb{R})$.

Smaller spaces of functions

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions.

Smaller spaces of functions

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions

Smaller spaces of functions

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions
- The set $C^{\infty}(\mathbb{R})$ of "smooth" functions (those which can be differentiated arbitrarily often)

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions
- The set $C^\infty(\mathbb{R})$ of "smooth" functions
  (those which can be differentiated arbitrarily often)
- the set $\mathbb{R}[x]$ of polynomial functions

Smaller spaces of functions

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions
- The set $C^{\infty}(\mathbb{R})$ of "smooth" functions
  (those which can be differentiated arbitrarily often)
- the set $\mathbb{R}[x]$ of polynomial functions
  (eg $p(x) = 1 + x + x^2 + x^3$ defines an element $p \in \mathbb{R}[x]$)

Smaller spaces of functions

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions
- The set $C^\infty(\mathbb{R})$ of "smooth" functions
  (those which can be differentiated arbitrarily often)
- the set $\mathbb{R}[x]$ of polynomial functions
  (eg $p(x) = 1 + x + x^2 + x^3$ defines an element $p \in \mathbb{R}[x]$)

If f and g are continuous then $f + g$ and tf are continuous, so $C(\mathbb{R})$ is a vector space.

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions
- The set $C^\infty(\mathbb{R})$ of "smooth" functions (those which can be differentiated arbitrarily often)
- the set $\mathbb{R}[x]$ of polynomial functions (eg $p(x) = 1 + x + x^2 + x^3$ defines an element $p \in \mathbb{R}[x]$)

If f and g are continuous then $f + g$ and tf are continuous, so $C(\mathbb{R})$ is a vector space.

If f and g are smooth then $f + g$ and tf are smooth, so $C^\infty(\mathbb{R})$ is a vector space.

Example ??: In practise, we do not generally want to consider the set $F(\mathbb{R})$ of *all* functions. Instead we consider

- The set $C(\mathbb{R})$ of continuous functions
- The set $C^\infty(\mathbb{R})$ of "smooth" functions
  (those which can be differentiated arbitrarily often)
- the set $\mathbb{R}[x]$ of polynomial functions
  (eg $p(x) = 1 + x + x^2 + x^3$ defines an element $p \in \mathbb{R}[x]$)

If f and g are continuous then $f + g$ and tf are continuous, so $C(\mathbb{R})$ is a vector space.

If f and g are smooth then $f + g$ and tf are smooth, so $C^\infty(\mathbb{R})$ is a vector space.

If f and g are polynomials then $f + g$ and tf are polynomials, so $\mathbb{R}[x]$ is a vector space.

Smaller spaces of functions

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

Smaller spaces of functions

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

We write $C[a, b]$ for the set of continuous functions $f : [a, b] \to \mathbb{R}$.

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

We write $C[a, b]$ for the set of continuous functions $f : [a, b] \to \mathbb{R}$.

For example, the rule $f(x) = 1/x$ defines a continuous function on the interval $[1, 2]$.

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

We write $C[a, b]$ for the set of continuous functions $f : [a, b] \to \mathbb{R}$.

For example, the rule $f(x) = 1/x$ defines a continuous function on the interval $[1, 2]$. (The only potential problem is at the point $x = 0$, but $0 \notin [1, 2]$, so we do not need to worry about it.)

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

We write $C[a, b]$ for the set of continuous functions $f\colon [a, b] \to \mathbb{R}$.

For example, the rule $f(x) = 1/x$ defines a continuous function on the interval $[1, 2]$. (The only potential problem is at the point $x = 0$, but $0 \notin [1, 2]$, so we do not need to worry about it.)

We thus have an element $f \in C[1, 2]$.

Smaller spaces of functions

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

We write $C[a, b]$ for the set of continuous functions $f \colon [a, b] \to \mathbb{R}$.

For example, the rule $f(x) = 1/x$ defines a continuous function on the interval $[1, 2]$. (The only potential problem is at the point $x = 0$, but $0 \notin [1, 2]$, so we do not need to worry about it.)

We thus have an element $f \in C[1, 2]$.

We can define another element $g \in C[1, 2]$ by $g(x) = 2/|x|$.

Smaller spaces of functions

Example ??: Let $[a, b]$ denote the interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$

We write $C[a, b]$ for the set of continuous functions $f : [a, b] \to \mathbb{R}$.

For example, the rule $f(x) = 1/x$ defines a continuous function on the interval $[1, 2]$. (The only potential problem is at the point $x = 0$, but $0 \notin [1, 2]$, so we do not need to worry about it.)

We thus have an element $f \in C[1, 2]$.

We can define another element $g \in C[1, 2]$ by $g(x) = 2/|x|$.

We actually have $g = 2f$, because f and g are defined as functions on $[1, 2]$, and $|x| = x$ for all $x \in [1, 2]$.

Spaces of matrices

Example ??: The set $M_2\mathbb{R}$ of 2×2 matrices (with real entries) is a vector space.

Example ??: The set $M_2\mathbb{R}$ of 2×2 matrices (with real entries) is a vector space. Indeed, if we add two such matrices, we get another 2×2 matrix, for example

$$\left[\begin{smallmatrix} 1 & 0 \\ 0 & 4 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0 & 2 \\ 3 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right].$$

Example ??: The set $M_2\mathbb{R}$ of 2×2 matrices (with real entries) is a vector space. Indeed, if we add two such matrices, we get another 2×2 matrix, for example

$$\left[\begin{smallmatrix} 1 & 0 \\ 0 & 4 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0 & 2 \\ 3 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right].$$

Similarly, if we multiply a 2×2 matrix by a real number, we get another 2×2 matrix, for example

$$7\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 7 & 14 \\ 21 & 28 \end{smallmatrix}\right].$$

Example ??: The set $M_2\mathbb{R}$ of 2×2 matrices (with real entries) is a vector space. Indeed, if we add two such matrices, we get another 2×2 matrix, for example

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Similarly, if we multiply a 2×2 matrix by a real number, we get another 2×2 matrix, for example

$$7 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 14 \\ 21 & 28 \end{bmatrix}.$$

We can identify $M_2\mathbb{R}$ with $\mathbb{R}^4$, by the rule

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \leftrightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Example ??: The set $M_2\mathbb{R}$ of 2×2 matrices (with real entries) is a vector space. Indeed, if we add two such matrices, we get another 2×2 matrix, for example

$$\left[\begin{smallmatrix} 1 & 0 \\ 0 & 4 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0 & 2 \\ 3 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right].$$

Similarly, if we multiply a 2×2 matrix by a real number, we get another 2×2 matrix, for example

$$7\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 7 & 14 \\ 21 & 28 \end{smallmatrix}\right].$$

We can identify $M_2\mathbb{R}$ with $\mathbb{R}^4$, by the rule

$$\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \leftrightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

More generally, for any n the set $M_n\mathbb{R}$ of $n \times n$ square matrices is a vector space, which can be identified with $\mathbb{R}^{n^2}$.

Example ??: The set $M_2\mathbb{R}$ of 2×2 matrices (with real entries) is a vector space. Indeed, if we add two such matrices, we get another 2×2 matrix, for example

$$\left[\begin{smallmatrix} 1 & 0 \\ 0 & 4 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0 & 2 \\ 3 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right].$$

Similarly, if we multiply a 2×2 matrix by a real number, we get another 2×2 matrix, for example

$$7\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 7 & 14 \\ 21 & 28 \end{smallmatrix}\right].$$

We can identify $M_2\mathbb{R}$ with $\mathbb{R}^4$, by the rule

$$\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \leftrightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

More generally, for any n the set $M_n\mathbb{R}$ of $n \times n$ square matrices is a vector space, which can be identified with $\mathbb{R}^{n^2}$.

More generally still, for any n and m, the set $M_{n,m}\mathbb{R}$ of $n \times m$ matrices is a vector space, which can be identified with $\mathbb{R}^{nm}$.

The set of all lists

Example ??: Let L be the set of all finite lists of real numbers.

For example, the lists $\mathbf{a} = (10, 20, 30, 40)$ and $\mathbf{b} = (5, 6, 7)$ and $\mathbf{c} = (1, \pi, \pi^2)$ define three elements $\mathbf{a}, \mathbf{b}, \mathbf{c} \in L$.

The set of all lists

Example ??: Let L be the set of all finite lists of real numbers.

For example, the lists $\mathbf{a} = (10, 20, 30, 40)$ and $\mathbf{b} = (5, 6, 7)$ and $\mathbf{c} = (1, \pi, \pi^2)$ define three elements $\mathbf{a}, \mathbf{b}, \mathbf{c} \in L$. Is L a vector space?

The set of all lists

Example ??: Let L be the set of all finite lists of real numbers.

For example, the lists $\mathbf{a} = (10, 20, 30, 40)$ and $\mathbf{b} = (5, 6, 7)$ and $\mathbf{c} = (1, \pi, \pi^2)$ define three elements $\mathbf{a}, \mathbf{b}, \mathbf{c} \in L$. Is L a vector space?

In trying to answer this question, we will reveal some inadequacies of Predefinition **??**.

Example ??: Let L be the set of all finite lists of real numbers.

For example, the lists $\mathbf{a} = (10, 20, 30, 40)$ and $\mathbf{b} = (5, 6, 7)$ and $\mathbf{c} = (1, \pi, \pi^2)$ define three elements $\mathbf{a}, \mathbf{b}, \mathbf{c} \in L$. Is L a vector space?

In trying to answer this question, we will reveal some inadequacies of Predefinition **??**.

It seems clear that L is closed under scalar multiplication: for example $100\mathbf{b} = (500, 600, 700)$, which is another element of L.

The set of all lists

Example ??: Let L be the set of all finite lists of real numbers.

For example, the lists $\mathbf{a} = (10, 20, 30, 40)$ and $\mathbf{b} = (5, 6, 7)$ and $\mathbf{c} = (1, \pi, \pi^2)$ define three elements $\mathbf{a}, \mathbf{b}, \mathbf{c} \in L$. Is L a vector space?

In trying to answer this question, we will reveal some inadequacies of Predefinition **??**.

It seems clear that L is closed under scalar multiplication: for example $100\mathbf{b} = (500, 600, 700)$, which is another element of L.

The real issue is closure under addition.
For example, is $\mathbf{a} + \mathbf{b}$ an element of L?

Example ??: Let L be the set of all finite lists of real numbers.

For example, the lists $\mathbf{a} = (10, 20, 30, 40)$ and $\mathbf{b} = (5, 6, 7)$ and $\mathbf{c} = (1, \pi, \pi^2)$ define three elements $\mathbf{a}, \mathbf{b}, \mathbf{c} \in L$. Is L a vector space?

In trying to answer this question, we will reveal some inadequacies of Predefinition **??**.

It seems clear that L is closed under scalar multiplication: for example $100\mathbf{b} = (500, 600, 700)$, which is another element of L.

The real issue is closure under addition.
For example, is $\mathbf{a} + \mathbf{b}$ an element of L?

We cannot answer this unless we know what $\mathbf{a} + \mathbf{b}$ means.

The set of all lists

The set of all lists

$L = \{$ all finite lists of real numbers $\}$ $\qquad \mathbf{a} = (10, 20, 30, 40) \qquad \mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean?

The set of all lists

$L = \{$ all finite lists of real numbers $\}$ $\qquad$ $\mathbf{a} = (10, 20, 30, 40)$ $\qquad$ $\mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean? There are at least three possible meanings:

The set of all lists

$L = \{$ all finite lists of real numbers $\}$ $\quad$ $\mathbf{a} = (10, 20, 30, 40)$ $\quad$ $\mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean? There are at least three possible meanings:

(1) $\mathbf{a} + \mathbf{b}$ could mean $(10, 20, 30, 40, 5, 6, 7)$
(the list $\mathbf{a}$ followed by the list $\mathbf{b}$).

The set of all lists

$L = \{$ all finite lists of real numbers $\}$ $\quad$ $\mathbf{a} = (10, 20, 30, 40)$ $\quad$ $\mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean? There are at least three possible meanings:

(1) $\mathbf{a} + \mathbf{b}$ could mean $(10, 20, 30, 40, 5, 6, 7)$
(the list $\mathbf{a}$ followed by the list $\mathbf{b}$).

(2) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37)$
(chop off $\mathbf{a}$ to make the lists the same length, then add them together).

The set of all lists

$L = \{$ all finite lists of real numbers $\}$ $\quad \mathbf{a} = (10, 20, 30, 40)$ $\quad \mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean? There are at least three possible meanings:

(1) $\mathbf{a} + \mathbf{b}$ could mean $(10, 20, 30, 40, 5, 6, 7)$
(the list **a** followed by the list **b**).

(2) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37)$
(chop off **a** to make the lists the same length, then add them together).

(3) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37, 40)$
(add zeros to the end of **b** to make the lists the same length, then add them together.)

The set of all lists

$L = \{$ all finite lists of real numbers $\}$ $\quad$ $\mathbf{a} = (10, 20, 30, 40)$ $\quad$ $\mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean? There are at least three possible meanings:

(1) $\mathbf{a} + \mathbf{b}$ could mean $(10, 20, 30, 40, 5, 6, 7)$
(the list $\mathbf{a}$ followed by the list $\mathbf{b}$).

(2) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37)$
(chop off $\mathbf{a}$ to make the lists the same length, then add them together).

(3) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37, 40)$
(add zeros to the end of $\mathbf{b}$ to make the lists the same length, then add them together.)

The point is that the expression $\mathbf{a} + \mathbf{b}$ does not have a meaning until we decide to give it one.

$L = \{$ all finite lists of real numbers $\}$ $\quad$ $\mathbf{a} = (10, 20, 30, 40)$ $\quad$ $\mathbf{b} = (5, 6, 7)$

What does $\mathbf{a} + \mathbf{b}$ mean? There are at least three possible meanings:

(1) $\mathbf{a} + \mathbf{b}$ could mean $(10, 20, 30, 40, 5, 6, 7)$
(the list $\mathbf{a}$ followed by the list $\mathbf{b}$).

(2) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37)$
(chop off $\mathbf{a}$ to make the lists the same length, then add them together).

(3) $\mathbf{a} + \mathbf{b}$ could mean $(15, 26, 37, 40)$
(add zeros to the end of $\mathbf{b}$ to make the lists the same length, then add them together.)

The point is that the expression $\mathbf{a} + \mathbf{b}$ does not have a meaning until we decide to give it one.

(Strictly speaking, the same is true of the expression $100\mathbf{b}$, but in that case there is only one reasonable possibility for what it should mean.)

The set of all lists

To avoid this kind of ambiguity, we should say that a vector space is a set *together with a definition of addition* etc.

The set of all lists

To avoid this kind of ambiguity, we should say that a vector space is a set *together with a definition of addition* etc.

Suppose we use the 3rd definition of addition, so $\mathbf{a} + \mathbf{b} = (15, 26, 37, 40)$.

To avoid this kind of ambiguity, we should say that a vector space is a set *together with a definition of addition* etc.

Suppose we use the 3rd definition of addition, so $\mathbf{a} + \mathbf{b} = (15, 26, 37, 40)$.

The ordinary rules of algebra would tell us that $(\mathbf{a} + (-1).\mathbf{a}) + \mathbf{b} = \mathbf{b}$.

The set of all lists

To avoid this kind of ambiguity, we should say that a vector space is a set *together with a definition of addition* etc.

Suppose we use the 3rd definition of addition, so $\mathbf{a}+\mathbf{b}=(15,26,37,40)$.

The ordinary rules of algebra would tell us that $(\mathbf{a}+(-1).\mathbf{a})+\mathbf{b}=\mathbf{b}$.

However, in fact we have

$$\begin{aligned}(\mathbf{a}+(-1).\mathbf{a})+\mathbf{b} &= ((10,20,30,40)+(-10,-20,-30,-40))+(5,6,7) \\ &= (0,0,0,0)+(5,6,7)=(5,6,7,0)\neq(5,6,7)=\mathbf{b}.\end{aligned}$$

Thus, the ordinary rules of algebra do not hold.

The set of all lists

To avoid this kind of ambiguity, we should say that a vector space is a set *together with a definition of addition* etc.

Suppose we use the 3rd definition of addition, so $\mathbf{a} + \mathbf{b} = (15, 26, 37, 40)$.

The ordinary rules of algebra would tell us that $(\mathbf{a} + (-1).\mathbf{a}) + \mathbf{b} = \mathbf{b}$.

However, in fact we have

$$\begin{aligned}(\mathbf{a} + (-1).\mathbf{a}) + \mathbf{b} &= ((10, 20, 30, 40) + (-10, -20, -30, -40)) + (5, 6, 7) \\ &= (0, 0, 0, 0) + (5, 6, 7) = (5, 6, 7, 0) \neq (5, 6, 7) = \mathbf{b}.\end{aligned}$$

Thus, the ordinary rules of algebra do not hold.

We do not want to deal with this kind of thing; we only want to consider sets where addition and scalar multiplication work in the usual way. We must therefore give a more careful definition of a vector space, which will allow us to say that L is not a vector space, so we need not think about it.

The set of all lists

To avoid this kind of ambiguity, we should say that a vector space is a set *together with a definition of addition* etc.

Suppose we use the 3rd definition of addition, so $\mathbf{a}+\mathbf{b}=(15,26,37,40)$.

The ordinary rules of algebra would tell us that $(\mathbf{a}+(-1).\mathbf{a})+\mathbf{b}=\mathbf{b}$.

However, in fact we have

$$\begin{aligned}(\mathbf{a}+(-1).\mathbf{a})+\mathbf{b} &= ((10,20,30,40)+(-10,-20,-30,-40))+(5,6,7) \\ &= (0,0,0,0)+(5,6,7)=(5,6,7,0)\neq(5,6,7)=\mathbf{b}.\end{aligned}$$

Thus, the ordinary rules of algebra do not hold.

We do not want to deal with this kind of thing; we only want to consider sets where addition and scalar multiplication work in the usual way. We must therefore give a more careful definition of a vector space, which will allow us to say that L is not a vector space, so we need not think about it.

(If we used either of the other definitions of addition then things would still go wrong; details are left as an exercise.)

A more precise definition

A more precise definition

Our next attempt at a definition is as follows:

A more precise definition

Our next attempt at a definition is as follows:

Predefinition ??: A *vector space* over $\mathbb{R}$ is a nonempty set V, together with a definition of what it means to add elements of V or multiply them by real numbers, such that

(a) If u and v are elements of V, then $u+v$ is an also an element of V.

(b) If u is an element of V and t is a real number, then tu is an element of V.

(c) All the usual algebraic rules for addition and multiplication hold.

A more precise definition

Our next attempt at a definition is as follows:

Predefinition ??: A *vector space* over $\mathbb{R}$ is a nonempty set V, together with a definition of what it means to add elements of V or multiply them by real numbers, such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If u is an element of V and t is a real number, then tu is an element of V.

(c) All the usual algebraic rules for addition and multiplication hold.

In the course we will be content with an informal understanding of the phrase "all the usual algebraic rules", but for completeness, we will give an explicit list of axioms.

The real definition

The real definition

Definition ??: A *vector space* over $\mathbb{R}$ is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by real numbers

Definition ??: A *vector space* over $\mathbb{R}$ is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by real numbers, such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is a real number, then tv is an element of V.

The real definition

Definition ??: A *vector space* over $\mathbb{R}$ is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by real numbers, such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is a real number, then tv is an element of V.

(c) For any elements $u, v, w \in V$ and any real numbers s, t, the following equations hold:

(1) $0 + v = v$
(2) $u + v = v + u$
(3) $u + (v + w) = (u + v) + w$
(4) $0u = 0$
(5) $1u = u$
(6) $(st)u = s(tu)$
(7) $(s + t)u = su + tu$
(8) $s(u + v) = su + sv$.

Definition ??: A *vector space* over $\mathbb{R}$ is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by real numbers, such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is a real number, then tv is an element of V.

(c) For any elements $u, v, w \in V$ and any real numbers s, t, the following equations hold:

(1) $0 + v = v$
(2) $u + v = v + u$
(3) $u + (v + w) = (u + v) + w$
(4) $0u = 0$
(5) $1u = u$
(6) $(st)u = s(tu)$
(7) $(s + t)u = su + tu$
(8) $s(u + v) = su + sv$.

Note that there are many rules that do not appear explicitly in the above list, such as the fact that $t(u + v - w/t) = tu + tv - w$, but it turns out that all such rules can be deduced from the ones listed. We will not discuss any such deductions.

Zero

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Thus 0 could mean

- the vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (if we are working with $\mathbb{R}^3$)

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Thus 0 could mean

- the vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (if we are working with $\mathbb{R}^3$)
- the zero matrix $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ (if we are working with $M_{2,3}\mathbb{R}$)

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Thus 0 could mean

- the vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (if we are working with $\mathbb{R}^3$)
- the zero matrix $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ (if we are working with $M_{2,3}\mathbb{R}$)
- the zero function (if we are working with $C(\mathbb{R})$)

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Thus 0 could mean

- the vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (if we are working with $\mathbb{R}^3$)
- the zero matrix $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ (if we are working with $M_{2,3}\mathbb{R}$)
- the zero function (if we are working with $C(\mathbb{R})$)

or whatever.

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Thus 0 could mean

- the vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (if we are working with $\mathbb{R}^3$)
- the zero matrix $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ (if we are working with $M_{2,3}\mathbb{R}$)
- the zero function (if we are working with $C(\mathbb{R})$)

or whatever.

Occasionally it will be important to distinguish between the zero elements in different vector spaces. In that case, we write 0_V for the zero element of V.

Zero

Remark ??: We will usually use the symbol 0 for the zero element of whatever vector space we are considering.

Thus 0 could mean

- the vector $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (if we are working with $\mathbb{R}^3$)
- the zero matrix $\left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ (if we are working with $M_{2,3}\mathbb{R}$)
- the zero function (if we are working with $C(\mathbb{R})$)

or whatever.

Occasionally it will be important to distinguish between the zero elements in different vector spaces. In that case, we write 0_V for the zero element of V.

For example:

$$0_{\mathbb{R}^2} = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \qquad\qquad 0_{M_2\mathbb{R}} = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right].$$

Vector spaces over other fields

Vector spaces over other fields

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

Definition ??: A *field* is a set K together with elements $0, 1 \in K$ and a definition of what it means to add or multiply two elements of K

Vector spaces over other fields

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

Definition ??: A *field* is a set K together with elements $0, 1 \in K$ and a definition of what it means to add or multiply two elements of K, such that:

(a) The usual rules of algebra are valid.

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

Definition ??: A *field* is a set K together with elements $0, 1 \in K$ and a definition of what it means to add or multiply two elements of K, such that:

(a) The usual rules of algebra are valid. More explicitly, for all $a, b, c \in K$ the following equations hold:

- $0 + a = a$
- $a + (b + c) = (a + b) + c$
- $a + b = b + a$
- $1.a = a$
- $a(bc) = (ab)c$
- $ab = ba$
- $a(b + c) = ab + ac$

Vector spaces over other fields

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

Definition ??: A *field* is a set K together with elements $0, 1 \in K$ and a definition of what it means to add or multiply two elements of K, such that:

(a) The usual rules of algebra are valid. More explicitly, for all $a, b, c \in K$ the following equations hold:

- $0 + a = a$
- $a + (b + c) = (a + b) + c$
- $a + b = b + a$
- $1.a = a$
- $a(bc) = (ab)c$
- $ab = ba$
- $a(b + c) = ab + ac$

(b) For every $a \in K$ there is an element $-a$ with $a + (-a) = 0$.

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

Definition ??: A *field* is a set K together with elements $0, 1 \in K$ and a definition of what it means to add or multiply two elements of K, such that:

(a) The usual rules of algebra are valid. More explicitly, for all $a, b, c \in K$ the following equations hold:
- $0 + a = a$
- $a + (b + c) = (a + b) + c$
- $a + b = b + a$
- $1.a = a$
- $a(bc) = (ab)c$
- $ab = ba$
- $a(b + c) = ab + ac$

(b) For every $a \in K$ there is an element $-a$ with $a + (-a) = 0$.

(c) For every $a \in K$ with $a \neq 0$ there is an element $a^{-1} \in K$ with $aa^{-1} = 1$.

One can also consider vector spaces over fields other than $\mathbb{R}$; the most important case for us will be the field $\mathbb{C}$ of complex numbers. We record the definitions for completeness.

Definition ??: A *field* is a set K together with elements $0, 1 \in K$ and a definition of what it means to add or multiply two elements of K, such that:

(a) The usual rules of algebra are valid. More explicitly, for all $a, b, c \in K$ the following equations hold:

- $0 + a = a$
- $a + (b + c) = (a + b) + c$
- $a + b = b + a$
- $1.a = a$
- $a(bc) = (ab)c$
- $ab = ba$
- $a(b + c) = ab + ac$

(b) For every $a \in K$ there is an element $-a$ with $a + (-a) = 0$.

(c) For every $a \in K$ with $a \neq 0$ there is an element $a^{-1} \in K$ with $aa^{-1} = 1$.

(d) $1 \neq 0$ (or equivalently, $K \neq \{0\}$).

Examples of fields

Examples of fields

Example ??: Recall that

$$\begin{aligned}
\mathbb{Z} &= \{ \text{ integers } \} = \{\dots, -2, -1, 0, 1, 2, 3, 4, \dots\} \\
\mathbb{Q} &= \{ \text{ rational numbers } \} = \{a/b \mid a, b \in \mathbb{Z}\ ,\ \ b \neq 0\} \\
\mathbb{R} &= \{ \text{ real numbers } \} \\
\mathbb{C} &= \{ \text{ complex numbers } \} = \{x + iy \mid x, y \in \mathbb{R}\},
\end{aligned}$$

Example ??: Recall that

$$\begin{aligned}
\mathbb{Z} &= \{ \text{ integers } \} = \{\dots, -2, -1, 0, 1, 2, 3, 4, \dots\} \\
\mathbb{Q} &= \{ \text{ rational numbers } \} = \{a/b \mid a, b \in \mathbb{Z}\ ,\ \ b \neq 0\} \\
\mathbb{R} &= \{ \text{ real numbers } \} \\
\mathbb{C} &= \{ \text{ complex numbers } \} = \{x + iy \mid x, y \in \mathbb{R}\},
\end{aligned}$$

so $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Example ??: Recall that

$$\begin{aligned}
\mathbb{Z} &= \{ \text{ integers } \} = \{\ldots, -2, -1, 0, 1, 2, 3, 4, \ldots\} \\
\mathbb{Q} &= \{ \text{ rational numbers } \} = \{a/b \mid a, b \in \mathbb{Z}\ ,\ \ b \neq 0\} \\
\mathbb{R} &= \{ \text{ real numbers } \} \\
\mathbb{C} &= \{ \text{ complex numbers } \} = \{x + iy \mid x, y \in \mathbb{R}\},
\end{aligned}$$

so $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Then $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Q}$ are fields.

Examples of fields

Example ??: Recall that

$$\begin{aligned} \mathbb{Z} &= \{ \text{ integers } \} = \{\dots, -2, -1, 0, 1, 2, 3, 4, \dots\} \\ \mathbb{Q} &= \{ \text{ rational numbers } \} = \{a/b \mid a, b \in \mathbb{Z}\ ,\ \ b \neq 0\} \\ \mathbb{R} &= \{ \text{ real numbers } \} \\ \mathbb{C} &= \{ \text{ complex numbers } \} = \{x + iy \mid x, y \in \mathbb{R}\}, \end{aligned}$$

so $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Then $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Q}$ are fields.

The ring $\mathbb{Z}$ is not a field, because axiom (c) is not satisfied: there is no element 2^{-1} *in the set* $\mathbb{Z}$ for which $2.2^{-1} = 1$.

Example ??: Recall that

$$\mathbb{Z} = \{ \text{ integers } \} = \{\dots, -2, -1, 0, 1, 2, 3, 4, \dots\}$$
$$\mathbb{Q} = \{ \text{ rational numbers } \} = \{a/b \mid a, b \in \mathbb{Z}\ ,\ \ b \neq 0\}$$
$$\mathbb{R} = \{ \text{ real numbers } \}$$
$$\mathbb{C} = \{ \text{ complex numbers } \} = \{x + iy \mid x, y \in \mathbb{R}\},$$

so $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Then $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Q}$ are fields.

The ring $\mathbb{Z}$ is not a field, because axiom (c) is not satisfied: there is no element 2^{-1} *in the set* $\mathbb{Z}$ for which $2.2^{-1} = 1$.

One can show that the ring $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is a prime number.

Vector spaces over other fields

Definition ??: A *vector space* over a field K is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by elements of K

Definition ??: A *vector space* over a field K is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by elements of K , such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is an element of K, then $tv \in V$.

Vector spaces over other fields

Definition ??: A *vector space* over a field K is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by elements of K , such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is an element of K, then $tv \in V$.

(c) For any elements $u, v, w \in V$ and any elements $s, t \in K$, the following equations hold:

(1) $0 + v = v$
(2) $u + v = v + u$
(3) $u + (v + w) = (u + v) + w$
(4) $0u = 0$
(5) $1u = u$
(6) $(st)u = s(tu)$
(7) $(s + t)u = su + tu$
(8) $s(u + v) = su + sv$.

Vector spaces over other fields

Definition ??: A *vector space* over a field K is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by elements of K , such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is an element of K, then $tv \in V$.

(c) For any elements $u, v, w \in V$ and any elements $s, t \in K$, the following equations hold:

(1) $0 + v = v$	(5) $1u = u$
(2) $u + v = v + u$	(6) $(st)u = s(tu)$
(3) $u + (v + w) = (u + v) + w$	(7) $(s + t)u = su + tu$
(4) $0u = 0$	(8) $s(u + v) = su + sv$.

Example ??: Almost all our examples work over any field K.

Definition ??: A *vector space* over a field K is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by elements of K , such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is an element of K, then $tv \in V$.

(c) For any elements $u, v, w \in V$ and any elements $s, t \in K$, the following equations hold:

(1) $0 + v = v$
(2) $u + v = v + u$
(3) $u + (v + w) = (u + v) + w$
(4) $0u = 0$
(5) $1u = u$
(6) $(st)u = s(tu)$
(7) $(s + t)u = su + tu$
(8) $s(u + v) = su + sv$.

Example ??: Almost all our examples work over any field K.
$M_4\mathbb{Q} = \{4 \times 4 \text{ matrices with entries in } \mathbb{Q}\}$ is a vector space over $\mathbb{Q}$.

Definition ??: A *vector space* over a field K is a set V, together with an element $0 \in V$ and a definition of what it means to add elements of V or multiply them by elements of K, such that

(a) If u and v are elements of V, then $u + v$ is an also an element of V.

(b) If v is an element of V and t is an element of K, then $tv \in V$.

(c) For any elements $u, v, w \in V$ and any elements $s, t \in K$, the following equations hold:

(1) $0 + v = v$	(5) $1u = u$
(2) $u + v = v + u$	(6) $(st)u = s(tu)$
(3) $u + (v + w) = (u + v) + w$	(7) $(s + t)u = su + tu$
(4) $0u = 0$	(8) $s(u + v) = su + sv$.

Example ??: Almost all our examples work over any field K.
$M_4\mathbb{Q} = \{4 \times 4 \text{ matrices with entries in } \mathbb{Q}\}$ is a vector space over $\mathbb{Q}$.
$\mathbb{C}[x] = \{\text{polynomials with complex coefficients}\}$ is a vector space over $\mathbb{C}$.

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).

Definition ??: Let V, W be vector spaces. Let $\phi \colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v+v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Remark ??: The definition can be reformulated slightly as follows. A map $\phi\colon V \to W$ is linear iff

(c) For any $t, t' \in \mathbb{R}$ and any $v, v' \in V$ we have $\phi(tv + t'v') = t\phi(v) + t'\phi(v')$.

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Remark ??: The definition can be reformulated slightly as follows. A map $\phi\colon V \to W$ is linear iff

(c) For any $t, t' \in \mathbb{R}$ and any $v, v' \in V$ we have
$\phi(tv + t'v') = t\phi(v) + t'\phi(v')$.

To show that this reformulation is valid, we must show that if (c) holds, then so do (a) and (b); and conversely, if (a) and (b) hold, then so does (c).

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Remark ??: The definition can be reformulated slightly as follows. A map $\phi\colon V \to W$ is linear iff

(c) For any $t, t' \in \mathbb{R}$ and any $v, v' \in V$ we have
$\phi(tv + t'v') = t\phi(v) + t'\phi(v')$.

To show that this reformulation is valid, we must show that if (c) holds, then so do (a) and (b); and conversely, if (a) and (b) hold, then so does (c).

This is left as an exercise.

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

Linear maps

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Remark ??: The definition can be reformulated slightly as follows. A map $\phi\colon V \to W$ is linear iff

(c) For any $t, t' \in \mathbb{R}$ and any $v, v' \in V$ we have
$\phi(tv + t'v') = t\phi(v) + t'\phi(v')$.

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Remark ??: The definition can be reformulated slightly as follows. A map $\phi\colon V \to W$ is linear iff

(c) For any $t, t' \in \mathbb{R}$ and any $v, v' \in V$ we have $\phi(tv + t'v') = t\phi(v) + t'\phi(v')$.

To show that this reformulation is valid, we must show that if (c) holds, then so do (a) and (b); and conversely, if (a) and (b) hold, then so does (c).

Definition ??: Let V, W be vector spaces. Let $\phi\colon V \to W$ be a function (so for each element $v \in V$ we have an element $\phi(v) \in W$).
We say that ϕ is *linear* if

(a) For any v and v' in V, we have $\phi(v + v') = \phi(v) + \phi(v')$ in W.

(b) For any $t \in \mathbb{R}$ and $v \in V$ we have $\phi(tv) = t\phi(v)$ in W.

By taking $t = v = 0$ in (b), we see that a linear map must satisfy $\phi(0) = 0$.

Further simple arguments also show that $\phi(v - v') = \phi(v) - \phi(v')$.

Remark ??: The definition can be reformulated slightly as follows. A map $\phi\colon V \to W$ is linear iff

(c) For any $t, t' \in \mathbb{R}$ and any $v, v' \in V$ we have
$\phi(tv + t'v') = t\phi(v) + t'\phi(v')$.

To show that this reformulation is valid, we must show that if (c) holds, then so do (a) and (b); and conversely, if (a) and (b) hold, then so does (c).

This is left as an exercise. $\bigcirc$

Linear maps from $\mathbb{R}$ to $\mathbb{R}$

Linear maps from $\mathbb{R}$ to $\mathbb{R}$

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

On the other hand:
$$f(x + x') = 2(x + x') = 2x + 2x' = f(x) + f(x')$$
$$f(tx) = 2tx = tf(x) \qquad \text{so } f \text{ is linear.}$$

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

On the other hand: $$f(x + x') = 2(x + x') = 2x + 2x' = f(x) + f(x')$$ $$f(tx) = 2tx = tf(x) \qquad \text{so } f \text{ is linear.}$$

Example ??: For any number $m \in \mathbb{R}$, we can define $\mu_m\colon \mathbb{R} \to \mathbb{R}$ by $\mu_m(x) = mx$

Example ??: Consider $f, g \colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

On the other hand:

$$f(x + x') = 2(x + x') = 2x + 2x' = f(x) + f(x')$$
$$f(tx) = 2tx = tf(x) \qquad \text{so } f \text{ is linear.}$$

Example ??: For any number $m \in \mathbb{R}$, we can define $\mu_m \colon \mathbb{R} \to \mathbb{R}$ by $\mu_m(x) = mx$ (so f in the previous example is μ_2).

Linear maps from $\mathbb{R}$ to $\mathbb{R}$

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

On the other hand:
$$f(x + x') = 2(x + x') = 2x + 2x' = f(x) + f(x')$$
$$f(tx) = 2tx = tf(x) \qquad \text{so } f \text{ is linear.}$$

Example ??: For any number $m \in \mathbb{R}$, we can define $\mu_m\colon \mathbb{R} \to \mathbb{R}$ by $\mu_m(x) = mx$ (so f in the previous example is μ_2). We have

$$\mu_m(x + x') = m(x + x') = mx + mx' = \mu_m(x) + \mu_m(x')$$
$$\mu_m(tx) = mtx = tmx = t\mu_m(x),$$

Linear maps from $\mathbb{R}$ to $\mathbb{R}$

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

On the other hand:
$$f(x + x') = 2(x + x') = 2x + 2x' = f(x) + f(x')$$
$$f(tx) = 2tx = tf(x) \qquad \text{so } f \text{ is linear.}$$

Example ??: For any number $m \in \mathbb{R}$, we can define $\mu_m\colon \mathbb{R} \to \mathbb{R}$ by $\mu_m(x) = mx$ (so f in the previous example is μ_2). We have

$$\mu_m(x + x') = m(x + x') = mx + mx' = \mu_m(x) + \mu_m(x')$$
$$\mu_m(tx) = mtx = tmx = t\mu_m(x),$$

so μ_m is linear.

Example ??: Consider $f, g\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x$ and $g(x) = x^2$.

$g(x + x') = x^2 + x'^2 + 2xx' \neq x^2 + x'^2 = g(x) + g(x')$, so g is not linear.

Similarly, $\sin(x + x') \neq \sin(x) + \sin(x')$ so sin is not linear.

On the other hand:

$$f(x + x') = 2(x + x') = 2x + 2x' = f(x) + f(x')$$
$$f(tx) = 2tx = tf(x) \qquad \text{so } f \text{ is linear.}$$

Example ??: For any number $m \in \mathbb{R}$, we can define $\mu_m\colon \mathbb{R} \to \mathbb{R}$ by $\mu_m(x) = mx$ (so f in the previous example is μ_2). We have

$$\mu_m(x + x') = m(x + x') = mx + mx' = \mu_m(x) + \mu_m(x')$$
$$\mu_m(tx) = mtx = tmx = t\mu_m(x),$$

so μ_m is linear. Note also that the graph of μ_m is a straight line of slope m through the origin; this is essentially the reason for the word "linear".

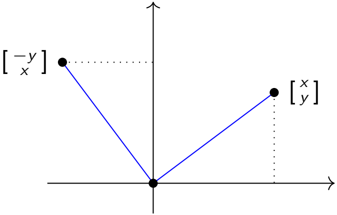
Rotation is linear

Rotation is linear

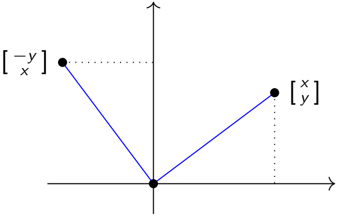
Example ??: For any $\mathbf{v} \in \mathbb{R}^2$, we let $\rho(\mathbf{v})$ be the vector obtained by rotating $\mathbf{v}$ through 90 degrees anticlockwise around the origin.

Rotation is linear

Example ??: For any $\mathbf{v} \in \mathbb{R}^2$, we let $\rho(\mathbf{v})$ be the vector obtained by rotating $\mathbf{v}$ through 90 degrees anticlockwise around the origin. It is well-known that the formula for this is $\rho\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right]$.



Example ??: For any $\mathbf{v} \in \mathbb{R}^2$, we let $\rho(\mathbf{v})$ be the vector obtained by rotating $\mathbf{v}$ through 90 degrees anticlockwise around the origin. It is well-known that the formula for this is $\rho\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right]$.



We thus have

$$\rho\left(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x' \\ y' \end{bmatrix}\right) = \rho\begin{bmatrix} x+x' \\ y+y' \end{bmatrix} = \begin{bmatrix} -y-y' \\ x+x' \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} + \begin{bmatrix} -y' \\ x' \end{bmatrix} = \rho\begin{bmatrix} x \\ y \end{bmatrix} + \rho\begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$\rho\left(t\begin{bmatrix} x \\ y \end{bmatrix}\right) = \rho\begin{bmatrix} tx \\ ty \end{bmatrix} = \begin{bmatrix} -ty \\ tx \end{bmatrix} = t\rho\begin{bmatrix} x \\ y \end{bmatrix},$$

so ρ is linear. ○

More general rotations

More general rotations

More generally, let $\mathrm{rot}_\theta(\mathbf{v})$ be the vector obtained by rotating $\mathbf{v}$ anticlockwise by an angle of θ around the origin.

More general rotations

More generally, let $\text{rot}_\theta(\mathbf{v})$ be the vector obtained by rotating $\mathbf{v}$ anticlockwise by an angle of θ around the origin.

Then

$$\text{rot}_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta)x - \sin(\theta)y \\ \sin(\theta)x + \cos(\theta)y \end{bmatrix}$$

More general rotations

More generally, let $\mathrm{rot}_\theta(\mathbf{v})$ be the vector obtained by rotating $\mathbf{v}$ anticlockwise by an angle of θ around the origin.

Then

$$\mathrm{rot}_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta)x - \sin(\theta)y \\ \sin(\theta)x + \cos(\theta)y \end{bmatrix}$$

Using this, we see that $\mathrm{rot}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map.

Reflection is linear

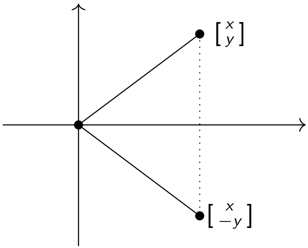
Reflection is linear

Example ??: For any $\mathbf{v} \in \mathbb{R}^2$, we let $\tau(\mathbf{v})$ be the vector obtained by reflecting $\mathbf{v}$ across the line $y = 0$.

Reflection is linear

Example ??: For any $\mathbf{v} \in \mathbb{R}^2$, we let $\tau(\mathbf{v})$ be the vector obtained by reflecting $\mathbf{v}$ across the line $y = 0$.

It is clear that the formula is $\tau \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x \\ -y \end{smallmatrix}\right]$, and using this we see that τ is linear.



More general reflections

More general reflections

More generally, let $\text{ref}_\theta(\mathbf{v})$ be the vector obtained by reflecting $\mathbf{v}$ across the line crossing the x-axis at an angle of $\theta/2$.

More general reflections

More generally, let $\mathrm{ref}_\theta(\mathbf{v})$ be the vector obtained by reflecting $\mathbf{v}$ across the line crossing the x-axis at an angle of $\theta/2$.

Then

$$\mathrm{ref}_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta)x+\sin(\theta)y \\ \sin(\theta)x-\cos(\theta)y \end{bmatrix}$$

More general reflections

More generally, let $\mathrm{ref}_\theta(\mathbf{v})$ be the vector obtained by reflecting $\mathbf{v}$ across the line crossing the x-axis at an angle of $\theta/2$.

Then

$$\mathrm{ref}_\theta \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta)x+\sin(\theta)y \\ \sin(\theta)x-\cos(\theta)y \end{bmatrix}$$

Using this, we see that $\mathrm{ref}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map.

Some nonlinear examples

Some nonlinear examples

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$

Some nonlinear examples

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta \left[\begin{smallmatrix} x \\ y \end{smallmatrix} \right] = \sqrt{x^2 + y^2}$.

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2 + y^2}$.
This is not linear, because $\theta(\mathbf{u} + \mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2+y^2}$.
This is not linear, because $\theta(\mathbf{u}+\mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u})+\theta(\mathbf{v}) = 1+1 = 2$.

Some nonlinear examples

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2+y^2}$.
This is not linear, because $\theta(\mathbf{u}+\mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u})+\theta(\mathbf{v}) = 1+1 = 2$.

Example ??: Define $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\sigma\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x+1 \\ y-1 \end{smallmatrix}\right]$.

Some nonlinear examples

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2+y^2}$.
This is not linear, because $\theta(\mathbf{u}+\mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u})+\theta(\mathbf{v}) = 1+1 = 2$.

Example ??: Define $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\sigma\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x+1 \\ y-1 \end{smallmatrix}\right]$.
Then σ is not linear, because $\sigma\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.

Some nonlinear examples

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2 + y^2}$.
This is not linear, because $\theta(\mathbf{u} + \mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u})+\theta(\mathbf{v}) = 1+1 = 2$.

Example ??: Define $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\sigma\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x+1 \\ y-1 \end{smallmatrix}\right]$.
Then σ is not linear, because $\sigma\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.

Example ??: Define $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\alpha\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} y^3/(x^2+y^2) \\ x^3/(x^2+y^2) \end{smallmatrix}\right]$.

Some nonlinear examples

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2 + y^2}$.
This is not linear, because $\theta(\mathbf{u} + \mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u}) + \theta(\mathbf{v}) = 1 + 1 = 2$.

Example ??: Define $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\sigma\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x+1 \\ y-1 \end{smallmatrix}\right]$.
Then σ is not linear, because $\sigma\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.

Example ??: Define $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\alpha\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} y^3/(x^2+y^2) \\ x^3/(x^2+y^2) \end{smallmatrix}\right]$.
(This does not really make sense when $x = y = 0$, but for that case we make the separate definition that $\alpha\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.)

Some nonlinear examples

Example ??: Define $\theta \colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2 + y^2}$.
This is not linear, because $\theta(\mathbf{u} + \mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u}) + \theta(\mathbf{v}) = 1 + 1 = 2$.

Example ??: Define $\sigma \colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\sigma \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x+1 \\ y-1 \end{smallmatrix}\right]$.
Then σ is not linear, because $\sigma \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.

Example ??: Define $\alpha \colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\alpha \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} y^3/(x^2+y^2) \\ x^3/(x^2+y^2) \end{smallmatrix}\right]$.
(This does not really make sense when $x = y = 0$, but for that case we make the separate definition that $\alpha \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.)
This map satisfies $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$, but it does not satisfy $\alpha(\mathbf{u} + \mathbf{v}) = \alpha(\mathbf{u}) + \alpha(\mathbf{v})$, so it is not linear.

Example ??: Define $\theta\colon \mathbb{R}^2 \to \mathbb{R}$ by $\theta(\mathbf{v}) = \|\mathbf{v}\|$ so $\theta\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \sqrt{x^2+y^2}$.
This is not linear, because $\theta(\mathbf{u}+\mathbf{v}) \neq \theta(\mathbf{u}) + \theta(\mathbf{v})$ in general.
Indeed, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right]$ then $\theta(\mathbf{u}+\mathbf{v}) = 0$ but $\theta(\mathbf{u})+\theta(\mathbf{v}) = 1+1 = 2$.

Example ??: Define $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\sigma\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x+1 \\ y-1 \end{smallmatrix}\right]$.
Then σ is not linear, because $\sigma\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.

Example ??: Define $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $\alpha\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} y^3/(x^2+y^2) \\ x^3/(x^2+y^2) \end{smallmatrix}\right]$.
(This does not really make sense when $x = y = 0$, but for that case we make the separate definition that $\alpha\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.)
This map satisfies $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$, but it does not satisfy $\alpha(\mathbf{u}+\mathbf{v}) = \alpha(\mathbf{u}) + \alpha(\mathbf{v})$, so it is not linear.

For example, if $\mathbf{u} = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ then $\alpha(\mathbf{u}) = \mathbf{v}$ and $\alpha(\mathbf{v}) = \mathbf{u}$ but $\alpha(\mathbf{u}+\mathbf{v}) = (\mathbf{u}+\mathbf{v})/2 \neq \alpha(\mathbf{u}) + \alpha(\mathbf{v})$.

Examples from vector algebra

Examples from vector algebra

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Examples from vector algebra

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

$$\det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix} = \det \begin{bmatrix} u_2 & u_3 \\ v_2 & v_3 \end{bmatrix} \mathbf{e}_1 - \det \begin{bmatrix} u_1 & u_3 \\ v_1 & v_3 \end{bmatrix} \mathbf{e}_2 + \det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \mathbf{e}_3 = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} = \mathbf{u} \times \mathbf{v}$$

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

To prove this we must show that $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$ and $\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w})$ and $\beta(t\mathbf{v}) = t\beta(\mathbf{v})$ and $\beta(\mathbf{v} + \mathbf{w}) = \beta(\mathbf{v}) + \beta(\mathbf{w})$.

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

To prove this we must show that $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$ and $\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w})$ and $\beta(t\mathbf{v}) = t\beta(\mathbf{v})$ and $\beta(\mathbf{v} + \mathbf{w}) = \beta(\mathbf{v}) + \beta(\mathbf{w})$.
We will write out only the last of these; the others are similar but easier.

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

To prove this we must show that $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$ and $\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w})$ and $\beta(t\mathbf{v}) = t\beta(\mathbf{v})$ and $\beta(\mathbf{v} + \mathbf{w}) = \beta(\mathbf{v}) + \beta(\mathbf{w})$.
We will write out only the last of these; the others are similar but easier.

$$\beta(\mathbf{v} + \mathbf{w}) = \beta \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix}$$

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

To prove this we must show that $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$ and $\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w})$ and $\beta(t\mathbf{v}) = t\beta(\mathbf{v})$ and $\beta(\mathbf{v} + \mathbf{w}) = \beta(\mathbf{v}) + \beta(\mathbf{w})$.
We will write out only the last of these; the others are similar but easier.

$$\beta(\mathbf{v} + \mathbf{w}) = \beta \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix} = \begin{bmatrix} a_2(v_3 + w_3) - a_3(v_2 + w_2) \\ a_3(v_1 + w_1) - a_1(v_3 + w_3) \\ a_1(v_2 + w_2) - a_2(v_1 + w_1) \end{bmatrix}$$

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

To prove this we must show that $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$ and $\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w})$ and $\beta(t\mathbf{v}) = t\beta(\mathbf{v})$ and $\beta(\mathbf{v} + \mathbf{w}) = \beta(\mathbf{v}) + \beta(\mathbf{w})$.
We will write out only the last of these; the others are similar but easier.

$$\beta(\mathbf{v} + \mathbf{w}) = \beta \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix} = \begin{bmatrix} a_2(v_3 + w_3) - a_3(v_2 + w_2) \\ a_3(v_1 + w_1) - a_1(v_3 + w_3) \\ a_1(v_2 + w_2) - a_2(v_1 + w_1) \end{bmatrix}$$

$$= \begin{bmatrix} a_2 v_3 - a_3 v_2 \\ a_3 v_1 - a_1 v_3 \\ a_1 v_2 - a_2 v_1 \end{bmatrix} + \begin{bmatrix} a_2 w_3 - a_3 w_2 \\ a_3 w_1 - a_1 w_3 \\ a_1 w_2 - a_2 w_1 \end{bmatrix}$$

Example ??: Given vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$, recall that the inner product and cross product are defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}.\mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Fix a vector $\mathbf{a} \in \mathbb{R}^3$. Define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}$ by $\alpha(\mathbf{v}) = \langle \mathbf{a}, \mathbf{v} \rangle$ and $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. Then both α and β are linear.

To prove this we must show that $\alpha(t\mathbf{v}) = t\alpha(\mathbf{v})$ and $\alpha(\mathbf{v} + \mathbf{w}) = \alpha(\mathbf{v}) + \alpha(\mathbf{w})$ and $\beta(t\mathbf{v}) = t\beta(\mathbf{v})$ and $\beta(\mathbf{v} + \mathbf{w}) = \beta(\mathbf{v}) + \beta(\mathbf{w})$.
We will write out only the last of these; the others are similar but easier.

$$\beta(\mathbf{v} + \mathbf{w}) = \beta \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix} = \begin{bmatrix} a_2(v_3+w_3) - a_3(v_2+w_2) \\ a_3(v_1+w_1) - a_1(v_3+w_3) \\ a_1(v_2+w_2) - a_2(v_1+w_1) \end{bmatrix}$$

$$= \begin{bmatrix} a_2 v_3 - a_3 v_2 \\ a_3 v_1 - a_1 v_3 \\ a_1 v_2 - a_2 v_1 \end{bmatrix} + \begin{bmatrix} a_2 w_3 - a_3 w_2 \\ a_3 w_1 - a_1 w_3 \\ a_1 w_2 - a_2 w_1 \end{bmatrix} = \beta(\mathbf{v}) + \beta(\mathbf{w}).$$

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

It is clear that $A(\mathbf{v} + \mathbf{v}') = A\mathbf{v} + A\mathbf{v}'$ and $At\mathbf{v} = tA\mathbf{v}$, so ϕ_A is a linear map.

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

It is clear that $A(\mathbf{v} + \mathbf{v}') = A\mathbf{v} + A\mathbf{v}'$ and $At\mathbf{v} = tA\mathbf{v}$, so ϕ_A is a linear map.

We will see later that every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ has this form.

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

It is clear that $A(\mathbf{v} + \mathbf{v}') = A\mathbf{v} + A\mathbf{v}'$ and $At\mathbf{v} = tA\mathbf{v}$, so ϕ_A is a linear map.

We will see later that every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ has this form.

In particular, if we put $R = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

It is clear that $A(\mathbf{v} + \mathbf{v}') = A\mathbf{v} + A\mathbf{v}'$ and $At\mathbf{v} = tA\mathbf{v}$, so ϕ_A is a linear map.

We will see later that every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ has this form.

In particular, if we put $R = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, we find that

$$R\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right] = \rho\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) \qquad\qquad T\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x \\ -y \end{smallmatrix}\right] = \tau\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right)$$

(where ρ and τ are as in Examples **??** and **??**).

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

It is clear that $A(\mathbf{v} + \mathbf{v}') = A\mathbf{v} + A\mathbf{v}'$ and $At\mathbf{v} = tA\mathbf{v}$, so ϕ_A is a linear map.

We will see later that every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ has this form.

In particular, if we put $R = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, we find that

$$R\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right] = \rho\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) \qquad\qquad T\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x \\ -y \end{smallmatrix}\right] = \tau\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right)$$

(where ρ and τ are as in Examples **??** and **??**).
This means that $\rho = \phi_R$ and $\tau = \phi_T$.

Multiplication by a matrix is linear

Example ??: Let A be a fixed $m \times n$ matrix.

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we can multiply A by $\mathbf{v}$ to get a vector $A\mathbf{v} \in \mathbb{R}^m$.

We can thus define $\phi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

It is clear that $A(\mathbf{v} + \mathbf{v}') = A\mathbf{v} + A\mathbf{v}'$ and $At\mathbf{v} = tA\mathbf{v}$, so ϕ_A is a linear map.

We will see later that every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ has this form.

In particular, if we put $R = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, we find that

$$R\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right] = \rho\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) \qquad T\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} x \\ -y \end{smallmatrix}\right] = \tau\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right)$$

(where ρ and τ are as in Examples **??** and **??**).
This means that $\rho = \phi_R$ and $\tau = \phi_T$.

More generally, $\mathrm{rot}_\theta = \phi_{R_\theta}$ and $\mathrm{ref}_\theta = \phi_{T_\theta}$, where

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \qquad T_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

Definite integration is linear

Definite integration is linear

Example ??: For any continuous function $f : \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

Example ??: For any continuous function $f : \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I : C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad\qquad I(p) = \int_0^1 x^2\,dx$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad\qquad I(p) = \int_0^1 x^2\,dx = [x^3/3]_0^1$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad\qquad I(p) = \int_0^1 x^2\, dx = [x^3/3]_0^1 = 1/3$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad I(p) = \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3$$

$$q(x) = 2x - 1$$

Definite integration is linear

Example ??: For any continuous function $f: \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I: C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad I(p) = \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3$$

$$q(x) = 2x - 1 \qquad I(q) = \int_0^1 2x - 1\,dx$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad I(p) = \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3$$

$$q(x) = 2x - 1 \qquad I(q) = \int_0^1 2x - 1\,dx = [x^2 - x]_0^1$$

Definite integration is linear

Example ??: For any continuous function $f : \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I : C(\mathbb{R}) \to \mathbb{R}$.

$$\begin{aligned} p(x) &= x^2 & I(p) &= \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3 \\ q(x) &= 2x - 1 & I(q) &= \int_0^1 2x - 1\,dx = [x^2 - x]_0^1 = 0 \end{aligned}$$

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$p(x) = x^2 \qquad I(p) = \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3$$

$$q(x) = 2x - 1 \qquad I(q) = \int_0^1 2x - 1\,dx = [x^2 - x]_0^1 = 0$$

$$r(x) = e^x$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$\begin{aligned}
p(x) &= x^2 & I(p) &= \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3 \\
q(x) &= 2x - 1 & I(q) &= \int_0^1 2x - 1\,dx = [x^2 - x]_0^1 = 0 \\
r(x) &= e^x & I(r) &= \int_0^1 e^x\,dx
\end{aligned}$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$
\begin{aligned}
p(x) &= x^2 & \qquad I(p) &= \int_0^1 x^2\, dx = [x^3/3]_0^1 = 1/3 \\
q(x) &= 2x - 1 & \qquad I(q) &= \int_0^1 2x - 1\, dx = [x^2 - x]_0^1 = 0 \\
r(x) &= e^x & \qquad I(r) &= \int_0^1 e^x\, dx = [e^x]_0^1
\end{aligned}
$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$\begin{aligned} p(x) &= x^2 & I(p) &= \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3 \\ q(x) &= 2x-1 & I(q) &= \int_0^1 2x-1\,dx = [x^2-x]_0^1 = 0 \\ r(x) &= e^x & I(r) &= \int_0^1 e^x\,dx = [e^x]_0^1 = e-1. \end{aligned}$$

Definite integration is linear

Example ??: For any continuous function $f\colon \mathbb{R} \to \mathbb{R}$, we write

$$I(f) = \int_0^1 f(x)dx \in \mathbb{R}.$$

This defines a map $I\colon C(\mathbb{R}) \to \mathbb{R}$.

$$\begin{aligned}
p(x) &= x^2 & I(p) &= \int_0^1 x^2\,dx = [x^3/3]_0^1 = 1/3 \\
q(x) &= 2x - 1 & I(q) &= \int_0^1 2x - 1\,dx = [x^2 - x]_0^1 = 0 \\
r(x) &= e^x & I(r) &= \int_0^1 e^x\,dx = [e^x]_0^1 = e - 1.
\end{aligned}$$

Using the obvious equations

$$\int_0^1 f(x) + g(x)dx = \int_0^1 f(x)dx + \int_0^1 g(x)dx$$
$$\int_0^1 tf(x)dx = t\int_0^1 f(x)dx$$

we see that $I(f+g) = I(f) + I(g)$ and $I(tf) = t\,I(f)$, so I is a linear map.

Linear differential operators

Linear differential operators

Definition ??: For smooth $f : \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

Linear differential operators

Definition ??: For smooth $f\colon \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

These are again smooth, so $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ and $L\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.

Linear differential operators

Definition ??: For smooth $f\colon \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

These are again smooth, so $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ and $L\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.

$$
\begin{aligned}
p(x) &= \sin(x) \\
q(x) &= \cos(x) \\
r(x) &= e^x
\end{aligned}
$$

Linear differential operators

Definition ??: For smooth $f\colon \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

These are again smooth, so $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ and $L\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.

$$\begin{aligned} p(x) &= \sin(x) & \qquad D(p) &= q \\ q(x) &= \cos(x) & D(q) &= -p \\ r(x) &= e^x & D(r) &= r \end{aligned}$$

Linear differential operators

Definition ??: For smooth $f\colon \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

These are again smooth, so $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ and $L\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.

$$
\begin{array}{lll}
p(x) = \sin(x) & D(p) = q & L(p) = 0 \\
q(x) = \cos(x) & D(q) = -p & L(q) = 0 \\
r(x) = e^x & D(r) = r & L(r) = 2r
\end{array}
$$

Linear differential operators

Definition ??: For smooth $f : \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

These are again smooth, so $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ and $L\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.

$$
\begin{aligned}
p(x) &= \sin(x) & D(p) &= q & L(p) &= 0 \\
q(x) &= \cos(x) & D(q) &= -p & L(q) &= 0 \\
r(x) &= e^x & D(r) &= r & L(r) &= 2r
\end{aligned}
$$

Using the equations $(f+g)' = f' + g'$ and $(tf)' = t\,f'$ we see that D is linear.

Definition ??: For smooth $f\colon \mathbb{R} \to \mathbb{R}$ put $D(f) = f'$ and $L(f) = f'' + f$.

These are again smooth, so $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ and $L\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$.

$$
\begin{array}{lll}
p(x) = \sin(x) & D(p) = q & L(p) = 0 \\
q(x) = \cos(x) & D(q) = -p & L(q) = 0 \\
r(x) = e^x & D(r) = r & L(r) = 2r
\end{array}
$$

Using the equations $(f+g)' = f' + g'$ and $(tf)' = t\,f'$ we see that D is linear. Similarly, we have

$$
\begin{aligned}
L(f+g) &= (f+g)'' + (f+g) = f'' + g'' + f + g \\
&= (f'' + f) + (g'' + g) = L(f) + L(g) \\
L(tf) &= (tf)'' + tf = t\,f'' + tf = tL(f).
\end{aligned}
$$

This shows that L is also linear. ○

Trace and determinant

Trace and determinant

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\mathrm{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

Trace and determinant

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\text{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

We thus have two functions $\text{trace}, \det \colon M_2\mathbb{R} \to \mathbb{R}$.

Trace and determinant

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\operatorname{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

We thus have two functions $\operatorname{trace}, \det\colon M_2\mathbb{R} \to \mathbb{R}$.

It is easy to see that $\operatorname{trace}(A + B) = \operatorname{trace}(A) + \operatorname{trace}(B)$ and $\operatorname{trace}(tA) = t\operatorname{trace}(A)$, so $\operatorname{trace}\colon M_2\mathbb{R} \to \mathbb{R}$ is a linear map.

Trace and determinant

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\mathrm{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

We thus have two functions $\mathrm{trace}, \det \colon M_2\mathbb{R} \to \mathbb{R}$.

It is easy to see that $\mathrm{trace}(A + B) = \mathrm{trace}(A) + \mathrm{trace}(B)$ and $\mathrm{trace}(tA) = t\,\mathrm{trace}(A)$, so $\mathrm{trace} \colon M_2\mathbb{R} \to \mathbb{R}$ is a linear map.

On the other hand, $\det(tA) = t^2 \det(A)$ and $\det(A + B) \neq \det(A) + \det(B)$ in general, so $\det \colon M_2\mathbb{R} \to \mathbb{R}$ is not a linear map.

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\mathrm{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

We thus have two functions $\mathrm{trace}, \det \colon M_2\mathbb{R} \to \mathbb{R}$.

It is easy to see that $\mathrm{trace}(A + B) = \mathrm{trace}(A) + \mathrm{trace}(B)$ and $\mathrm{trace}(tA) = t\,\mathrm{trace}(A)$, so $\mathrm{trace} \colon M_2\mathbb{R} \to \mathbb{R}$ is a linear map.

On the other hand, $\det(tA) = t^2 \det(A)$ and $\det(A + B) \neq \det(A) + \det(B)$ in general, so $\det \colon M_2\mathbb{R} \to \mathbb{R}$ is not a linear map.

For a specific counterexample, take $\quad A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \quad$ and $\quad B = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$

Trace and determinant

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\operatorname{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

We thus have two functions $\operatorname{trace}, \det \colon M_2\mathbb{R} \to \mathbb{R}$.

It is easy to see that $\operatorname{trace}(A + B) = \operatorname{trace}(A) + \operatorname{trace}(B)$ and $\operatorname{trace}(tA) = t \operatorname{trace}(A)$, so $\operatorname{trace} \colon M_2\mathbb{R} \to \mathbb{R}$ is a linear map.

On the other hand, $\det(tA) = t^2 \det(A)$ and $\det(A + B) \neq \det(A) + \det(B)$ in general, so $\det \colon M_2\mathbb{R} \to \mathbb{R}$ is not a linear map.

For a specific counterexample, take $A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$

$\det(A) = \det(B) = 0$ but $\det(A + B) = 1$, so $\det(A + B) \neq \det(A) + \det(B)$.

Example ??: For any 2×2 matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the trace and determinant are defined by $\mathrm{trace}(A) = a + d \in \mathbb{R}$ and $\det(A) = ad - bc \in \mathbb{R}$.

We thus have two functions $\mathrm{trace}, \det \colon M_2\mathbb{R} \to \mathbb{R}$.

It is easy to see that $\mathrm{trace}(A + B) = \mathrm{trace}(A) + \mathrm{trace}(B)$ and $\mathrm{trace}(tA) = t\,\mathrm{trace}(A)$, so $\mathrm{trace} \colon M_2\mathbb{R} \to \mathbb{R}$ is a linear map.

On the other hand, $\det(tA) = t^2 \det(A)$ and $\det(A + B) \neq \det(A) + \det(B)$ in general, so $\det \colon M_2\mathbb{R} \to \mathbb{R}$ is not a linear map.

For a specific counterexample, take $\quad A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \quad$ and $\quad B = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$

$\det(A) = \det(B) = 0$ but $\det(A + B) = 1$, so $\det(A + B) \neq \det(A) + \det(B)$.

None of this is really restricted to 2×2 matrices. For any n we have a map $\mathrm{trace} \colon M_n\mathbb{R} \to \mathbb{R}$ given by $\mathrm{trace}(A) = \sum_{i=1}^{n} A_{ii}$, which is again linear. We also have a determinant map $\det \colon M_n\mathbb{R} \to \mathbb{R}$ which satisfies $\det(tI) = t^n$; this shows that det is not linear, except in the silly case where $n = 1$. ◯

Matrix inversion is not linear

Matrix inversion is not linear

Example ??: "Define" $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$

Matrix inversion is not linear

Example ??: "Define" $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$, so

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \begin{bmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{bmatrix}.$$

Matrix inversion is not linear

Example ??: "Define" $\phi \colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$, so

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \begin{bmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{bmatrix}.$$

This is not a linear map, simply because it is not a well-defined map at all: the "definition" does not make sense when $ad - bc = 0$.

Matrix inversion is not linear

Example ??: "Define" $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$, so

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{smallmatrix}\right].$$

This is not a linear map, simply because it is not a well-defined map at all: the "definition" does not make sense when $ad - bc = 0$.

Even if it were well-defined, it would not be linear

Matrix inversion is not linear

Example ??: "Define" $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$, so

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{smallmatrix}\right].$$

This is not a linear map, simply because it is not a well-defined map at all: the "definition" does not make sense when $ad - bc = 0$.

Even if it were well-defined, it would not be linear, because $\phi(I + I) = (2I)^{-1} = I/2$

Matrix inversion is not linear

Example ??: "Define" $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$, so

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{smallmatrix}\right].$$

This is not a linear map, simply because it is not a well-defined map at all: the "definition" does not make sense when $ad - bc = 0$.

Even if it were well-defined, it would not be linear, because $\phi(I+I) = (2I)^{-1} = I/2$, whereas $\phi(I) + \phi(I) = 2I$

Matrix inversion is not linear

Example ??: "Define" $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^{-1}$, so

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{smallmatrix}\right].$$

This is not a linear map, simply because it is not a well-defined map at all: the "definition" does not make sense when $ad - bc = 0$.

Even if it were well-defined, it would not be linear, because $\phi(I + I) = (2I)^{-1} = I/2$, whereas $\phi(I) + \phi(I) = 2I$, so $\phi(I + I) \neq \phi(I) + \phi(I)$.
$\bigcirc$

Row reduction is not linear

Row reduction is not linear

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

Row reduction is not linear

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

For example, we have the following sequence of reductions:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix}$$

Row reduction is not linear

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

For example, we have the following sequence of reductions:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & 0 & -12 \end{bmatrix}$$

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

For example, we have the following sequence of reductions:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -12 \end{bmatrix}$$

Row reduction is not linear

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

For example, we have the following sequence of reductions:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

For example, we have the following sequence of reductions:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

which shows that

$$\phi \begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Row reduction is not linear

Example ??: Define $\phi\colon M_3\mathbb{R} \to M_3\mathbb{R}$ by

$$\phi(A) = \text{ the row reduced echelon form of } A.$$

For example, we have the following sequence of reductions:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

which shows that

$$\phi \begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 6 \\ 7 & 14 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The map is not linear, because $\phi(I) = I$ and also $\phi(2I) = I$, so $\phi(2I) \neq 2\phi(I)$.

Transposition is linear

Transposition is linear

Example ??: We can define a map $\mathrm{trans}\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\mathrm{trans}(A) = A^T$.

Transposition is linear

Example ??: We can define a map trans: $M_n\mathbb{R} \to M_n\mathbb{R}$ by $\text{trans}(A) = A^T$.

Here as usual, A^T is the transpose of A, which is obtained by flipping A across the main diagonal.

Transposition is linear

Example ??: We can define a map trans: $M_n\mathbb{R} \to M_n\mathbb{R}$ by $\text{trans}(A) = A^T$.

Here as usual, A^T is the transpose of A, which is obtained by flipping A across the main diagonal.

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}.$$

Transposition is linear

Example ??: We can define a map trans: $M_n\mathbb{R} \to M_n\mathbb{R}$ by $\mathrm{trans}(A) = A^T$.

Here as usual, A^T is the transpose of A, which is obtained by flipping A across the main diagonal.

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}.$$

In general, we have $(A^T)_{ij} = A_{ji}$.

Transposition is linear

Example ??: We can define a map trans: $M_n\mathbb{R} \to M_n\mathbb{R}$ by $\text{trans}(A) = A^T$.

Here as usual, A^T is the transpose of A, which is obtained by flipping A across the main diagonal.

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}.$$

In general, we have $(A^T)_{ij} = A_{ji}$.

It is clear that $(A+B)^T = A^T + B^T$ and $(tA)^T = tA^T$

Transposition is linear

Example ??: We can define a map trans: $M_n\mathbb{R} \to M_n\mathbb{R}$ by $\text{trans}(A) = A^T$.

Here as usual, A^T is the transpose of A, which is obtained by flipping A across the main diagonal.

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}.$$

In general, we have $(A^T)_{ij} = A_{ji}$.

It is clear that $(A + B)^T = A^T + B^T$ and $(tA)^T = tA^T$,

so trans: $M_n\mathbb{R} \to M_n\mathbb{R}$ is a linear map.

e.g. $\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right)^T = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}^T = \begin{bmatrix} a+a' & c+c' \\ b+b' & d+d' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^T + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}^T$

Isomorphisms

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1}\colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

Isomorphisms

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1}\colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

(ϕ^{-1} is automatically a *linear* map - we leave this as an exercise.)

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1}\colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

(ϕ^{-1} is automatically a *linear* map - we leave this as an exercise.)

Say that V and W are *isomorphic* if there is an isomorphism from V to W.

Isomorphisms

Definition ??:
A linear map $\phi \colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1} \colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

(ϕ^{-1} is automatically a *linear* map - we leave this as an exercise.)

Say that V and W are *isomorphic* if there is an isomorphism from V to W.

Example ??: We can now rephrase part of Example **??** as follows:

Isomorphisms

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1}\colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

(ϕ^{-1} is automatically a *linear* map - we leave this as an exercise.)

Say that V and W are *isomorphic* if there is an isomorphism from V to W.

Example ??: We can now rephrase part of Example **??** as follows:

There is an isomorphism $\phi\colon M_2\mathbb{R} \to \mathbb{R}^4$ given by

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1}\colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

(ϕ^{-1} is automatically a *linear* map - we leave this as an exercise.)

Say that V and W are *isomorphic* if there is an isomorphism from V to W.

Example ??: We can now rephrase part of Example **??** as follows:

There is an isomorphism $\phi\colon M_2\mathbb{R} \to \mathbb{R}^4$ given by

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

so $M_2\mathbb{R}$ is isomorphic to $\mathbb{R}^4$.

Definition ??:
A linear map $\phi\colon V \to W$ is an *isomorphism* if it is a bijection,

so there is an inverse map $\phi^{-1}\colon W \to V$ with $\phi(\phi^{-1}(w)) = w$ for all $w \in W$, and $\phi^{-1}(\phi(v)) = v$ for all $v \in V$.

(ϕ^{-1} is automatically a *linear* map - we leave this as an exercise.)

Say that V and W are *isomorphic* if there is an isomorphism from V to W.

Example ??: We can now rephrase part of Example **??** as follows:

There is an isomorphism $\phi\colon M_2\mathbb{R} \to \mathbb{R}^4$ given by

$$\phi \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

so $M_2\mathbb{R}$ is isomorphic to $\mathbb{R}^4$.

Similarly, the space $M_{p,q}\mathbb{R}$ is isomorphic to $\mathbb{R}^{pq}$.

Physical vectors

Example ??: Let U be the space of physical vectors, as in Example **??**. A choice of axes and length units gives rise to an isomorphism from $\mathbb{R}^3$ to U.

Example ??: Let U be the space of physical vectors, as in Example **??**. A choice of axes and length units gives rise to an isomorphism from $\mathbb{R}^3$ to U.

More explicitly, choose a point P on the surface of the earth (for example, the base of the Eiffel Tower) and put

$$\begin{aligned}\mathbf{u} &= \text{ the vector of length 1 km pointing east from } P\\ \mathbf{v} &= \text{ the vector of length 1 km pointing north from } P\\ \mathbf{w} &= \text{ the vector of length 1 km pointing vertically upwards from } P.\end{aligned}$$

Example ??: Let U be the space of physical vectors, as in Example **??**. A choice of axes and length units gives rise to an isomorphism from $\mathbb{R}^3$ to U.

More explicitly, choose a point P on the surface of the earth (for example, the base of the Eiffel Tower) and put

$$\begin{aligned} \mathbf{u} &= \text{ the vector of length 1 km pointing east from } P \\ \mathbf{v} &= \text{ the vector of length 1 km pointing north from } P \\ \mathbf{w} &= \text{ the vector of length 1 km pointing vertically upwards from } P. \end{aligned}$$

Define $\phi\colon \mathbb{R}^3 \to U$ by $\phi(x, y, z) = x\mathbf{u} + y\mathbf{v} + z\mathbf{w}$. Then ϕ is an isomorphism.

Example ??: Let U be the space of physical vectors, as in Example **??**. A choice of axes and length units gives rise to an isomorphism from $\mathbb{R}^3$ to U.

More explicitly, choose a point P on the surface of the earth (for example, the base of the Eiffel Tower) and put

$$\begin{aligned}\mathbf{u} &= \text{ the vector of length 1 km pointing east from } P\\ \mathbf{v} &= \text{ the vector of length 1 km pointing north from } P\\ \mathbf{w} &= \text{ the vector of length 1 km pointing vertically upwards from } P.\end{aligned}$$

Define $\phi\colon \mathbb{R}^3 \to U$ by $\phi(x, y, z) = x\mathbf{u} + y\mathbf{v} + z\mathbf{w}$. Then ϕ is an isomorphism.

We will be able to give more interesting examples of isomorphisms after we have learnt about subspaces. ◯

Subspaces

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

Subspaces

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W.

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W. (In other words, W is closed under addition.)

Subspaces

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W. (In other words, W is closed under addition.)

(c) Whenever u lies in W and t lies in $\mathbb{R}$, the element tu also lies in W.

Subspaces

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W. (In other words, W is closed under addition.)

(c) Whenever u lies in W and t lies in $\mathbb{R}$, the element tu also lies in W. (In other words, W is closed under scalar multiplication.)

Subspaces

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W. (In other words, W is closed under addition.)

(c) Whenever u lies in W and t lies in $\mathbb{R}$, the element tu also lies in W. (In other words, W is closed under scalar multiplication.)

These conditions mean that W is itself a vector space.

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W. (In other words, W is closed under addition.)

(c) Whenever u lies in W and t lies in $\mathbb{R}$, the element tu also lies in W. (In other words, W is closed under scalar multiplication.)

These conditions mean that W is itself a vector space.

Remark ??: Strictly speaking, a vector space is a set *together with a definition of addition and scalar multiplication* such that certain identities hold.

Definition ??: Let V be a vector space. A *vector subspace* (or just *subspace*) of V is a subset $W \subseteq V$ such that

(a) $0 \in W$

(b) Whenever u and v lie in W, the element $u + v$ also lies in W. (In other words, W is closed under addition.)

(c) Whenever u lies in W and t lies in $\mathbb{R}$, the element tu also lies in W. (In other words, W is closed under scalar multiplication.)

These conditions mean that W is itself a vector space.

Remark ??: Strictly speaking, a vector space is a set *together with a definition of addition and scalar multiplication* such that certain identities hold.

We should therefore specify that addition in W is to be defined using the same rule as for V, and similarly for scalar multiplication. $\bigcirc$

Reformulation

Remark ??: W is a subspace iff (a) $0 \in W$

(b) Whenever $u, v \in W$, the element $u + v$ also lies in W.

(c) Whenever $u \in W$ and $t \in \mathbb{R}$, the element tu also lies in W.

Reformulation

Remark ??: W is a subspace iff (a) $0 \in W$

(b) Whenever $u, v \in W$, the element $u + v$ also lies in W.

(c) Whenever $u \in W$ and $t \in \mathbb{R}$, the element tu also lies in W.

Reformulation: a subset $W \subseteq V$ is a subspace iff (a) $0 \in W$ and

(d) Whenever $u, v \in W$ and $t, s \in \mathbb{R}$ we have $tu + sv \in W$.

Remark ??: W is a subspace iff (a) $0 \in W$

(b) Whenever $u, v \in W$, the element $u + v$ also lies in W.

(c) Whenever $u \in W$ and $t \in \mathbb{R}$, the element tu also lies in W.

Reformulation: a subset $W \subseteq V$ is a subspace iff (a) $0 \in W$ and

(d) Whenever $u, v \in W$ and $t, s \in \mathbb{R}$ we have $tu + sv \in W$.

To show that this reformulation is valid, we must check that if condition (d) holds then so do (b) and (c);

Reformulation

Remark ??: W is a subspace iff (a) $0 \in W$

(b) Whenever $u, v \in W$, the element $u + v$ also lies in W.

(c) Whenever $u \in W$ and $t \in \mathbb{R}$, the element tu also lies in W.

Reformulation: a subset $W \subseteq V$ is a subspace iff (a) $0 \in W$ and

(d) Whenever $u, v \in W$ and $t, s \in \mathbb{R}$ we have $tu + sv \in W$.

To show that this reformulation is valid, we must check that if condition (d) holds then so do (b) and (c); and that if (b) and (c) hold then so does (d).

Remark ??: W is a subspace iff (a) $0 \in W$

(b) Whenever $u, v \in W$, the element $u + v$ also lies in W.

(c) Whenever $u \in W$ and $t \in \mathbb{R}$, the element tu also lies in W.

Reformulation: a subset $W \subseteq V$ is a subspace iff (a) $0 \in W$ and

(d) Whenever $u, v \in W$ and $t, s \in \mathbb{R}$ we have $tu + sv \in W$.

To show that this reformulation is valid, we must check that if condition (d) holds then so do (b) and (c); and that if (b) and (c) hold then so does (d).

In fact, conditions (b) is the special cases of (d) where $t = s = 1$, and condition (c) is the special case of (d) where $v = 0$; so if (d) holds then so do (b) and (c).

Remark ??: W is a subspace iff (a) $0 \in W$

(b) Whenever $u, v \in W$, the element $u + v$ also lies in W.

(c) Whenever $u \in W$ and $t \in \mathbb{R}$, the element tu also lies in W.

Reformulation: a subset $W \subseteq V$ is a subspace iff (a) $0 \in W$ and

(d) Whenever $u, v \in W$ and $t, s \in \mathbb{R}$ we have $tu + sv \in W$.

To show that this reformulation is valid, we must check that if condition (d) holds then so do (b) and (c); and that if (b) and (c) hold then so does (d).

In fact, conditions (b) is the special cases of (d) where $t = s = 1$, and condition (c) is the special case of (d) where $v = 0$; so if (d) holds then so do (b) and (c).

Conversely, suppose that (b) and (c) hold, and that $u, v \in W$ and $t, s \in \mathbb{R}$. Then condition (c) tells us that $tu \in W$, and similarly that $sv \in W$. Given these, condition (b) tells us that $tu + sv \in W$; we conclude that condition (d) holds, as required. $\bigcirc$

Examples of subspaces

Example ??: For any vector space V, there are two silly examplesof subspaces of V: $\{0\}$ is always a subspace of V, and V itself is always a subspace of V.

Example ??: For any vector space V, there are two silly examplesof subspaces of V: $\{0\}$ is always a subspace of V, and V itself is always a subspace of V.

Example ??: Any straight line through the origin is a subspace of $\mathbb{R}^2$. These are the only subspaces of $\mathbb{R}^2$ (except for the two silly examples).

Examples of subspaces

Example ??: For any vector space V, there are two silly examplesof subspaces of V: $\{0\}$ is always a subspace of V, and V itself is always a subspace of V.

Example ??: Any straight line through the origin is a subspace of $\mathbb{R}^2$. These are the only subspaces of $\mathbb{R}^2$ (except for the two silly examples).

Example ??: In $\mathbb{R}^3$, any straight line through the origin is a subspace, and any plane through the origin is also a subspace. These are the only subspaces of $\mathbb{R}^3$ (except for the two silly examples).

Trace-free matrices

Trace-free matrices

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \text{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

Trace-free matrices

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \text{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

To check this, we first note that $0 \in W$.

Trace-free matrices

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

To check this, we first note that $0 \in W$. Suppose that $A, A' \in W$ and $t, t' \in \mathbb{R}$.

Trace-free matrices

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \text{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

To check this, we first note that $0 \in W$. Suppose that $A, A' \in W$ and $t, t' \in \mathbb{R}$. We then have $\text{trace}(A) = \text{trace}(A') = 0$ (because $A, A' \in W$)

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \text{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

To check this, we first note that $0 \in W$. Suppose that $A, A' \in W$ and $t, t' \in \mathbb{R}$. We then have $\text{trace}(A) = \text{trace}(A') = 0$ (because $A, A' \in W$) and so

$$\text{trace}(tA + t'A') = t\,\text{trace}(A) + t'\,\text{trace}(A') = t.0 + t'.0 = 0,$$

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \text{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

To check this, we first note that $0 \in W$. Suppose that $A, A' \in W$ and $t, t' \in \mathbb{R}$. We then have $\text{trace}(A) = \text{trace}(A') = 0$ (because $A, A' \in W$) and so

$$\text{trace}(tA + t'A') = t\,\text{trace}(A) + t'\,\text{trace}(A') = t.0 + t'.0 = 0,$$

so $tA + t'A' \in W$.

Example ??: The set $W = \{A \in M_2\mathbb{R} \mid \mathrm{trace}(A) = 0\}$ is a subspace of $M_2\mathbb{R}$.

To check this, we first note that $0 \in W$. Suppose that $A, A' \in W$ and $t, t' \in \mathbb{R}$. We then have $\mathrm{trace}(A) = \mathrm{trace}(A') = 0$ (because $A, A' \in W$) and so

$$\mathrm{trace}(tA + t'A') = t\,\mathrm{trace}(A) + t'\,\mathrm{trace}(A') = t.0 + t'.0 = 0,$$

so $tA + t'A' \in W$.

Thus, conditions (a) and (d) in Remark **??** are satisfied, showing that W is a subspace as claimed. $\bigcirc$

Spaces of polynomials

Example ??: Recall that $\mathbb{R}[x]$ is the set of all polynomial functions of x

Spaces of polynomials

Example ??: Recall that $\mathbb{R}[x]$ is the set of all polynomial functions of x

(so the functions $p(x) = x + 1$ and $q(x) = (x+1)^5 - (x-1)^5$ and $r(x) = 1 + 4x^4 + 8x^8$ define elements $p, q, r \in \mathbb{R}[x]$).

Example ??: Recall that $\mathbb{R}[x]$ is the set of all polynomial functions of x

(so the functions $p(x) = x + 1$ and $q(x) = (x+1)^5 - (x-1)^5$ and $r(x) = 1 + 4x^4 + 8x^8$ define elements $p, q, r \in \mathbb{R}[x]$).

It is clear that the sum of two polynomials is another polynomial, and any polynomial multiplied by a constant is also a polynomial, so $\mathbb{R}[x]$ is a subspace of the vector space $F(\mathbb{R})$ of all functions on $\mathbb{R}$.

Example ??: Recall that $\mathbb{R}[x]$ is the set of all polynomial functions of x

(so the functions $p(x) = x + 1$ and $q(x) = (x + 1)^5 - (x - 1)^5$ and $r(x) = 1 + 4x^4 + 8x^8$ define elements $p, q, r \in \mathbb{R}[x]$).

It is clear that the sum of two polynomials is another polynomial, and any polynomial multiplied by a constant is also a polynomial, so $\mathbb{R}[x]$ is a subspace of the vector space $F(\mathbb{R})$ of all functions on $\mathbb{R}$.

We write $\mathbb{R}[x]_{\leq d}$ for the set of polynomials of degree at most d

Example ??: Recall that $\mathbb{R}[x]$ is the set of all polynomial functions of x

(so the functions $p(x) = x + 1$ and $q(x) = (x + 1)^5 - (x - 1)^5$ and $r(x) = 1 + 4x^4 + 8x^8$ define elements $p, q, r \in \mathbb{R}[x]$).

It is clear that the sum of two polynomials is another polynomial, and any polynomial multiplied by a constant is also a polynomial, so $\mathbb{R}[x]$ is a subspace of the vector space $F(\mathbb{R})$ of all functions on $\mathbb{R}$.

We write $\mathbb{R}[x]_{\leq d}$ for the set of polynomials of degree at most d, so a general element $f \in \mathbb{R}[x]_{\leq d}$ has the form

$$f(x) = a_0 + a_1 x + \ldots + a_d x^d = \sum_{i=0}^{d} a_i x^i$$

for some $a_0, \ldots, a_d \in \mathbb{R}$. It is easy to see that this is a subspace of $\mathbb{R}[x]$.

Spaces of polynomials

Example ??: Recall that $\mathbb{R}[x]$ is the set of all polynomial functions of x

(so the functions $p(x) = x + 1$ and $q(x) = (x+1)^5 - (x-1)^5$ and $r(x) = 1 + 4x^4 + 8x^8$ define elements $p, q, r \in \mathbb{R}[x]$).

It is clear that the sum of two polynomials is another polynomial, and any polynomial multiplied by a constant is also a polynomial, so $\mathbb{R}[x]$ is a subspace of the vector space $F(\mathbb{R})$ of all functions on $\mathbb{R}$.

We write $\mathbb{R}[x]_{\leq d}$ for the set of polynomials of degree at most d, so a general element $f \in \mathbb{R}[x]_{\leq d}$ has the form

$$f(x) = a_0 + a_1 x + \ldots + a_d x^d = \sum_{i=0}^{d} a_i x^i$$

for some $a_0, \ldots, a_d \in \mathbb{R}$. It is easy to see that this is a subspace of $\mathbb{R}[x]$.

If we let f correspond to the vector $\begin{bmatrix} a_0 & \cdots & a_d \end{bmatrix}^T \in \mathbb{R}^{d+1}$, we get a one-to-one correspondence between $\mathbb{R}[x]_{\leq d}$ and $\mathbb{R}^{d+1}$. ○

More precisely, there is an isomorphism $\phi\colon \mathbb{R}^{d+1} \to \mathbb{R}[x]_{\leq d}$ given by

$$\phi\left(\begin{bmatrix} a_0 \\ \vdots \\ a_d \end{bmatrix}\right) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d = \sum_{i=0}^{d} a_i x^i.$$

More precisely, there is an isomorphism $\phi\colon \mathbb{R}^{d+1} \to \mathbb{R}[x]_{\leq d}$ given by

$$\phi\left(\begin{bmatrix} a_0 \\ \vdots \\ a_d \end{bmatrix}\right) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d = \sum_{i=0}^{d} a_i x^i.$$

Remark ??: It is a common mistake to think that $\mathbb{R}[x]_{\leq d}$ is isomorphic to $\mathbb{R}^d$ (rather than $\mathbb{R}^{d+1}$), but this is not correct.

More precisely, there is an isomorphism $\phi\colon \mathbb{R}^{d+1} \to \mathbb{R}[x]_{\leq d}$ given by

$$\phi\left(\begin{bmatrix} a_0 \\ \vdots \\ a_d \end{bmatrix}\right) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d = \sum_{i=0}^{d} a_i x^i.$$

Remark ??: It is a common mistake to think that $\mathbb{R}[x]_{\leq d}$ is isomorphic to $\mathbb{R}^d$ (rather than $\mathbb{R}^{d+1}$), but this is not correct.

Note that the list $0, 1, 2, 3$ has four entries (not three), and similarly, the list $0, 1, 2, \ldots, d$ has $d + 1$ entries (not d).

Even and odd functions

Even and odd functions

Example ??: A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

Even and odd functions

Example ??: A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

eg $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, so cos is even and sin is odd.

Even and odd functions

Example ??: A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

eg $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, so cos is even and sin is odd.

(Of course, most functions are neither even nor odd.)

Example ??: A function $f: \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

eg $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, so cos is even and sin is odd.

(Of course, most functions are neither even nor odd.)

We write EF for the set of even functions, so EF is a subset of the set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$, and $\cos \in EF$.

Even and odd functions

Example ??: A function $f\colon \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

eg $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, so cos is even and sin is odd.

(Of course, most functions are neither even nor odd.)

We write EF for the set of even functions, so EF is a subset of the set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$, and $\cos \in EF$.

If f and g are even, it is clear that $f + g$ is also even. If f is even and t is a constant, then it is clear that tf is also even; and the zero function is certainly even as well.

Example ??: A function $f\colon \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

eg $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, so cos is even and sin is odd.

(Of course, most functions are neither even nor odd.)

We write EF for the set of even functions, so EF is a subset of the set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$, and $\cos \in EF$.

If f and g are even, it is clear that $f + g$ is also even. If f is even and t is a constant, then it is clear that tf is also even; and the zero function is certainly even as well.

This shows that EF is actually a subspace of $F(\mathbb{R})$.

Example ??: A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x.

eg $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$, so cos is even and sin is odd.

(Of course, most functions are neither even nor odd.)

We write EF for the set of even functions, so EF is a subset of the set $F(\mathbb{R})$ of all functions from $\mathbb{R}$ to $\mathbb{R}$, and $\cos \in EF$.

If f and g are even, it is clear that $f + g$ is also even. If f is even and t is a constant, then it is clear that tf is also even; and the zero function is certainly even as well.

This shows that EF is actually a subspace of $F(\mathbb{R})$.

Similarly, the set OF of odd functions is a subspace of $F(\mathbb{R})$. $\bigcirc$

Solutions of differential equations

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

Example ??: Let V be the vector space of smooth functions $u(x, t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.
  This governs the flow of heat along an iron bar.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.
  This governs the flow of heat along an iron bar.
- We say that u solves the *Korteweg-de Vries Equation* if
  $\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} - 6u\frac{\partial u}{\partial x} = 0$.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.
  This governs the flow of heat along an iron bar.
- We say that u solves the *Korteweg-de Vries Equation* if
  $\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} - 6u\frac{\partial u}{\partial x} = 0$.
  This governs the propagation of large waves in shallow water.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.
  This governs the flow of heat along an iron bar.
- We say that u solves the *Korteweg-de Vries Equation* if
  $\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} - 6u\frac{\partial u}{\partial x} = 0$.
  This governs the propagation of large waves in shallow water.

The set of solutions of the Wave Equation is a subspace of V, as is the set of solutions to the Heat Equation.

Example ??: Let V be the vector space of smooth functions $u(x,t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.
  This governs the flow of heat along an iron bar.
- We say that u solves the *Korteweg-de Vries Equation* if $\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} - 6u\frac{\partial u}{\partial x} = 0$.
  This governs the propagation of large waves in shallow water.

The set of solutions of the Wave Equation is a subspace of V, as is the set of solutions to the Heat Equation.

However, the sum of two solutions to the KdV equation does not satisfy the KdV equation, so the set of solutions is not a subspace of V.

Example ??: Let V be the vector space of smooth functions $u(x, t)$ in two variables x and t (to be thought of as position and time).

- We say that u solves the *Wave Equation* if $\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$.
  This equation governs the propagation of small waves in deep water, or of electromagnetic waves in empty space.
- We say that u solves the *Heat Equation* if $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$.
  This governs the flow of heat along an iron bar.
- We say that u solves the *Korteweg-de Vries Equation* if $\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} - 6u\frac{\partial u}{\partial x} = 0$.
  This governs the propagation of large waves in shallow water.

The set of solutions of the Wave Equation is a subspace of V, as is the set of solutions to the Heat Equation.

However, the sum of two solutions to the KdV equation does not satisfy the KdV equation, so the set of solutions is not a subspace of V.

The Wave and Heat equations are *linear*, but the KdV equation is not. ◯

Solutions of differential equations

Solutions of differential equations

The distinction between linear and nonlinear differential equations is of fundamental importance in physics.

Solutions of differential equations

The distinction between linear and nonlinear differential equations is of fundamental importance in physics.

Linear equations can generally be solved analytically, or by efficient computer algorithms, but nonlinear equations require far more computing power.

Solutions of differential equations

The distinction between linear and nonlinear differential equations is of fundamental importance in physics.

Linear equations can generally be solved analytically, or by efficient computer algorithms, but nonlinear equations require far more computing power.

The equations of electromagnetism are linear, which explains why hundreds of different radio, TV and mobile phone channels can coexist, together with visible light (which is also a form of electromagnetic radiation), with little or no interference.

Solutions of differential equations

The distinction between linear and nonlinear differential equations is of fundamental importance in physics.

Linear equations can generally be solved analytically, or by efficient computer algorithms, but nonlinear equations require far more computing power.

The equations of electromagnetism are linear, which explains why hundreds of different radio, TV and mobile phone channels can coexist, together with visible light (which is also a form of electromagnetic radiation), with little or no interference.

The motion of fluids and gasses is governed by the Navier-Stokes equation, which is nonlinear; because of this, massive supercomputers are needed for weather forecasting, climate modelling, and aircraft design.

Example ??: Consider the following sets of 3×3 matrices:

Example ??: Consider the following sets of 3×3 matrices:

$U_0 = \{$ symmetric matrices $\}$ $= \{A \in M_3\mathbb{R} \mid A^T = A\}$

Example ??: Consider the following sets of 3×3 matrices:

$$U_0 = \{ \text{ symmetric matrices } \} = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

$$U_1 = \{ \text{ antisymmetric matrices } \} = \{A \in M_3\mathbb{R} \mid A^T = -A\}$$

Subspaces of matrices

Example ??: Consider the following sets of 3×3 matrices:

$$U_0 = \{ \text{ symmetric matrices } \} = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

$$U_1 = \{ \text{ antisymmetric matrices } \} = \{A \in M_3\mathbb{R} \mid A^T = -A\}$$

$$U_2 = \{ \text{ trace-free matrices } \} = \{A \in M_3\mathbb{R} \mid \operatorname{trace}(A) = 0\}$$

Subspaces of matrices

Example ??: Consider the following sets of 3×3 matrices:

$$
\begin{aligned}
U_0 &= \{\text{ symmetric matrices }\} && = \{A \in M_3\mathbb{R} \mid A^T = A\} \\
U_1 &= \{\text{ antisymmetric matrices }\} && = \{A \in M_3\mathbb{R} \mid A^T = -A\} \\
U_2 &= \{\text{ trace-free matrices }\} && = \{A \in M_3\mathbb{R} \mid \operatorname{trace}(A) = 0\} \\
U_3 &= \{\text{ diagonal matrices }\} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \neq j\}
\end{aligned}
$$

Subspaces of matrices

Example ??: Consider the following sets of 3×3 matrices:

$$
\begin{aligned}
U_0 &= \{ \text{ symmetric matrices } \} & &= \{A \in M_3\mathbb{R} \mid A^T = A\} \\
U_1 &= \{ \text{ antisymmetric matrices } \} & &= \{A \in M_3\mathbb{R} \mid A^T = -A\} \\
U_2 &= \{ \text{ trace-free matrices } \} & &= \{A \in M_3\mathbb{R} \mid \operatorname{trace}(A) = 0\} \\
U_3 &= \{ \text{ diagonal matrices } \} & &= \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \neq j\} \\
U_4 &= \{ \text{ strictly upper-triangular matrices } \} & &= \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}
\end{aligned}
$$

Subspaces of matrices

Example ??: Consider the following sets of 3×3 matrices:

$$
\begin{aligned}
U_0 &= \{ \text{ symmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = A\} \\
U_1 &= \{ \text{ antisymmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = -A\} \\
U_2 &= \{ \text{ trace-free matrices } \} && = \{A \in M_3\mathbb{R} \mid \operatorname{trace}(A) = 0\} \\
U_3 &= \{ \text{ diagonal matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \neq j\} \\
U_4 &= \{ \text{ strictly upper-triangular matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\} \\
U_5 &= \{ \text{ invertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\}
\end{aligned}
$$

Example ??: Consider the following sets of 3×3 matrices:

$$
\begin{aligned}
U_0 &= \{ \text{ symmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = A\} \\
U_1 &= \{ \text{ antisymmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = -A\} \\
U_2 &= \{ \text{ trace-free matrices } \} && = \{A \in M_3\mathbb{R} \mid \operatorname{trace}(A) = 0\} \\
U_3 &= \{ \text{ diagonal matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \neq j\} \\
U_4 &= \{ \text{ strictly upper-triangular matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\} \\
U_5 &= \{ \text{ invertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\} \\
U_6 &= \{ \text{ noninvertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) = 0\}
\end{aligned}
$$

Subspaces of matrices

Example ??: Consider the following sets of 3×3 matrices:

$$
\begin{aligned}
U_0 &= \{ \text{ symmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = A\} \\
U_1 &= \{ \text{ antisymmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = -A\} \\
U_2 &= \{ \text{ trace-free matrices } \} && = \{A \in M_3\mathbb{R} \mid \text{trace}(A) = 0\} \\
U_3 &= \{ \text{ diagonal matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \neq j\} \\
U_4 &= \{ \text{ strictly upper-triangular matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\} \\
U_5 &= \{ \text{ invertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\} \\
U_6 &= \{ \text{ noninvertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) = 0\}
\end{aligned}
$$

Then $U_0, \dots, U_4$ are all subspaces of $M_3\mathbb{R}$.

Subspaces of matrices

Example ??: Consider the following sets of 3×3 matrices:

$$
\begin{aligned}
U_0 &= \{ \text{ symmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = A\} \\
U_1 &= \{ \text{ antisymmetric matrices } \} && = \{A \in M_3\mathbb{R} \mid A^T = -A\} \\
U_2 &= \{ \text{ trace-free matrices } \} && = \{A \in M_3\mathbb{R} \mid \text{trace}(A) = 0\} \\
U_3 &= \{ \text{ diagonal matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \neq j\} \\
U_4 &= \{ \text{ strictly upper-triangular matrices } \} && = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\} \\
U_5 &= \{ \text{ invertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\} \\
U_6 &= \{ \text{ noninvertible matrices } \} && = \{A \in M_3\mathbb{R} \mid \det(A) = 0\}
\end{aligned}
$$

Then $U_0, \dots, U_4$ are all subspaces of $M_3\mathbb{R}$.

We will prove this for U_0 and U_4; the other cases are similar.

$$U_0 = \{ \text{ symmetric matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

Subspaces of matrices

$$U_0 = \{ \text{ symmetric matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

It is clear that $0^T = 0$, so $0 \in U_0$.

Subspaces of matrices

$$U_0 = \{ \text{ symmetric matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

It is clear that $0^T = 0$, so $0 \in U_0$.
Suppose that $A, B \in U_0$ (so $A^T = A$ and $B^T = B$) and $s, t \in \mathbb{R}$.

Subspaces of matrices

$$U_0 = \{ \text{ symmetric matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

It is clear that $0^T = 0$, so $0 \in U_0$.
Suppose that $A, B \in U_0$ (so $A^T = A$ and $B^T = B$) and $s, t \in \mathbb{R}$. Then

$$(sA + tB)^T = sA^T + tB^T = sA + tB$$

$$U_0 = \{ \text{ symmetric matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

It is clear that $0^T = 0$, so $0 \in U_0$.
Suppose that $A, B \in U_0$ (so $A^T = A$ and $B^T = B$) and $s, t \in \mathbb{R}$. Then

$$(sA + tB)^T = sA^T + tB^T = sA + tB$$

so $sA + tB \in U_0$.

$$U_0 = \{ \text{ symmetric matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid A^T = A\}$$

It is clear that $0^T = 0$, so $0 \in U_0$.
Suppose that $A, B \in U_0$ (so $A^T = A$ and $B^T = B$) and $s, t \in \mathbb{R}$. Then

$$(sA + tB)^T = sA^T + tB^T = sA + tB$$

so $sA + tB \in U_0$.

So U_0 is a subspace. $\bigcirc$

$$U_4 = \{\text{ strictly upper-triangular matrices }\} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Subspaces of matrices

$U_4 = \{$ strictly upper-triangular matrices $\} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0$ when $i \geq j\}$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

Subspaces of matrices

$U_4 = \{$ strictly upper-triangular matrices $\} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

$$sA + tB$$

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

$$sA + tB = s \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & b_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

$$sA + tB = s \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & b_{23} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & sa_{12}+tb_{12} & sa_{13}+tb_{13} \\ 0 & 0 & sa_{23}+tb_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

$$sA + tB = s\begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix} + t\begin{bmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & b_{23} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & sa_{12}+tb_{12} & sa_{13}+tb_{13} \\ 0 & 0 & sa_{23}+tb_{23} \\ 0 & 0 & 0 \end{bmatrix},$$

which shows that $sA + tB$ is again strictly upper triangular

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

$$sA + tB = s\begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix} + t\begin{bmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & b_{23} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & sa_{12}+tb_{12} & sa_{13}+tb_{13} \\ 0 & 0 & sa_{23}+tb_{23} \\ 0 & 0 & 0 \end{bmatrix},$$

which shows that $sA + tB$ is again strictly upper triangular, and so is an element of U_4.

Subspaces of matrices

$$U_4 = \{ \text{ strictly upper-triangular matrices } \} = \{A \in M_3\mathbb{R} \mid A_{ij} = 0 \text{ when } i \geq j\}$$

The elements of U_4 are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

The zero matrix is an element of U_4 (with $a_{12} = a_{13} = a_{23} = 0$).

Suppose that $A, B \in U_4$ and $s, t \in \mathbb{R}$.

$$sA + tB = s \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & b_{23} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & sa_{12}+tb_{12} & sa_{13}+tb_{13} \\ 0 & 0 & sa_{23}+tb_{23} \\ 0 & 0 & 0 \end{bmatrix},$$

which shows that $sA + tB$ is again strictly upper triangular, and so is an element of U_4.

Thus U_4 is also a subspace.

Subspaces of matrices

$$U_5 = \{\text{ invertible matrices }\} = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\}$$

$$U_6 = \{\text{ noninvertible matrices }\} = \{A \in M_3\mathbb{R} \mid \det(A) = 0\}$$

$$U_5 = \{ \text{ invertible matrices } \} = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\}$$
$$U_6 = \{ \text{ noninvertible matrices } \} = \{A \in M_3\mathbb{R} \mid \det(A) = 0\}$$

U_5 is not a subspace, because it does not contain the zero matrix.

Subspaces of matrices

$$U_5 = \{ \text{ invertible matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid \det(A) \neq 0\}$$
$$U_6 = \{ \text{ noninvertible matrices } \} \qquad = \{A \in M_3\mathbb{R} \mid \det(A) = 0\}$$

U_5 is not a subspace, because it does not contain the zero matrix.

U_6 is not a subspace: if we put

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

then $A, B \in U_6$ but $A + B = I \notin U_6$.

Sums of subspaces

Definition ??: Let U be a vector space, and let V and W be subspaces of U. We put

$$V + W = \{u \in U \mid u = v + w \text{ for some } v \in V \text{ and } w \in W\}.$$

Sums of subspaces

Definition ??: Let U be a vector space, and let V and W be subspaces of U. We put

$$V + W = \{u \in U \mid u = v + w \text{ for some } v \in V \text{ and } w \in W\}.$$

Example ??: If $U = \mathbb{R}^3$ and $V = \{\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \mid x \in \mathbb{R}\}$ and $W = \{\begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \mid z \in \mathbb{R}\}$

then $V + W = \{\begin{bmatrix} x \\ 0 \\ z \end{bmatrix} \mid x, z \in \mathbb{R}\}$

Definition ??: Let U be a vector space, and let V and W be subspaces of U. We put

$$V + W = \{u \in U \mid u = v + w \text{ for some } v \in V \text{ and } w \in W\}.$$

Example ??: If $U = \mathbb{R}^3$ and $V = \{\left[\begin{smallmatrix} x \\ 0 \\ 0 \end{smallmatrix}\right] \mid x \in \mathbb{R}\}$ and $W = \{\left[\begin{smallmatrix} 0 \\ 0 \\ z \end{smallmatrix}\right] \mid z \in \mathbb{R}\}$

then $V + W = \{\left[\begin{smallmatrix} x \\ 0 \\ z \end{smallmatrix}\right] \mid x, z \in \mathbb{R}\}$

Example ??: If $U = M_2\mathbb{R}$ and

$$V = \{\left[\begin{smallmatrix} a & b \\ 0 & 0 \end{smallmatrix}\right] \mid a, b \in \mathbb{R}\} \qquad W = \{\left[\begin{smallmatrix} 0 & b \\ 0 & d \end{smallmatrix}\right] \mid b, d \in \mathbb{R}\}$$

then

$$V + W = \{\left[\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right] \mid a, b, d \in \mathbb{R}\}$$

Intersections of subspaces

Intersections of subspaces

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$:

Intersections of subspaces

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$: As V is a subspace we have $0 \in V$, and as W is a subspace we have $0 \in W$, so $0 \in V \cap W$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$: As V is a subspace we have $0 \in V$, and as W is a subspace we have $0 \in W$, so $0 \in V \cap W$.

Next, suppose we have $x, y \in V \cap W$ and $s, t \in \mathbb{R}$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$: As V is a subspace we have $0 \in V$, and as W is a subspace we have $0 \in W$, so $0 \in V \cap W$.

Next, suppose we have $x, y \in V \cap W$ and $s, t \in \mathbb{R}$. Then $x, y \in V$ and V is a subspace, so $sx + ty \in V$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$: As V is a subspace we have $0 \in V$, and as W is a subspace we have $0 \in W$, so $0 \in V \cap W$.

Next, suppose we have $x, y \in V \cap W$ and $s, t \in \mathbb{R}$. Then $x, y \in V$ and V is a subspace, so $sx + ty \in V$. Similarly, we have $x, y \in W$ and W is a subspace so $sx + ty \in W$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$: As V is a subspace we have $0 \in V$, and as W is a subspace we have $0 \in W$, so $0 \in V \cap W$.

Next, suppose we have $x, y \in V \cap W$ and $s, t \in \mathbb{R}$. Then $x, y \in V$ and V is a subspace, so $sx + ty \in V$. Similarly, we have $x, y \in W$ and W is a subspace so $sx + ty \in W$. This shows that $sx + ty \in V \cap W$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V \cap W$: As V is a subspace we have $0 \in V$, and as W is a subspace we have $0 \in W$, so $0 \in V \cap W$.

Next, suppose we have $x, y \in V \cap W$ and $s, t \in \mathbb{R}$. Then $x, y \in V$ and V is a subspace, so $sx + ty \in V$. Similarly, we have $x, y \in W$ and W is a subspace so $sx + ty \in W$. This shows that $sx + ty \in V \cap W$.

This works for all x, y, s and t, so $V \cap W$ is a subspace. $\bigcirc$

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:

Sums of subspaces

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.
We then have $tv + t'v' \in V$ (because V is a subspace)

Sums of subspaces

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.
We then have $tv + t'v' \in V$ (because V is a subspace)
and $tw + t'w' \in W$ (because W is a subspace).

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.
We then have $tv + t'v' \in V$ (because V is a subspace)
and $tw + t'w' \in W$ (because W is a subspace).

We also have

$$tx + t'x' = t(v + w) + t'(v' + w') = (tv + t'v') + (tw + t'w')$$

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.
We then have $tv + t'v' \in V$ (because V is a subspace)
and $tw + t'w' \in W$ (because W is a subspace).

We also have

$$tx + t'x' = t(v + w) + t'(v' + w') = (tv + t'v') + (tw + t'w')$$

with $tv + t'v' \in V$ and $tw + t'w' \in W$

Sums of subspaces

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.
We then have $tv + t'v' \in V$ (because V is a subspace)
and $tw + t'w' \in W$ (because W is a subspace).

We also have

$$tx + t'x' = t(v + w) + t'(v' + w') = (tv + t'v') + (tw + t'w')$$

with $tv + t'v' \in V$ and $tw + t'w' \in W$, so $tx + t'x' \in V + W$.

Sums of subspaces

Proposition ??: Let U be a vector space, and let V and W be subspaces of U. Then both $V \cap W$ and $V + W$ are subspaces of U.

Proof for $V + W$:
We can write 0 as $0 + 0$ with $0 \in V$ and $0 \in W$, so $0 \in V + W$.

Now suppose we have $x, x' \in V + W$ and $t, t' \in \mathbb{R}$.
As $x \in V + W$ we can find $v \in V$ and $w \in W$ such that $x = v + w$.
As $x' \in V + W$ we can also find $v' \in V$ and $w' \in W$ such that $x' = v' + w'$.
We then have $tv + t'v' \in V$ (because V is a subspace)
and $tw + t'w' \in W$ (because W is a subspace).

We also have

$$tx + t'x' = t(v + w) + t'(v' + w') = (tv + t'v') + (tw + t'w')$$

with $tv + t'v' \in V$ and $tw + t'w' \in W$, so $tx + t'x' \in V + W$.

As this works for all x, x', t and t', we conclude that $V + W$ is a subspace. ○

Two planes

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$$V = \{[x,y,z]^T \mid x + 2y + 3z = 0\} \qquad W = \{[x,y,z]^T \mid 3x + 2y + z = 0\}.$$

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x, y, z]^T \mid x + 2y + 3z = 0\}$ $\quad$ $W = \{[x, y, z]^T \mid 3x + 2y + z = 0\}$.

Claim: $V \cap W = \{[x, -2x, x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x, y, z]^T \mid x + 2y + 3z = 0\}$ $\qquad$ $W = \{[x, y, z]^T \mid 3x + 2y + z = 0\}$.

Claim: $V \cap W = \{[x, -2x, x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x, y, z]^T \in V \cap W$ iff $x + 2y + 3z = 0$ and also $3x + 2y + z = 0$.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x, y, z]^T \mid x + 2y + 3z = 0\}$ $\qquad$ $W = \{[x, y, z]^T \mid 3x + 2y + z = 0\}$.

Claim: $V \cap W = \{[x, -2x, x]^T \mid t \in \mathbb{R}\}$ $\quad$ and $\quad$ $V + W = \mathbb{R}^3$.

Indeed, $[x, y, z]^T \in V \cap W$ iff $x + 2y + 3z = 0$ and also $3x + 2y + z = 0$.
If we subtract these two equations and divide by two, we find that $z = x$.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x,y,z]^T \mid x+2y+3z=0\}$ $\qquad$ $W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$

Claim: $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3.$

Indeed, $[x,y,z]^T \in V \cap W$ iff $x+2y+3z=0$ and also $3x+2y+z=0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$$V = \{[x, y, z]^T \mid x + 2y + 3z = 0\} \qquad W = \{[x, y, z]^T \mid 3x + 2y + z = 0\}.$$

Claim: $V \cap W = \{[x, -2x, x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x, y, z]^T \in V \cap W$ iff $x + 2y + 3z = 0$ and also $3x + 2y + z = 0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.
Conversely, if $y = -2x$ and $z = x$ we see that both equations are satisfied.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$$V = \{[x,y,z]^T \mid x+2y+3z=0\} \qquad W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$$

Claim: $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x,y,z]^T \in V \cap W$ iff $x+2y+3z=0$ and also $3x+2y+z=0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.
Conversely, if $y = -2x$ and $z = x$ we see that both equations are satisfied.
It follows that $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ as claimed.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x, y, z]^T \mid x + 2y + 3z = 0\}$ $\qquad W = \{[x, y, z]^T \mid 3x + 2y + z = 0\}$.

Claim: $V \cap W = \{[x, -2x, x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x, y, z]^T \in V \cap W$ iff $x + 2y + 3z = 0$ and also $3x + 2y + z = 0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.
Conversely, if $y = -2x$ and $z = x$ we see that both equations are satisfied.
It follows that $V \cap W = \{[x, -2x, x]^T \mid t \in \mathbb{R}\}$ as claimed.

Next, consider an arbitrary vector $\mathbf{u} = [x, y, z]^T \in \mathbb{R}^3$.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x,y,z]^T \mid x+2y+3z=0\}$ $\qquad$ $W = \{[x,y,z]^T \mid 3x+2y+z=0\}$.

Claim: $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x,y,z]^T \in V \cap W$ iff $x+2y+3z=0$ and also $3x+2y+z=0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.
Conversely, if $y = -2x$ and $z = x$ we see that both equations are satisfied.
It follows that $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ as claimed.

Next, consider an arbitrary vector $\mathbf{u} = [x,y,z]^T \in \mathbb{R}^3$. Put

$\mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix}$ $\qquad$ $\mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$V = \{[x,y,z]^T \mid x + 2y + 3z = 0\}$ $\qquad$ $W = \{[x,y,z]^T \mid 3x + 2y + z = 0\}$.

Claim: $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x,y,z]^T \in V \cap W$ iff $x + 2y + 3z = 0$ and also $3x + 2y + z = 0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.
Conversely, if $y = -2x$ and $z = x$ we see that both equations are satisfied.
It follows that $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ as claimed.

Next, consider an arbitrary vector $\mathbf{u} = [x,y,z]^T \in \mathbb{R}^3$. Put

$$\mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix} \qquad \mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$$

Then $\mathbf{u} = \mathbf{v} + \mathbf{w}$ with $\mathbf{v} \in V$ and $\mathbf{w} \in W$, so $\mathbf{u} \in V + W$.

Two planes

Example ??: Take $U = \mathbb{R}^3$ and

$$V = \{[x,y,z]^T \mid x+2y+3z=0\} \qquad W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$$

Claim: $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ and $V + W = \mathbb{R}^3$.

Indeed, $[x,y,z]^T \in V \cap W$ iff $x+2y+3z=0$ and also $3x+2y+z=0$.
If we subtract these two equations and divide by two, we find that $z = x$.
If we feed this back into the first equation, we see that $y = -2x$.
Conversely, if $y = -2x$ and $z = x$ we see that both equations are satisfied.
It follows that $V \cap W = \{[x,-2x,x]^T \mid t \in \mathbb{R}\}$ as claimed.

Next, consider an arbitrary vector $\mathbf{u} = [x,y,z]^T \in \mathbb{R}^3$. Put

$$\mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix} \qquad \mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$$

Then $\mathbf{u} = \mathbf{v} + \mathbf{w}$ with $\mathbf{v} \in V$ and $\mathbf{w} \in W$, so $\mathbf{u} \in V + W$.
This works for any $\mathbf{u} \in \mathbb{R}^3$, so $\mathbb{R}^3 = V + W$. $\bigcirc$

Two planes

$$V = \{[x,y,z]^T \mid x+2y+3z=0\} \qquad W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$$

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix} \qquad \mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$$

Two planes

$$V = \{[x,y,z]^T \mid x+2y+3z=0\} \qquad W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$$

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix} \qquad \mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$$

$(12x+8y+4z)+2(3x+2y+z)+3(-6x-4y-2z)=0$, so $\mathbf{v} \in V$

Two planes

$$V = \{[x,y,z]^T \mid x+2y+3z=0\} \qquad W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$$

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix} \qquad \mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$$

$(12x+8y+4z)+2(3x+2y+z)+3(-6x-4y-2z)=0$, so $\mathbf{v} \in V$

$3(-8y-4z)+2(-3x+10y-z)+(6x+4y+14z)=0$, so $\mathbf{w} \in W$

Two planes

$$V = \{[x,y,z]^T \mid x+2y+3z=0\} \qquad W = \{[x,y,z]^T \mid 3x+2y+z=0\}.$$

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \mathbf{v} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z \\ 3x+2y+z \\ -6x-4y-2z \end{bmatrix} \qquad \mathbf{w} = \frac{1}{12}\begin{bmatrix} -8y-4z \\ -3x+10y-z \\ 6x+4y+14z \end{bmatrix}$$

$(12x+8y+4z)+2(3x+2y+z)+3(-6x-4y-2z)=0$, so $\mathbf{v} \in V$

$3(-8y-4z)+2(-3x+10y-z)+(6x+4y+14z)=0$, so $\mathbf{w} \in W$

$$\mathbf{v}+\mathbf{w} = \frac{1}{12}\begin{bmatrix} 12x+8y+4z-8y-4z \\ 3x+2y+z-3x+10y-z \\ -6x-4y-2z+6x+4y+14z \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12x \\ 12y \\ 12z \end{bmatrix} = \mathbf{u}$$

Subspaces of polynomials

Example ??: Take $U = \mathbb{R}[x]_{\leq 4}$ and

$$V = \{f \in U \mid f(0) = f'(0) = 0\} \qquad W = \{f \in U \mid f(-x) = f(x) \text{ for all } x\}.$$

Example ??: Take $U = \mathbb{R}[x]_{\leq 4}$ and

$$V = \{f \in U \mid f(0) = f'(0) = 0\} \qquad W = \{f \in U \mid f(-x) = f(x) \text{ for all } x\}.$$

Then

$$\begin{aligned} U &= \{a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 \mid a_0, \dots, a_4 \in \mathbb{R}\} \\ V &= \{a_2 x^2 + a_3 x^3 + a_4 x^4 \mid a_2, a_3, a_4 \in \mathbb{R}\} \\ W &= \{a_0 + a_2 x^2 + a_4 x^4 \mid a_0, a_2, a_4 \in \mathbb{R}\} \end{aligned}$$

Example ??: Take $U = \mathbb{R}[x]_{\leq 4}$ and

$$V = \{f \in U \mid f(0) = f'(0) = 0\} \qquad W = \{f \in U \mid f(-x) = f(x) \text{ for all } x\}.$$

Then

$$\begin{aligned} U &= \{a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \mid a_0, \dots, a_4 \in \mathbb{R}\} \\ V &= \{a_2x^2 + a_3x^3 + a_4x^4 \mid a_2, a_3, a_4 \in \mathbb{R}\} \\ W &= \{a_0 + a_2x^2 + a_4x^4 \mid a_0, a_2, a_4 \in \mathbb{R}\} \end{aligned}$$

From this we see that

$$V \cap W = \{a_2x^2 + a_4x^4 \mid a_2, a_4 \in \mathbb{R}\}$$

Example ??: Take $U = \mathbb{R}[x]_{\leq 4}$ and

$$V = \{f \in U \mid f(0) = f'(0) = 0\} \qquad W = \{f \in U \mid f(-x) = f(x) \text{ for all } x\}.$$

Then

$$\begin{aligned} U &= \{a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \mid a_0, \dots, a_4 \in \mathbb{R}\} \\ V &= \{a_2x^2 + a_3x^3 + a_4x^4 \mid a_2, a_3, a_4 \in \mathbb{R}\} \\ W &= \{a_0 + a_2x^2 + a_4x^4 \mid a_0, a_2, a_4 \in \mathbb{R}\} \end{aligned}$$

From this we see that

$$\begin{aligned} V \cap W &= \{a_2x^2 + a_4x^4 \mid a_2, a_4 \in \mathbb{R}\} \\ V + W &= \{a_0 + a_2x^2 + a_3x^3 + a_4x^4 \mid a_0, a_2, a_3, a_4 \in \mathbb{R}\}. \end{aligned}$$

Example ??: Take $U = \mathbb{R}[x]_{\leq 4}$ and

$$V = \{f \in U \mid f(0) = f'(0) = 0\} \qquad W = \{f \in U \mid f(-x) = f(x) \text{ for all } x\}.$$

Then

$$\begin{aligned} U &= \{a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \mid a_0, \dots, a_4 \in \mathbb{R}\} \\ V &= \{a_2x^2 + a_3x^3 + a_4x^4 \mid a_2, a_3, a_4 \in \mathbb{R}\} \\ W &= \{a_0 + a_2x^2 + a_4x^4 \mid a_0, a_2, a_4 \in \mathbb{R}\} \end{aligned}$$

From this we see that

$$\begin{aligned} V \cap W &= \{a_2x^2 + a_4x^4 \mid a_2, a_4 \in \mathbb{R}\} \\ V + W &= \{a_0 + a_2x^2 + a_3x^3 + a_4x^4 \mid a_0, a_2, a_3, a_4 \in \mathbb{R}\}. \end{aligned}$$

In particular, the polynomial $f(x) = x$ does not lie in $V + W$, so $V + W \neq U$.○

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map.

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\ker(\phi) = \{u \in U \mid \phi(u) = 0\}$$

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned} \ker(\phi) &= \{u \in U \mid \phi(u) = 0\} \\ \mathrm{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}. \end{aligned}$$

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned}\ker(\phi) &= \{u \in U \mid \phi(u) = 0\} \\ \mathrm{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}.\end{aligned}$$

Example ??: Define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ z \end{bmatrix}$.

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned} \ker(\phi) &= \{u \in U \mid \phi(u) = 0\} \\ \mathrm{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}. \end{aligned}$$

Example ??: Define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi\left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ y \\ z \end{smallmatrix}\right]$. Then

$$\ker(\pi) = \{\left[\begin{smallmatrix} x \\ 0 \\ 0 \end{smallmatrix}\right] \mid x \in \mathbb{R}\}$$

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned}\ker(\phi) &= \{u \in U \mid \phi(u) = 0\}\\ \text{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}.\end{aligned}$$

Example ??: Define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi\left[\begin{smallmatrix}x\\y\\z\end{smallmatrix}\right] = \left[\begin{smallmatrix}0\\y\\z\end{smallmatrix}\right]$. Then

$$\ker(\pi) = \{\left[\begin{smallmatrix}x\\0\\0\end{smallmatrix}\right] \mid x \in \mathbb{R}\}$$

$$\text{image}(\pi) = \{\left[\begin{smallmatrix}0\\y\\z\end{smallmatrix}\right] \mid y, z \in \mathbb{R}\}$$

Kernels and images are subspaces

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map.

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\text{image}(\phi)$ is a subspace of V.

Proof for $\ker(\phi)$**:** We have $\phi(0_U) = 0_V$, which shows that $0_U \in \ker(\phi)$

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\ker(\phi)$**:** We have $\phi(0_U) = 0_V$, which shows that $0_U \in \ker(\phi)$
Next, suppose that $u, u' \in \ker(\phi)$

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\ker(\phi)$**:** We have $\phi(0_U) = 0_V$, which shows that $0_U \in \ker(\phi)$
Next, suppose that $u, u' \in \ker(\phi)$, which means that $\phi(u) = \phi(u') = 0$.
Suppose also that $t, t' \in \mathbb{R}$

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\operatorname{image}(\phi)$ is a subspace of V.

Proof for $\ker(\phi)$**:** We have $\phi(0_U) = 0_V$, which shows that $0_U \in \ker(\phi)$
Next, suppose that $u, u' \in \ker(\phi)$, which means that $\phi(u) = \phi(u') = 0$.
Suppose also that $t, t' \in \mathbb{R}$

As ϕ is linear this implies that
$\phi(tu + t'u') = t\phi(u) + t'\phi(u') = t.0 + t'.0 = 0,$

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\ker(\phi)$**:** We have $\phi(0_U) = 0_V$, which shows that $0_U \in \ker(\phi)$
Next, suppose that $u, u' \in \ker(\phi)$, which means that $\phi(u) = \phi(u') = 0$.
Suppose also that $t, t' \in \mathbb{R}$

As ϕ is linear this implies that
$\phi(tu + t'u') = t\phi(u) + t'\phi(u') = t.0 + t'.0 = 0$,
so $tu + t'u' \in \ker(\phi)$.

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\ker(\phi)$**:** We have $\phi(0_U) = 0_V$, which shows that $0_U \in \ker(\phi)$
Next, suppose that $u, u' \in \ker(\phi)$, which means that $\phi(u) = \phi(u') = 0$.
Suppose also that $t, t' \in \mathbb{R}$

As ϕ is linear this implies that
$\phi(tu + t'u') = t\phi(u) + t'\phi(u') = t.0 + t'.0 = 0$,
so $tu + t'u' \in \ker(\phi)$.

As this works for all $u, u' \in \ker(\phi)$ and $t, t' \in \mathbb{R}$, we deduce that $\ker(\phi)$ is a subspace.

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\mathrm{image}(\phi)$**:**

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\mathrm{image}(\phi)$**:**
We have $\phi(0_U) = 0_V$, which shows that $0_V \in \mathrm{image}(\phi)$

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\mathrm{image}(\phi)$**:**
We have $\phi(0_U) = 0_V$, which shows that $0_V \in \mathrm{image}(\phi)$

Now suppose we have $v, v' \in \mathrm{image}(\phi)$ and $t, t' \in \mathbb{R}$.

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\operatorname{image}(\phi)$ is a subspace of V.

Proof for $\operatorname{image}(\phi)$**:**
We have $\phi(0_U) = 0_V$, which shows that $0_V \in \operatorname{image}(\phi)$

Now suppose we have $v, v' \in \operatorname{image}(\phi)$ and $t, t' \in \mathbb{R}$.

As $v, v' \in \operatorname{image}(\phi)$, we can find $x, x' \in U$ with $\phi(x) = v$ and $\phi(x') = v'$.

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\mathrm{image}(\phi)$**:**
We have $\phi(0_U) = 0_V$, which shows that $0_V \in \mathrm{image}(\phi)$

Now suppose we have $v, v' \in \mathrm{image}(\phi)$ and $t, t' \in \mathbb{R}$.

As $v, v' \in \mathrm{image}(\phi)$, we can find $x, x' \in U$ with $\phi(x) = v$ and $\phi(x') = v'$.
We thus have $tx + t'x' \in U$

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\mathrm{image}(\phi)$ is a subspace of V.

Proof for $\mathrm{image}(\phi)$**:**
We have $\phi(0_U) = 0_V$, which shows that $0_V \in \mathrm{image}(\phi)$

Now suppose we have $v, v' \in \mathrm{image}(\phi)$ and $t, t' \in \mathbb{R}$.

As $v, v' \in \mathrm{image}(\phi)$, we can find $x, x' \in U$ with $\phi(x) = v$ and $\phi(x') = v'$.
We thus have $tx + t'x' \in U$, and as ϕ is linear we have
$\phi(tx + t'x') = t\phi(x) + t'\phi(x') = tv + t'v'$
This shows that $tv + t'v' \in \mathrm{image}(\phi)$.

Kernels and images are subspaces

Proposition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then $\ker(\phi)$ is a subspace of U, and $\text{image}(\phi)$ is a subspace of V.

Proof for $\text{image}(\phi)$**:**
We have $\phi(0_U) = 0_V$, which shows that $0_V \in \text{image}(\phi)$

Now suppose we have $v, v' \in \text{image}(\phi)$ and $t, t' \in \mathbb{R}$.

As $v, v' \in \text{image}(\phi)$, we can find $x, x' \in U$ with $\phi(x) = v$ and $\phi(x') = v'$.
We thus have $tx + t'x' \in U$, and as ϕ is linear we have
$\phi(tx + t'x') = t\phi(x) + t'\phi(x') = tv + t'v'$
This shows that $tv + t'v' \in \text{image}(\phi)$.

As this works for all $v, v' \in \text{image}(\phi)$ and $t, t' \in \mathbb{R}$, we deduce that $\text{image}(\phi)$ is a subspace.

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map.

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\ker(\phi) = \{u \in U \mid \phi(u) = 0\}$$

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned} \ker(\phi) &= \{u \in U \mid \phi(u) = 0\} \\ \mathrm{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}. \end{aligned}$$

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned} \ker(\phi) &= \{u \in U \mid \phi(u) = 0\} \\ \operatorname{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}. \end{aligned}$$

Example ??: Define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ z \end{bmatrix}$.

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\begin{aligned} \ker(\phi) &= \{u \in U \mid \phi(u) = 0\} \\ \mathrm{image}(\phi) &= \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}. \end{aligned}$$

Example ??: Define $\pi : \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ z \end{bmatrix}$. Then

$$\ker(\pi) = \{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \mid x \in \mathbb{R} \}$$

Kernels and images

Definition ??: Let U and V be vector spaces, and let $\phi\colon U \to V$ be a linear map. Then we write

$$\ker(\phi) = \{u \in U \mid \phi(u) = 0\}$$
$$\mathrm{image}(\phi) = \{v \in V \mid v = \phi(u) \text{ for some } u \in U\}.$$

Example ??: Define $\pi : \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ z \end{bmatrix}$. Then

$$\ker(\pi) = \{\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \mid x \in \mathbb{R}\}$$
$$\mathrm{image}(\pi) = \{\begin{bmatrix} 0 \\ y \\ z \end{bmatrix} \mid y, z \in \mathbb{R}\}$$

An example of kernels and images

An example of kernels and images

Example ??: Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x, y, z]^T) = [2x - z, 2y - 8x, 2z - y]^T$.

An example of kernels and images

Example ??: Define $\phi \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x, y, z]^T) = [2x - z, 2y - 8x, 2z - y]^T$.

Then

$$\ker(\phi) = \{[x, y, z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\}$$

An example of kernels and images

Example ??: Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x, y, z]^T) = [2x - z, 2y - 8x, 2z - y]^T$.

Then

$$\ker(\phi) = \{[x, y, z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\} = \{[t, 4t, 2t]^T \mid t \in \mathbb{R}\}$$

An example of kernels and images

Example ??: Define $\phi \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x, y, z]^T) = [2x - z, 2y - 8x, 2z - y]^T$.

Then

$$\ker(\phi) = \{[x, y, z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\} = \{[t, 4t, 2t]^T \mid t \in \mathbb{R}\}$$

image(ϕ)

An example of kernels and images

Example ??: Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x,y,z]^T) = [2x - z, 2y - 8x, 2z - y]^T$.

Then

$$\ker(\phi) = \{[x,y,z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\} = \{[t,4t,2t]^T \mid t \in \mathbb{R}\}$$
$$\text{image}(\phi) = \{[u,v,w]^T \in \mathbb{R}^3 \mid 4u + v + 2w = 0\}$$

An example of kernels and images

Example ??: Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x,y,z]^T) = [2x-z, 2y-8x, 2z-y]^T$.

Then

$$\begin{aligned} \ker(\phi) &= \{[x,y,z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\} = \{[t,4t,2t]^T \mid t \in \mathbb{R}\} \\ \mathrm{image}(\phi) &= \{[u,v,w]^T \in \mathbb{R}^3 \mid 4u + v + 2w = 0\} = \{[u,v,-2u-v/2]^T \mid u,v \in \mathbb{R}^2\}. \end{aligned}$$

An example of kernels and images

Example ??: Define $\phi \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x,y,z]^T) = [2x-z, 2y-8x, 2z-y]^T$.

Then

$$\ker(\phi) = \{[x,y,z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\} = \{[t, 4t, 2t]^T \mid t \in \mathbb{R}\}$$
$$\text{image}(\phi) = \{[u,v,w]^T \in \mathbb{R}^3 \mid 4u + v + 2w = 0\} = \{[u, v, -2u - v/2]^T \mid u, v \in \mathbb{R}^2\}.$$

So $\ker(\phi)$ is a line through the origin (and thus a one-dimensional subspace)

An example of kernels and images

Example ??: Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi([x, y, z]^T) = [2x - z, 2y - 8x, 2z - y]^T$.

Then

$$\begin{aligned} \ker(\phi) &= \{[x, y, z]^T \in \mathbb{R}^3 \mid z = 2x,\ y = 4x,\ (2z = y)\} = \{[t, 4t, 2t]^T \mid t \in \mathbb{R}\} \\ \mathrm{image}(\phi) &= \{[u, v, w]^T \in \mathbb{R}^3 \mid 4u + v + 2w = 0\} = \{[u, v, -2u - v/2]^T \mid u, v \in \mathbb{R}^2\}. \end{aligned}$$

So $\ker(\phi)$ is a line through the origin (and thus a one-dimensional subspace) and $\mathrm{image}(\phi)$ is a plane through the origin (and thus a two-dimensional subspace). ○

Antisymmetrisation

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix.

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear). Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear). Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear). Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$.

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).
Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\operatorname{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\operatorname{image}(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$

Antisymmetrisation

Example ??: Define $\phi \colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear). Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that image$(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$.

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$.
This shows that $\phi(A)$ is always antisymmetric

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$.
This shows that $\phi(A)$ is always antisymmetric, so $\text{image}(\phi) \subseteq W$.

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear). Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that image$(\phi) = W$. For any A we have $\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$. This shows that $\phi(A)$ is always antisymmetric, so image$(\phi) \subseteq W$. Next, if B is antisymmetric then $B^T = -B$

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that image$(\phi) = W$. For any A we have $\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$. This shows that $\phi(A)$ is always antisymmetric, so image$(\phi) \subseteq W$. Next, if B is antisymmetric then $B^T = -B$ so $\phi(B/2) = B/2 - B^T/2 = B/2 + B/2 = B$.

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$. For any A we have $\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$. This shows that $\phi(A)$ is always antisymmetric, so $\text{image}(\phi) \subseteq W$. Next, if B is antisymmetric then $B^T = -B$ so $\phi(B/2) = B/2 - B^T/2 = B/2 + B/2 = B$. Thus B is ϕ(something)

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$. This shows that $\phi(A)$ is always antisymmetric, so $\text{image}(\phi) \subseteq W$. Next, if B is antisymmetric then $B^T = -B$ so $\phi(B/2) = B/2 - B^T/2 = B/2 + B/2 = B$. Thus B is $\phi(\text{something})$, so $B \in \text{image}(\phi)$.

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\text{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\text{image}(\phi) = W$. For any A we have
$\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$.
This shows that $\phi(A)$ is always antisymmetric, so $\text{image}(\phi) \subseteq W$. Next, if B is antisymmetric then $B^T = -B$ so $\phi(B/2) = B/2 - B^T/2 = B/2 + B/2 = B$.
Thus B is ϕ(something), so $B \in \text{image}(\phi)$. This shows that $W \subseteq \text{image}(\phi)$

Antisymmetrisation

Example ??: Define $\phi\colon M_n\mathbb{R} \to M_n\mathbb{R}$ by $\phi(A) = A - A^T$ (which is linear).

Then clearly $\phi(A) = 0$ iff $A = A^T$ iff A is a symmetric matrix. Thus

$$\ker(\phi) = \{n \times n \text{ symmetric matrices }\}.$$

We claim that also

$$\mathrm{image}(\phi) = \{n \times n \text{ antisymmetric matrices }\}.$$

For brevity, we write W for the set of antisymmetric matrices, so we must show that $\mathrm{image}(\phi) = W$. For any A we have $\phi(A)^T = (A - A^T)^T = A^T - A^{TT}$, but $A^{TT} = A$, so $\phi(A)^T = A^T - A = -\phi(A)$. This shows that $\phi(A)$ is always antisymmetric, so $\mathrm{image}(\phi) \subseteq W$. Next, if B is antisymmetric then $B^T = -B$ so $\phi(B/2) = B/2 - B^T/2 = B/2 + B/2 = B$. Thus B is $\phi(\text{something})$, so $B \in \mathrm{image}(\phi)$. This shows that $W \subseteq \mathrm{image}(\phi)$, so $W = \mathrm{image}(\phi)$ as claimed. ○

A polynomial example

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$.

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly:
$\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly:
$\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words
$b = a+b = 2a+b = 0$

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly:
$\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words
$b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$.

A polynomial example

Example ??: Define $\phi \colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$. This means that $\ker(\phi) = \{0\}$.

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$.
This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\mathrm{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$. This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\mathrm{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \mathrm{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$.

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly:
$\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$.
This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\text{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \text{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$. This means that $u - 2v + w = b - 2(a+b) + 2a + b = 0$, as required.

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$. This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\mathrm{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \mathrm{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$. This means that $u - 2v + w = b - 2(a+b) + 2a + b = 0$, as required.

Conversely, suppose that we have a vector $[u,v,w]^T \in \mathbb{R}^3$ with $u - 2v + w = 0$.

A polynomial example

Example ??: Define $\phi \colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$. This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\text{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \text{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$. This means that $u - 2v + w = b - 2(a+b) + 2a + b = 0$, as required.

Conversely, suppose that we have a vector $[u,v,w]^T \in \mathbb{R}^3$ with $u - 2v + w = 0$. We then have $w = 2v - u$

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$. This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\mathrm{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \mathrm{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$. This means that $u - 2v + w = b - 2(a+b) + 2a + b = 0$, as required.

Conversely, suppose that we have a vector $[u,v,w]^T \in \mathbb{R}^3$ with $u - 2v + w = 0$. We then have $w = 2v - u$ and so

$$\phi((v-u)x+u) = \begin{bmatrix} u \\ (v-u)+u \\ 2(v-u)+u \end{bmatrix} = \begin{bmatrix} u \\ v \\ 2v-u \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly:
$\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$.
This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\mathrm{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \mathrm{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$. This means that $u - 2v + w = b - 2(a+b) + 2a + b = 0$, as required.

Conversely, suppose that we have a vector $[u,v,w]^T \in \mathbb{R}^3$ with $u - 2v + w = 0$. We then have $w = 2v - u$ and so

$$\phi((v-u)x + u) = \begin{bmatrix} u \\ (v-u)+u \\ 2(v-u)+u \end{bmatrix} = \begin{bmatrix} u \\ v \\ 2v-u \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

so $[u,v,w]^T$ is in the image of ϕ. $\bigcirc$

A polynomial example

Example ??: Define $\phi\colon \mathbb{R}[x]_{\leq 1} \to \mathbb{R}^3$ by $\phi(f) = [f(0), f(1), f(2)]^T$. Explicitly: $\phi(ax+b) = [b, a+b, 2a+b]^T = a[0,1,2]^T + b[1,1,1]^T$.
If $ax+b \in \ker(\phi)$ then we must have $\phi(ax+b) = 0$, or in other words $b = a+b = 2a+b = 0$, which implies that $a = b = 0$ and so $ax+b = 0$.
This means that $\ker(\phi) = \{0\}$.

Next, we claim that $\mathrm{image}(\phi) = \{[u,v,w]^T \mid u - 2v + w = 0\}$.
Indeed, if $[u,v,w]^T \in \mathrm{image}(\phi)$ then we must have $[u,v,w] = \phi(ax+b) = [b, a+b, 2a+b]$ for some $a, b \in \mathbb{R}$. This means that $u - 2v + w = b - 2(a+b) + 2a + b = 0$, as required.

Conversely, suppose that we have a vector $[u,v,w]^T \in \mathbb{R}^3$ with $u - 2v + w = 0$. We then have $w = 2v - u$ and so

$$\phi((v-u)x+u) = \begin{bmatrix} u \\ (v-u)+u \\ 2(v-u)+u \end{bmatrix} = \begin{bmatrix} u \\ v \\ 2v-u \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix},$$

so $[u,v,w]^T$ is in the image of ϕ. ○

Another example

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x - 1)^2.$$

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x - 1)^2.$$

It follows that $\ker(\phi) = \{a(x - 1)^2 \mid a \in \mathbb{R}\}$.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x - 1)^2.$$

It follows that $\ker(\phi) = \{a(x-1)^2 \mid a \in \mathbb{R}\}$. In particular, $\ker(\phi)$ is nonzero, so ϕ is not injective.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x-1)^2.$$

It follows that $\ker(\phi) = \{a(x-1)^2 \mid a \in \mathbb{R}\}$. In particular, $\ker(\phi)$ is nonzero, so ϕ is not injective. Explicitly, we have $x^2 + 1 \neq 2x$ but
$\phi(x^2 + 1) = [2, 2]^T = \phi(2x)$.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x-1)^2.$$

It follows that $\ker(\phi) = \{a(x-1)^2 \mid a \in \mathbb{R}\}$. In particular, $\ker(\phi)$ is nonzero, so ϕ is not injective. Explicitly, we have $x^2 + 1 \neq 2x$ but
$\phi(x^2 + 1) = [2, 2]^T = \phi(2x)$.

On the other hand, we claim that ϕ is surjective.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x-1)^2.$$

It follows that $\ker(\phi) = \{a(x-1)^2 \mid a \in \mathbb{R}\}$. In particular, $\ker(\phi)$ is nonzero, so ϕ is not injective. Explicitly, we have $x^2 + 1 \neq 2x$ but
$\phi(x^2 + 1) = [2, 2]^T = \phi(2x)$.

On the other hand, we claim that ϕ is surjective. Indeed, for any vector $\mathbf{a} = [u, v]^T \in \mathbb{R}^2$ we check that

$$\phi(vx + u - v) = [v + u - v, v]^T = [u, v]^T = \mathbf{a}$$

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x - 1)^2.$$

It follows that $\ker(\phi) = \{a(x - 1)^2 \mid a \in \mathbb{R}\}$. In particular, $\ker(\phi)$ is nonzero, so ϕ is not injective. Explicitly, we have $x^2 + 1 \neq 2x$ but
$\phi(x^2 + 1) = [2, 2]^T = \phi(2x)$.

On the other hand, we claim that ϕ is surjective. Indeed, for any vector $\mathbf{a} = [u, v]^T \in \mathbb{R}^2$ we check that

$$\phi(vx + u - v) = [v + u - v, v]^T = [u, v]^T = \mathbf{a}$$

so $\mathbf{a}$ is ϕ(something) as required.

Another example

Define $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ by $\phi(f) = [f(1), f'(1)]^T$.
Explicitly: $\phi(ax^2 + bx + c) = [a + b + c, 2a + b]^T$.
It follows that $ax^2 + bx + c$ lies in $\ker(\phi)$ iff $a + b + c = 0 = 2a + b$, which gives $b = -2a$ and $c = -a - b = -a + 2a = a$, so

$$ax^2 + bx + c = ax^2 - 2ax + a = a(x^2 - 2x + 1) = a(x - 1)^2.$$

It follows that $\ker(\phi) = \{a(x-1)^2 \mid a \in \mathbb{R}\}$. In particular, $\ker(\phi)$ is nonzero, so ϕ is not injective. Explicitly, we have $x^2 + 1 \neq 2x$ but
$\phi(x^2 + 1) = [2, 2]^T = \phi(2x)$.

On the other hand, we claim that ϕ is surjective. Indeed, for any vector $\mathbf{a} = [u, v]^T \in \mathbb{R}^2$ we check that

$$\phi(vx + u - v) = [v + u - v, v]^T = [u, v]^T = \mathbf{a},$$

so $\mathbf{a}$ is ϕ(something) as required.

Another example

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x+2y+4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x+2y+4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Then

$$\ker(\phi) = \{[x, y, z]^T \in \mathbb{R}^3 \mid x + 2y + 4z = 0\}$$

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x+2y+4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Then

$$\ker(\phi) = \{[x,y,z]^T \in \mathbb{R}^3 \mid x+2y+4z=0\} = \{[-2y-4z,y,z]^T \mid y,z \in \mathbb{R}\}$$

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x + 2y + 4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Then

$$\ker(\phi) = \{[x, y, z]^T \in \mathbb{R}^3 \mid x + 2y + 4z = 0\} = \{[-2y - 4z, y, z]^T \mid y, z \in \mathbb{R}\}$$

$\mathrm{image}(\phi)$

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x + 2y + 4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Then

$$\begin{aligned} \ker(\phi) &= \{[x, y, z]^T \in \mathbb{R}^3 \mid x + 2y + 4z = 0\} = \{[-2y - 4z, y, z]^T \mid y, z \in \mathbb{R}\} \\ \mathrm{image}(\phi) &= \{[t, 2t, 4t]^T \mid t \in \mathbb{R}\} \end{aligned}$$

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x+2y+4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Then

$$\ker(\phi) = \{[x,y,z]^T \in \mathbb{R}^3 \mid x+2y+4z = 0\} = \{[-2y-4z, y, z]^T \mid y, z \in \mathbb{R}\}$$
$$\operatorname{image}(\phi) = \{[t, 2t, 4t]^T \mid t \in \mathbb{R}\}$$

So $\ker(\phi)$ is a plane through the origin (and thus a two-dimensional subspace)

Another example

Define $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+4z \\ 2x+4y+8z \\ 4x+8y+16z \end{bmatrix} = (x + 2y + 4z) \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Then

$$\begin{aligned} \ker(\phi) &= \{[x, y, z]^T \in \mathbb{R}^3 \mid x + 2y + 4z = 0\} = \{[-2y - 4z, y, z]^T \mid y, z \in \mathbb{R}\} \\ \text{image}(\phi) &= \{[t, 2t, 4t]^T \mid t \in \mathbb{R}\} \end{aligned}$$

So $\ker(\phi)$ is a plane through the origin (and thus a two-dimensional subspace) and $\text{image}(\phi)$ is a line through the origin (and thus a one-dimensional subspace). $\bigcirc$

Injective and surjective maps

Injective and surjective maps

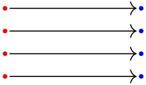
$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.

Injective and surjective maps

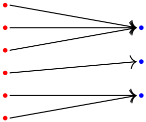
$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.

Injective and surjective maps

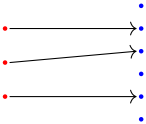
$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.



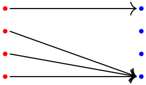
Injective and surjective



Surjective, not injective



Injective, not surjective



Neither surjective nor injective

Injective and surjective maps

$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.

Injective and surjective maps

$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.

Proposition ??: Let $\phi\colon U \to V$ be a linear map between vector spaces. Then ϕ is injective iff $\ker(\phi) = \{0\}$, and ϕ is surjective iff $\mathrm{image}(\phi) = V$.

Injective and surjective maps

$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.

Proposition ??: Let $\phi\colon U \to V$ be a linear map between vector spaces. Then ϕ is injective iff $\ker(\phi) = \{0\}$, and ϕ is surjective iff $\mathrm{image}(\phi) = V$.
Proof:

- Suppose that ϕ is injective, so whenever $\phi(u) = \phi(u')$ we have $u = u'$. Suppose that $u \in \ker(\phi)$. Then $\phi(u) = 0 = \phi(0)$. As ϕ is injective and $\phi(u) = \phi(0)$, we must have $u = 0$. Thus $\ker(\phi) = \{0\}$, as claimed.

Injective and surjective maps

$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.

Proposition ??: Let $\phi\colon U \to V$ be a linear map between vector spaces. Then ϕ is injective iff $\ker(\phi) = \{0\}$, and ϕ is surjective iff $\text{image}(\phi) = V$.
Proof:

- Suppose that ϕ is injective, so whenever $\phi(u) = \phi(u')$ we have $u = u'$. Suppose that $u \in \ker(\phi)$. Then $\phi(u) = 0 = \phi(0)$. As ϕ is injective and $\phi(u) = \phi(0)$, we must have $u = 0$. Thus $\ker(\phi) = \{0\}$, as claimed.
- Conversely, suppose that $\ker(\phi) = \{0\}$. Suppose that $\phi(u) = \phi(u')$. Then $\phi(u - u') = \phi(u) - \phi(u') = 0$, so $u - u' \in \ker(\phi) = \{0\}$, so $u - u' = 0$, so $u = u'$. This means that ϕ is injective.

Injective and surjective maps

$\phi\colon U \to V$ is *surjective* if every $v \in V$ has the form $\phi(u)$ for some $u \in U$.
$\phi\colon U \to V$ is said to be *injective* if whenever $\phi(u) = \phi(u')$ we have $u = u'$.

Proposition ??: Let $\phi\colon U \to V$ be a linear map between vector spaces. Then ϕ is injective iff $\ker(\phi) = \{0\}$, and ϕ is surjective iff $\mathrm{image}(\phi) = V$.
Proof:

- Suppose that ϕ is injective, so whenever $\phi(u) = \phi(u')$ we have $u = u'$. Suppose that $u \in \ker(\phi)$. Then $\phi(u) = 0 = \phi(0)$. As ϕ is injective and $\phi(u) = \phi(0)$, we must have $u = 0$. Thus $\ker(\phi) = \{0\}$, as claimed.
- Conversely, suppose that $\ker(\phi) = \{0\}$. Suppose that $\phi(u) = \phi(u')$. Then $\phi(u - u') = \phi(u) - \phi(u') = 0$, so $u - u' \in \ker(\phi) = \{0\}$, so $u - u' = 0$, so $u = u'$. This means that ϕ is injective.
- Recall that $\mathrm{image}(\phi)$ is the set of those $v \in V$ such that $v = \phi(u)$ for some $u \in U$. Thus $\mathrm{image}(\phi) = V$ iff *every* element $v \in V$ has the form $\phi(u)$ for some $u \in U$, which is precisely what it means for ϕ to be surjective. $\bigcirc$

Corollary ??: $\phi \colon U \to V$ is an isomorphism iff $\ker(\phi) = 0$ and $\mathrm{image}(\phi) = V$.

Direct sums

Direct sums

Definition ??: Let V and W be vector spaces. We define $V \oplus W$ to be the set of pairs (v, w) with $v \in V$ and $w \in W$. Addition and scalar multiplication are defined in the obvious way:

$$(v, w) + (v', w') = (v + v', w + w')$$
$$t.(v, w) = (tv, tw).$$

This makes $V \oplus W$ into a vector space, called the *direct sum* of V and W. We may sometimes use the notation $V \times W$ instead of $V \oplus W$.

Definition ??: Let V and W be vector spaces. We define $V \oplus W$ to be the set of pairs (v, w) with $v \in V$ and $w \in W$. Addition and scalar multiplication are defined in the obvious way:

$$(v, w) + (v', w') = (v + v', w + w')$$
$$t.(v, w) = (tv, tw).$$

This makes $V \oplus W$ into a vector space, called the *direct sum* of V and W. We may sometimes use the notation $V \times W$ instead of $V \oplus W$.

Example ??: An element of $\mathbb{R}^p \oplus \mathbb{R}^q$ is a pair $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ is a list of p real numbers, and $\mathbf{y}$ is a list of q real numbers. Such a pair is essentially the same thing as a list of $p + q$ real numbers, so $\mathbb{R}^p \oplus \mathbb{R}^q = \mathbb{R}^{p+q}$. ◯

Two subspaces

Two subspaces

Now suppose that V and W are subspaces of a third space U.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma\colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma \colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma \colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma \colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$.

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma\colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$. This means that $x \in V$ and $y \in W$ and $y = -x$.

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma\colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$. This means that $x \in V$ and $y \in W$ and $y = -x$. Note then that $x = -y$ and $y \in W$ so $x \in W$.

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma \colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$. This means that $x \in V$ and $y \in W$ and $y = -x$. Note then that $x = -y$ and $y \in W$ so $x \in W$. We also have $x \in V$, so $x \in V \cap W$.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma\colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$. This means that $x \in V$ and $y \in W$ and $y = -x$. Note then that $x = -y$ and $y \in W$ so $x \in W$. We also have $x \in V$, so $x \in V \cap W$. This shows that $\ker(\sigma) = \{(x, -x) \mid x \in V \cap W\}$, as claimed.

Two subspaces

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma\colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$. This means that $x \in V$ and $y \in W$ and $y = -x$. Note then that $x = -y$ and $y \in W$ so $x \in W$. We also have $x \in V$, so $x \in V \cap W$. This shows that $\ker(\sigma) = \{(x, -x) \mid x \in V \cap W\}$, as claimed. If $V \cap W = 0$ then we get $\ker(\sigma) = 0$, so σ is injective (by Proposition **??**).

Now suppose that V and W are subspaces of a third space U. We then have a space $V \oplus W$ as above, and also a subspace $V + W \leq U$ as in Definition **??**. We need to understand the relationship between these.

Proposition ??: The rule $\sigma(v, w) = v + w$ defines a linear map $\sigma\colon V \oplus W \to U$, whose image is $V + W$, and whose kernel is the space $X = \{(x, -x) \in V \oplus W \mid x \in V \cap W\}$. Thus, if $V \cap W = 0$ then $\ker(\sigma) = 0$ and σ gives an isomorphism $V \oplus W \to V + W$.

Proof: We leave it as an exercise to check that σ is a linear map. The image is the set of things of the form $v + w$ for some $v \in V$ and $w \in W$, which is precisely the definition of $V + W$. The kernel is the set of pairs $(x, y) \in V \oplus W$ for which $x + y = 0$. This means that $x \in V$ and $y \in W$ and $y = -x$. Note then that $x = -y$ and $y \in W$ so $x \in W$. We also have $x \in V$, so $x \in V \cap W$. This shows that $\ker(\sigma) = \{(x, -x) \mid x \in V \cap W\}$, as claimed. If $V \cap W = 0$ then we get $\ker(\sigma) = 0$, so σ is injective (by Proposition **??**). If we regard it as a map to $V + W$ (rather than to U) then it is also surjective, so it is an isomorphism $V \oplus W \to V + W$, as claimed. $\bigcirc$

Internal direct sums

Remark ??: If $V \cap W = 0$ and $V + W = U$ then σ gives an isomorphism $V \oplus W \to U$. In this situation it is common to say that $U = V \oplus W$.

Remark ??: If $V \cap W = 0$ and $V + W = U$ then σ gives an isomorphism $V \oplus W \to U$. In this situation it is common to say that $U = V \oplus W$.

This is not strictly true (because U is only isomorphic to $V \oplus W$, not equal to it), but it is a harmless abuse of language.

Internal direct sums

Remark ??: If $V \cap W = 0$ and $V + W = U$ then σ gives an isomorphism $V \oplus W \to U$. In this situation it is common to say that $U = V \oplus W$.

This is not strictly true (because U is only isomorphic to $V \oplus W$, not equal to it), but it is a harmless abuse of language.

Sometimes people call $V \oplus W$ the *external direct sum* of V and W, and they say that U is the *internal direct sum* of V and W if $U = V + W$ and $V \cap W = 0$. ○

Remark ??: If $V \cap W = 0$ and $V + W = U$ then the map $\sigma(v, w) = v + w$ gives an isomorphism $V \oplus W \to U$. In this situation it is common to say that $U = V \oplus W$.

Remark ??: If $V \cap W = 0$ and $V + W = U$ then the map $\sigma(v, w) = v + w$ gives an isomorphism $V \oplus W \to U$. In this situation it is common to say that $U = V \oplus W$.

This is not strictly true (because U is only isomorphic to $V \oplus W$, not equal to it), but it is a harmless abuse of language.

Internal direct sums

Remark ??: If $V \cap W = 0$ and $V + W = U$ then the map $\sigma(v, w) = v + w$ gives an isomorphism $V \oplus W \to U$. In this situation it is common to say that $U = V \oplus W$.

This is not strictly true (because U is only isomorphic to $V \oplus W$, not equal to it), but it is a harmless abuse of language.

Sometimes people call $V \oplus W$ the *external direct sum* of V and W, and they say that U is the *internal direct sum* of V and W if $U = V + W$ and $V \cap W = 0$. ○

Odd and even functions

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$. Thus $EF \cap OF = 0$, as required.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$. Thus $EF \cap OF = 0$, as required. Next, consider an arbitrary function $g \in F$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$. Thus $EF \cap OF = 0$, as required. Next, consider an arbitrary function $g \in F$. Put

$$g_+(x) = (g(x) + g(-x))/2 \qquad g_-(x) = (g(x) - g(-x))/2.$$

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$. Thus $EF \cap OF = 0$, as required. Next, consider an arbitrary function $g \in F$. Put

$$g_+(x) = (g(x) + g(-x))/2 \qquad\qquad g_-(x) = (g(x) - g(-x))/2.$$

Then

$$g_+(-x) = (g(-x)+g(x))/2 = g_+(x) \qquad g_-(-x) = (g(-x)-g(x))/2 = -g_-(x),$$

so $g_+ \in EF$ and $g_- \in OF$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$. Thus $EF \cap OF = 0$, as required. Next, consider an arbitrary function $g \in F$. Put

$$g_+(x) = (g(x) + g(-x))/2 \qquad\qquad g_-(x) = (g(x) - g(-x))/2.$$

Then

$$g_+(-x) = (g(-x)+g(x))/2 = g_+(x) \qquad g_-(-x) = (g(-x)-g(x))/2 = -g_-(x),$$

so $g_+ \in EF$ and $g_- \in OF$. It is also clear from the formulae that $g = g_+ + g_-$, so $g \in EF + OF$.

Example ??: Consider the space F of all functions from $\mathbb{R}$ to $\mathbb{R}$, and the subspaces EF and OF of even functions and odd functions.

We claim that $F = EF \oplus OF$.

To prove this, we must check that $EF \cap OF = 0$ and $EF + OF = F$.

Suppose that $f \in EF \cap OF$. Then for any x we have $f(x) = f(-x)$ (because $f \in EF$), but $f(-x) = -f(x)$ (because $f \in OF$), so $f(x) = -f(x)$, so $f(x) = 0$. Thus $EF \cap OF = 0$, as required. Next, consider an arbitrary function $g \in F$. Put

$$g_+(x) = (g(x) + g(-x))/2 \qquad g_-(x) = (g(x) - g(-x))/2.$$

Then

$$g_+(-x) = (g(-x)+g(x))/2 = g_+(x) \qquad g_-(-x) = (g(-x)-g(x))/2 = -g_-(x),$$

so $g_+ \in EF$ and $g_- \in OF$. It is also clear from the formulae that $g = g_+ + g_-$, so $g \in EF + OF$. This shows that $EF + OF = F$, so $F = EF \oplus OF$ as claimed.
$\bigcirc$

Trace free matrices

Trace free matrices

Example ??: Put $U = M_2\mathbb{R}$ and

$$\begin{aligned} V &= \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\} \\ W &= \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}. \end{aligned}$$

Trace free matrices

Example ??: Put $U = M_2\mathbb{R}$ and

$$\begin{aligned} V &= \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\} \\ W &= \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}. \end{aligned}$$

We claim that $U = V \oplus W$.

Trace free matrices

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$.

Example ??: Put $U = M_2\mathbb{R}$ and

$$\begin{aligned} V &= \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\} \\ W &= \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}. \end{aligned}$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \mathrm{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$, but $\mathrm{trace}(A) = 0$ (because $A \in V$) whereas $\mathrm{trace}(tI) = 2t$, so we must have $t = 0$

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \mathrm{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$, but $\mathrm{trace}(A) = 0$ (because $A \in V$) whereas $\mathrm{trace}(tI) = 2t$, so we must have $t = 0$, which means that $A = 0$.

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \mathrm{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$, but $\mathrm{trace}(A) = 0$ (because $A \in V$) whereas $\mathrm{trace}(tI) = 2t$, so we must have $t = 0$, which means that $A = 0$. This shows that $V \cap W = 0$.

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$, but $\operatorname{trace}(A) = 0$ (because $A \in V$) whereas $\operatorname{trace}(tI) = 2t$, so we must have $t = 0$, which means that $A = 0$. This shows that $V \cap W = 0$.

Next, consider an arbitrary matrix $B = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in U$.

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \operatorname{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$, but $\operatorname{trace}(A) = 0$ (because $A \in V$) whereas $\operatorname{trace}(tI) = 2t$, so we must have $t = 0$, which means that $A = 0$. This shows that $V \cap W = 0$.

Next, consider an arbitrary matrix $B = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in U$. We can write this as $B = C + D$, where

$$C = \begin{bmatrix} (p-s)/2 & q \\ r & (s-p)/2 \end{bmatrix} \in V$$
$$D = \begin{bmatrix} (p+s)/2 & 0 \\ 0 & (p+s)/2 \end{bmatrix} = \frac{p+s}{2} I \in W.$$

Example ??: Put $U = M_2\mathbb{R}$ and

$$V = \{A \in M_2\mathbb{R} \mid \text{trace}(A) = 0\} = \{\left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] \mid a, b, c \in \mathbb{R}\}$$
$$W = \{tI \mid t \in \mathbb{R}\} = \{\left[\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right] \mid t \in \mathbb{R}\}.$$

We claim that $U = V \oplus W$. To check this, first suppose that $A \in V \cap W$. As $A \in W$ we have $A = tI$ for some $t \in \mathbb{R}$, but $\text{trace}(A) = 0$ (because $A \in V$) whereas $\text{trace}(tI) = 2t$, so we must have $t = 0$, which means that $A = 0$. This shows that $V \cap W = 0$.
Next, consider an arbitrary matrix $B = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in U$. We can write this as $B = C + D$, where

$$C = \begin{bmatrix} (p-s)/2 & q \\ r & (s-p)/2 \end{bmatrix} \in V$$
$$D = \begin{bmatrix} (p+s)/2 & 0 \\ 0 & (p+s)/2 \end{bmatrix} = \frac{p+s}{2} I \in W.$$

This shows that $U = V + W$. ○

Two randomly-chosen vectors in $\mathbb{R}^2$ will generally not be parallel; it is an important special case if they happen to point in the same direction. Similarly, given three vectors u, v and w in $\mathbb{R}^3$, there will usually not be any plane that contains all three vectors. This means that we can get from the origin to any point by travelling a certain (possibly negative) distance in the direction of u, then a certain distance in the direction of v, then a certain distance in the direction of w. The case where u, v and w all lie in a common plane will have special geometric significance in any purely mathematical problem, and will often have special physical significance in applied problems. Our task in this section is to generalise these ideas, and study the corresponding special cases in an arbitrary vector space V. The abstract picture will be illuminating even in the case of $\mathbb{R}^2$ and $\mathbb{R}^3$.

Linear independence

Definition ??: Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ be a list of elements of V.

Linear independence

Definition ??: Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ be a list of elements of V.

A *linear relation* between the v_i's is a vector $[\lambda_1, \ldots, \lambda_n]^T \in \mathbb{R}^n$ such that $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$.

Definition ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V.

A *linear relation* between the v_i's is a vector $[\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$.

The vector $[0, \dots, 0]^T$ is obviously a linear relation, called the *trivial relation*.

Linear independence

Definition ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V.

A *linear relation* between the v_i's is a vector $[\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$.

The vector $[0, \dots, 0]^T$ is obviously a linear relation, called the *trivial relation*.

If there is a nontrivial linear relation, we say that the list $\mathcal{V}$ is *linearly dependent*.

Linear independence

Definition ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V.

A *linear relation* between the v_i's is a vector $[\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$.

The vector $[0, \dots, 0]^T$ is obviously a linear relation, called the *trivial relation*.

If there is a nontrivial linear relation, we say that the list $\mathcal{V}$ is *linearly dependent*.

Otherwise, if the only relation is the trivial one, we say that the list $\mathcal{V}$ is *linearly independent*. ◯

Linear independence examples

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Linear independence examples

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation

Linear independence examples

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Linear independence examples

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Example ??: Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Linear independence examples

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Example ??: Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3]^T$ such that $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$

Linear independence examples

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Example ??: Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3]^T$ such that $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$, or equivalently

$$\begin{bmatrix} \lambda_1 \\ \lambda_1 + \lambda_2 \\ \lambda_1 + \lambda_2 + \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Example ??: Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3]^T$ such that $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$, or equivalently

$$\begin{bmatrix} \lambda_1 \\ \lambda_1+\lambda_2 \\ \lambda_1+\lambda_2+\lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

From this we see that $\lambda_1 = 0$, then from the equation $\lambda_1 + \lambda_2 = 0$ we see that $\lambda_2 = 0$, then from the equation $\lambda_1 + \lambda_2 + \lambda_3 = 0$ we see that $\lambda_3 = 0$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Example ??: Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3]^T$ such that $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$, or equivalently

$$\begin{bmatrix} \lambda_1 \\ \lambda_1+\lambda_2 \\ \lambda_1+\lambda_2+\lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

From this we see that $\lambda_1 = 0$, then from the equation $\lambda_1 + \lambda_2 = 0$ we see that $\lambda_2 = 0$, then from the equation $\lambda_1 + \lambda_2 + \lambda_3 = 0$ we see that $\lambda_3 = 0$. Thus, the only linear relation is the trivial one where $[\lambda_1, \lambda_2, \lambda_3]^T = [0, 0, 0]^T$

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Then $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = 0$, so $[1, -2, 1]^T$ is a nontrivial linear relation, so the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly dependent.

Example ??: Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3]^T$ such that $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$, or equivalently

$$\begin{bmatrix} \lambda_1 \\ \lambda_1 + \lambda_2 \\ \lambda_1 + \lambda_2 + \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

From this we see that $\lambda_1 = 0$, then from the equation $\lambda_1 + \lambda_2 = 0$ we see that $\lambda_2 = 0$, then from the equation $\lambda_1 + \lambda_2 + \lambda_3 = 0$ we see that $\lambda_3 = 0$. Thus, the only linear relation is the trivial one where $[\lambda_1, \lambda_2, \lambda_3]^T = [0, 0, 0]^T$, so our vectors v_1, v_2, v_3 are linearly independent. ◯

Linear independence of polynomials

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent.

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad\qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad\qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x, or equivalently

$$\lambda_0 + \lambda_1 + \lambda_2 = 0, \qquad 2\lambda_1 + 4\lambda_2 = 0, \qquad \lambda_1 + 4\lambda_2 = 0.$$

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad\qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad\qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x, or equivalently

$$\lambda_0 + \lambda_1 + \lambda_2 = 0, \qquad 2\lambda_1 + 4\lambda_2 = 0, \qquad \lambda_1 + 4\lambda_2 = 0.$$

Subtracting the last two equations gives $\lambda_1 = 0$

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x, or equivalently

$$\lambda_0 + \lambda_1 + \lambda_2 = 0, \qquad 2\lambda_1 + 4\lambda_2 = 0, \qquad \lambda_1 + 4\lambda_2 = 0.$$

Subtracting the last two equations gives $\lambda_1 = 0$, putting this in the last equation gives $\lambda_2 = 0$

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad\qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad\qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x, or equivalently

$$\lambda_0 + \lambda_1 + \lambda_2 = 0, \qquad 2\lambda_1 + 4\lambda_2 = 0, \qquad \lambda_1 + 4\lambda_2 = 0.$$

Subtracting the last two equations gives $\lambda_1 = 0$, putting this in the last equation gives $\lambda_2 = 0$, and now the first equation gives $\lambda_0 = 0$.

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x, or equivalently

$$\lambda_0 + \lambda_1 + \lambda_2 = 0, \qquad 2\lambda_1 + 4\lambda_2 = 0, \qquad \lambda_1 + 4\lambda_2 = 0.$$

Subtracting the last two equations gives $\lambda_1 = 0$, putting this in the last equation gives $\lambda_2 = 0$, and now the first equation gives $\lambda_0 = 0$. Thus, the only linear relation is $[\lambda_0, \lambda_1, \lambda_2]^T = [0, 0, 0]^T$

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I claim that the list p_0, p_1, p_2 is linearly independent. Indeed, a linear relation between them is a vector $[\lambda_0, \lambda_1, \lambda_2]^T$ such that $\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 = 0$, or equivalently

$$(\lambda_0 + \lambda_1 + \lambda_2)x^2 + (2\lambda_1 + 4\lambda_2)x + (\lambda_1 + 4\lambda_2) = 0$$

for all x, or equivalently

$$\lambda_0 + \lambda_1 + \lambda_2 = 0, \qquad 2\lambda_1 + 4\lambda_2 = 0, \qquad \lambda_1 + 4\lambda_2 = 0.$$

Subtracting the last two equations gives $\lambda_1 = 0$, putting this in the last equation gives $\lambda_2 = 0$, and now the first equation gives $\lambda_0 = 0$. Thus, the only linear relation is $[\lambda_0, \lambda_1, \lambda_2]^T = [0, 0, 0]^T$, so the list p_0, p_1, p_2 is independent.

○

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I next claim, however, that the list p_0, p_1, p_2, p_3 is linearly dependent.

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I next claim, however, that the list p_0, p_1, p_2, p_3 is linearly dependent.

Indeed, you can check that

$$p_3 - 3p_2 + 3p_1 - p_0 = 0$$

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad\qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad\qquad p_3(x) = x^2 + 6x + 9.$$

I next claim, however, that the list p_0, p_1, p_2, p_3 is linearly dependent.

Indeed, you can check that

$$p_3 - 3p_2 + 3p_1 - p_0 = 0$$

so $[1, -3, 3, -1]^T$ is a nontrivial linear relation.

Linear independence of polynomials

Example ??: Consider the polynomials $p_n(x) = (x+n)^2$, so

$$p_0(x) = x^2 \qquad p_1(x) = x^2 + 2x + 1$$
$$p_2(x) = x^2 + 4x + 4 \qquad p_3(x) = x^2 + 6x + 9.$$

I next claim, however, that the list p_0, p_1, p_2, p_3 is linearly dependent.

Indeed, you can check that

$$p_3 - 3p_2 + 3p_1 - p_0 = 0$$

so $[1, -3, 3, -1]^T$ is a nontrivial linear relation.

(The entries in this list are the coefficients in the expansion of $(T-1)^3 = T^3 - 3T^2 + 3T - 1$; this is not a coincidence, but the explanation would take us too far afield.) ○

Example ??: Consider the functions

$$f_1(x) = e^x$$
$$f_2(x) = e^{-x}$$
$$f_3(x) = \sinh(x)$$
$$f_4(x) = \cosh(x).$$

Example ??: Consider the functions

$$\begin{aligned} f_1(x) &= e^x \\ f_2(x) &= e^{-x} \\ f_3(x) &= \sinh(x) \\ f_4(x) &= \cosh(x). \end{aligned}$$

These are linearly dependent

Linear dependence of functions

Example ??: Consider the functions

$$
\begin{aligned}
f_1(x) &= e^x \\
f_2(x) &= e^{-x} \\
f_3(x) &= \sinh(x) \\
f_4(x) &= \cosh(x).
\end{aligned}
$$

These are linearly dependent, because $\sinh(x)$ is by definition just $(e^x - e^{-x})/2$, so

$$f_1 - f_2 - 2f_3 = e^x - e^{-x} - (e^x - e^{-x}) = 0$$

Example ??: Consider the functions

$$f_1(x) = e^x$$
$$f_2(x) = e^{-x}$$
$$f_3(x) = \sinh(x)$$
$$f_4(x) = \cosh(x).$$

These are linearly dependent, because $\sinh(x)$ is by definition just $(e^x - e^{-x})/2$, so

$$f_1 - f_2 - 2f_3 = e^x - e^{-x} - (e^x - e^{-x}) = 0$$

so $[1, -1, -2, 0]^T$ is a nontrivial linear relation.

Example ??: Consider the functions

$$f_1(x) = e^x$$
$$f_2(x) = e^{-x}$$
$$f_3(x) = \sinh(x)$$
$$f_4(x) = \cosh(x).$$

These are linearly dependent, because $\sinh(x)$ is by definition just $(e^x - e^{-x})/2$, so

$$f_1 - f_2 - 2f_3 = e^x - e^{-x} - (e^x - e^{-x}) = 0$$

so $[1, -1, -2, 0]^T$ is a nontrivial linear relation. Similarly, we have $\cosh(x) = (e^x + e^{-x})/2$

Example ??: Consider the functions

$$\begin{aligned} f_1(x) &= e^x \\ f_2(x) &= e^{-x} \\ f_3(x) &= \sinh(x) \\ f_4(x) &= \cosh(x). \end{aligned}$$

These are linearly dependent, because $\sinh(x)$ is by definition just $(e^x - e^{-x})/2$, so

$$f_1 - f_2 - 2f_3 = e^x - e^{-x} - (e^x - e^{-x}) = 0$$

so $[1, -1, -2, 0]^T$ is a nontrivial linear relation. Similarly, we have $\cosh(x) = (e^x + e^{-x})/2$, so $f_4 = \frac{1}{2} f_1 + \frac{1}{2} f_2$

Example ??: Consider the functions

$$\begin{aligned} f_1(x) &= e^x \\ f_2(x) &= e^{-x} \\ f_3(x) &= \sinh(x) \\ f_4(x) &= \cosh(x). \end{aligned}$$

These are linearly dependent, because $\sinh(x)$ is by definition just $(e^x - e^{-x})/2$, so

$$f_1 - f_2 - 2f_3 = e^x - e^{-x} - (e^x - e^{-x}) = 0$$

so $[1, -1, -2, 0]^T$ is a nontrivial linear relation. Similarly, we have $\cosh(x) = (e^x + e^{-x})/2$, so $f_4 = \frac{1}{2}f_1 + \frac{1}{2}f_2$, so $[\frac{1}{2}, \frac{1}{2}, 0, -1]^T$ is another linear relation. ○

Linear independence of matrices

Linear independence of matrices

Example ??: Consider the matrices

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example ??: Consider the matrices

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]^T$ such that $\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4$ is the zero matrix.

Linear independence of matrices

Example ??: Consider the matrices

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]^T$ such that $\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4$ is the zero matrix. But

$$\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4 = \begin{bmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{bmatrix}$$

Example ??: Consider the matrices

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right].$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]^T$ such that $\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4$ is the zero matrix. But

$$\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4 = \left[\begin{smallmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{smallmatrix}\right]$$

and this is only the zero matrix if $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.

Linear independence of matrices

Example ??: Consider the matrices

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right].$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]^T$ such that $\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4$ is the zero matrix. But

$$\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4 = \left[\begin{smallmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{smallmatrix}\right]$$

and this is only the zero matrix if $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.

Thus, the only linear relation is the trivial one

Linear independence of matrices

Example ??: Consider the matrices

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

A linear relation between these is a vector $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]^T$ such that $\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4$ is the zero matrix. But

$$\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 + \lambda_4 E_4 = \begin{bmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{bmatrix}$$

and this is only the zero matrix if $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.

Thus, the only linear relation is the trivial one, showing that $E_1, \ldots, E_4$ are linearly independent. ◯

Linear independence and the map $\mu_{\mathcal{V}}$

Linear independence and the map $\mu_{\mathcal{V}}$

Remark ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V.

Linear independence and the map $\mu_{\mathcal{V}}$

Remark ??: Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ be a list of elements of V. We have a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$, given by

$$\mu_{\mathcal{V}}([\lambda_1, \ldots, \lambda_n]^T) = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Remark ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V. We have a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$, given by

$$\mu_{\mathcal{V}}([\lambda_1, \dots, \lambda_n]^T) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

By definition, a linear relation between the v_i's is just a vector $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = 0$

Remark ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V. We have a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$, given by

$$\mu_{\mathcal{V}}([\lambda_1, \dots, \lambda_n]^T) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

By definition, a linear relation between the v_i's is just a vector $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = 0$, or in other words, an element of the kernel of $\mu_{\mathcal{V}}$.

Linear independence and the map $\mu_{\mathcal{V}}$

Remark ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V. We have a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$, given by

$$\mu_{\mathcal{V}}([\lambda_1, \dots, \lambda_n]^T) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

By definition, a linear relation between the v_i's is just a vector $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = 0$, or in other words, an element of the kernel of $\mu_{\mathcal{V}}$.

Thus, $\mathcal{V}$ is linearly independent iff $\ker(\mu_{\mathcal{V}}) = \{0\}$

Linear independence and the map $\mu_{\mathcal{V}}$

Remark ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements of V. We have a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$, given by

$$\mu_{\mathcal{V}}([\lambda_1, \dots, \lambda_n]^T) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

By definition, a linear relation between the v_i's is just a vector $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T \in \mathbb{R}^n$ such that $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = 0$, or in other words, an element of the kernel of $\mu_{\mathcal{V}}$.

Thus, $\mathcal{V}$ is linearly independent iff $\ker(\mu_{\mathcal{V}}) = \{0\}$ iff $\mu_{\mathcal{V}}$ is injective (by Proposition **??**). ◯

The Wronskian

Definition ??: Let $C^\infty(\mathbb{R})$ be the vector space of smooth functions $f : \mathbb{R} \to \mathbb{R}$.

The Wronskian

Definition ??: Let $C^\infty(\mathbb{R})$ be the vector space of smooth functions $f : \mathbb{R} \to \mathbb{R}$. Given $f_1, \ldots, f_n \in C^\infty(\mathbb{R})$, their *Wronskian matrix* is the matrix $WM(f_1, \ldots, f_n)$ whose entries are the derivatives $f_i^{(j)}$ for $i = 1, \ldots, n$ and $j = 0, \ldots, n-1$.

Definition ??: Let $C^\infty(\mathbb{R})$ be the vector space of smooth functions $f : \mathbb{R} \to \mathbb{R}$. Given $f_1, \dots, f_n \in C^\infty(\mathbb{R})$, their *Wronskian matrix* is the matrix $WM(f_1, \dots, f_n)$ whose entries are the derivatives $f_i^{(j)}$ for $i = 1, \dots, n$ and $j = 0, \dots, n-1$. For example, in the case $n = 4$, we have

$$WM(f_1, f_2, f_3, f_4) = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 \\ f_1' & f_2' & f_3' & f_4' \\ f_1'' & f_2'' & f_3'' & f_4'' \\ f_1''' & f_2''' & f_3''' & f_4''' \end{bmatrix}.$$

Definition ??: Let $C^\infty(\mathbb{R})$ be the vector space of smooth functions $f:\mathbb{R}\to\mathbb{R}$. Given $f_1,\dots,f_n\in C^\infty(\mathbb{R})$, their *Wronskian matrix* is the matrix $WM(f_1,\dots,f_n)$ whose entries are the derivatives $f_i^{(j)}$ for $i=1,\dots,n$ and $j=0,\dots,n-1$. For example, in the case $n=4$, we have

$$WM(f_1,f_2,f_3,f_4)=\begin{bmatrix} f_1 & f_2 & f_3 & f_4 \\ f_1' & f_2' & f_3' & f_4' \\ f_1'' & f_2'' & f_3'' & f_4'' \\ f_1''' & f_2''' & f_3''' & f_4''' \end{bmatrix}.$$

The *Wronskian* of $f_1,\dots,f_n$ is the determinant of the Wronskian matrix; it is written $W(f_1,\dots,f_n)$.

The Wronskian

Definition ??: Let $C^\infty(\mathbb{R})$ be the vector space of smooth functions $f : \mathbb{R} \to \mathbb{R}$. Given $f_1, \ldots, f_n \in C^\infty(\mathbb{R})$, their *Wronskian matrix* is the matrix $WM(f_1, \ldots, f_n)$ whose entries are the derivatives $f_i^{(j)}$ for $i = 1, \ldots, n$ and $j = 0, \ldots, n-1$. For example, in the case $n = 4$, we have

$$WM(f_1, f_2, f_3, f_4) = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 \\ f_1' & f_2' & f_3' & f_4' \\ f_1'' & f_2'' & f_3'' & f_4'' \\ f_1''' & f_2''' & f_3''' & f_4''' \end{bmatrix}.$$

The *Wronskian* of $f_1, \ldots, f_n$ is the determinant of the Wronskian matrix; it is written $W(f_1, \ldots, f_n)$.

Note that the entries in the Wronskian matrix are all functions, so the determinant is again a function. ◯

Wronskian example

Wronskian example

Example ??: Consider the functions exp, sin and cos

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$.

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix}$$

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix}$$

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix}$$

$$= \exp.(-\cos^2 - \sin^2)$$

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix}$$

$$= \exp.(-\cos^2 - \sin^2) - \exp.(-\sin . \cos + \sin . \cos)$$

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix}$$

$$= \exp.(-\cos^2 - \sin^2) - \exp.(-\sin . \cos + \sin . \cos) + \exp.(-\sin^2 - \cos^2)$$

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix}$$

$$= \exp.(-\cos^2 - \sin^2) - \exp.(-\sin\,.\,\cos + \sin\,.\,\cos) + \exp.(-\sin^2 - \cos^2)$$

Wronskian example

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$W(\exp, \sin, \cos) = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix}$$

$$= \exp.(-\cos^2 - \sin^2) - \exp.(-\sin . \cos + \sin . \cos) + \exp.(-\sin^2 - \cos^2)$$

$$= \exp .(-1) - \exp .(0) + \exp .(-1)$$

Example ??: Consider the functions exp, sin and cos, so $\exp' = \exp$ and $\sin' = \cos$ and $\cos' = -\sin$ and $\sin^2 + \cos^2 = 1$. We have

$$
\begin{aligned}
W(\exp, \sin, \cos) &= \det \begin{bmatrix} \exp & \sin & \cos \\ \exp' & \sin' & \cos' \\ \exp'' & \sin'' & \cos'' \end{bmatrix} = \det \begin{bmatrix} \exp & \sin & \cos \\ \exp & \cos & -\sin \\ \exp & -\sin & -\cos \end{bmatrix} \\
&= \exp.(-\cos^2 - \sin^2) - \exp.(-\sin . \cos + \sin . \cos) + \exp.(-\sin^2 - \cos^2) \\
&= \exp .(-1) - \exp .(0) + \exp .(-1) \\
&= -2 \exp .
\end{aligned}
$$

The Wronskian and linear dependence

Proposition ??:
If $f_1, \ldots, f_n$ are linearly dependent, then $W(f_1, \ldots, f_n) = 0$.

Proposition ??:
If $f_1, \ldots, f_n$ are linearly dependent, then $W(f_1, \ldots, f_n) = 0$.
(The function $w = W(f_1, \ldots, f_n)$ is the zero function, ie $w(x) = 0$ for all x.)

Proposition ??:
If $f_1, \ldots, f_n$ are linearly dependent, then $W(f_1, \ldots, f_n) = 0$.
(The function $w = W(f_1, \ldots, f_n)$ is the zero function, ie $w(x) = 0$ for all x.)

Proof for $n = 3$:

Proposition ??:
If $f_1, \ldots, f_n$ are linearly dependent, then $W(f_1, \ldots, f_n) = 0$.
(The function $w = W(f_1, \ldots, f_n)$ is the zero function, ie $w(x) = 0$ for all x.)

Proof for $n = 3$:
If f_1, f_2, f_3 are linearly dependent, then there are numbers $\lambda_1, \lambda_2, \lambda_3$ (not all zero) such that $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$ is the zero function

The Wronskian and linear dependence

Proposition ??:
If $f_1, \ldots, f_n$ are linearly dependent, then $W(f_1, \ldots, f_n) = 0$.
(The function $w = W(f_1, \ldots, f_n)$ is the zero function, ie $w(x) = 0$ for all x.)

Proof for $n = 3$**:**
If f_1, f_2, f_3 are linearly dependent, then there are numbers $\lambda_1, \lambda_2, \lambda_3$ (not all zero) such that $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$ is the zero function, which means that

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0 \qquad \text{(for all } x\text{)}$$

Proposition ??:
If $f_1, \ldots, f_n$ are linearly dependent, then $W(f_1, \ldots, f_n) = 0$.
(The function $w = W(f_1, \ldots, f_n)$ is the zero function, ie $w(x) = 0$ for all x.)

Proof for $n = 3$:
If f_1, f_2, f_3 are linearly dependent, then there are numbers $\lambda_1, \lambda_2, \lambda_3$ (not all zero) such that $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$ is the zero function, which means that

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0 \qquad \text{(for all } x\text{)}$$

We can differentiate to get

$$\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) = 0$$

Proposition ??:
If $f_1, \dots, f_n$ are linearly dependent, then $W(f_1, \dots, f_n) = 0$.
(The function $w = W(f_1, \dots, f_n)$ is the zero function, ie $w(x) = 0$ for all x.)

Proof for $n = 3$:
If f_1, f_2, f_3 are linearly dependent, then there are numbers $\lambda_1, \lambda_2, \lambda_3$ (not all zero) such that $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$ is the zero function, which means that

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0 \qquad \text{(for all } x\text{)}$$

We can differentiate to get

$$\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) = 0$$

and again to get

$$\lambda_1 f_1''(x) + \lambda_2 f_2''(x) + \lambda_3 f_3''(x) = 0$$

The Wronskian and linear dependence

$$\begin{aligned}
\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) &= 0\\
\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) &= 0\\
\lambda_1 f_1''(x) + \lambda_2 f_2''(x) + \lambda_3 f_3''(x) &= 0
\end{aligned}$$

The Wronskian and linear dependence

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0$$
$$\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) = 0$$
$$\lambda_1 f_1''(x) + \lambda_2 f_2''(x) + \lambda_3 f_3''(x) = 0$$

so

$$\lambda_1 \begin{bmatrix} f_1(x) \\ f_1'(x) \\ f_1''(x) \end{bmatrix} + \lambda_2 \begin{bmatrix} f_2(x) \\ f_2'(x) \\ f_2''(x) \end{bmatrix} + \lambda_3 \begin{bmatrix} f_3(x) \\ f_3'(x) \\ f_3''(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The Wronskian and linear dependence

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0$$
$$\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) = 0$$
$$\lambda_1 f_1''(x) + \lambda_2 f_2''(x) + \lambda_3 f_3''(x) = 0$$

so

$$\lambda_1 \begin{bmatrix} f_1(x) \\ f_1'(x) \\ f_1''(x) \end{bmatrix} + \lambda_2 \begin{bmatrix} f_2(x) \\ f_2'(x) \\ f_2''(x) \end{bmatrix} + \lambda_3 \begin{bmatrix} f_3(x) \\ f_3'(x) \\ f_3''(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so the columns of the matrix

$$WM(f_1, f_2, f_3) = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{bmatrix}.$$

are linearly dependent

The Wronskian and linear dependence

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0$$
$$\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) = 0$$
$$\lambda_1 f_1''(x) + \lambda_2 f_2''(x) + \lambda_3 f_3''(x) = 0$$

so

$$\lambda_1 \begin{bmatrix} f_1(x) \\ f_1'(x) \\ f_1''(x) \end{bmatrix} + \lambda_2 \begin{bmatrix} f_2(x) \\ f_2'(x) \\ f_2''(x) \end{bmatrix} + \lambda_3 \begin{bmatrix} f_3(x) \\ f_3'(x) \\ f_3''(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so the columns of the matrix

$$WM(f_1, f_2, f_3) = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{bmatrix}.$$

are linearly dependent, so

$$W(f_1, f_2, f_3) = \det(WM(f_1, f_2, f_3)) = 0.$$

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x) = 0$$
$$\lambda_1 f_1'(x) + \lambda_2 f_2'(x) + \lambda_3 f_3'(x) = 0$$
$$\lambda_1 f_1''(x) + \lambda_2 f_2''(x) + \lambda_3 f_3''(x) = 0$$

so

$$\lambda_1 \begin{bmatrix} f_1(x) \\ f_1'(x) \\ f_1''(x) \end{bmatrix} + \lambda_2 \begin{bmatrix} f_2(x) \\ f_2'(x) \\ f_2''(x) \end{bmatrix} + \lambda_3 \begin{bmatrix} f_3(x) \\ f_3'(x) \\ f_3''(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so the columns of the matrix

$$WM(f_1, f_2, f_3) = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{bmatrix}.$$

are linearly dependent, so

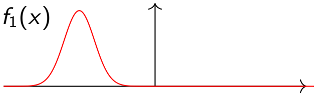
$$W(f_1, f_2, f_3) = \det(WM(f_1, f_2, f_3)) = 0.$$

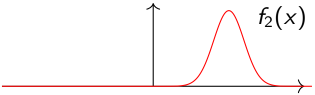
Corollary ??:
If $W(f_1, \ldots, f_n) \neq 0$, then $f_1, \ldots, f_n$ are linearly independent. □○

A Wronskian counterexample

A Wronskian counterexample

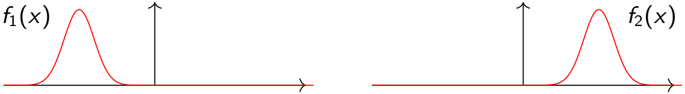
Remark ??: Consider a pair of smooth functions like this:





A Wronskian counterexample

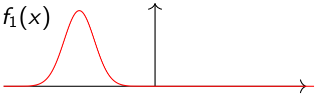
Remark ??: Consider a pair of smooth functions like this:

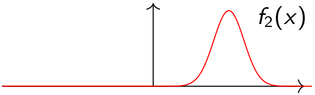


Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:

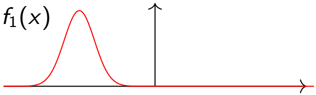


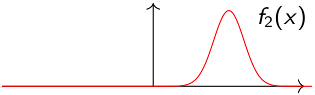


Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$.

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:

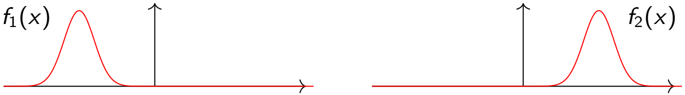




Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.)

A Wronskian counterexample

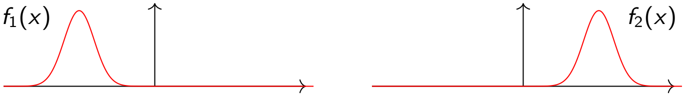
Remark ??: Consider a pair of smooth functions like this:



Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix

$WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$

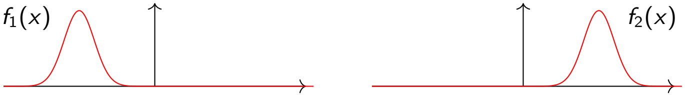
Remark ??: Consider a pair of smooth functions like this:



Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero.

A Wronskian counterexample

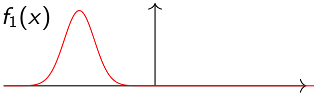
Remark ??: Consider a pair of smooth functions like this:

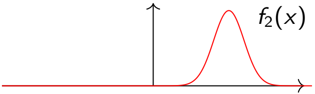


Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero. For $x \geq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} 0 & f_2(x) \\ 0 & f_2'(x) \end{bmatrix}$

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:

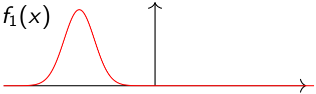


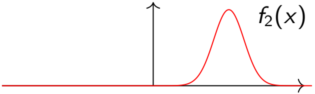


Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero. For $x \geq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} 0 & f_2(x) \\ 0 & f_2'(x) \end{bmatrix}$, so the determinant is again zero.

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:

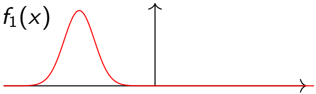


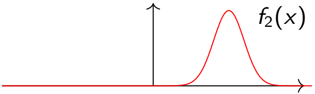


Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero. For $x \geq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} 0 & f_2(x) \\ 0 & f_2'(x) \end{bmatrix}$, so the determinant is again zero. Thus $W(f_1, f_2)(x) = 0$ for all x

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:

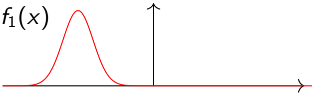


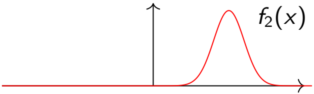


Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero. For $x \geq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} 0 & f_2(x) \\ 0 & f_2'(x) \end{bmatrix}$, so the determinant is again zero. Thus $W(f_1, f_2)(x) = 0$ for all x, but f_1 and f_2 are not linearly dependent.

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:

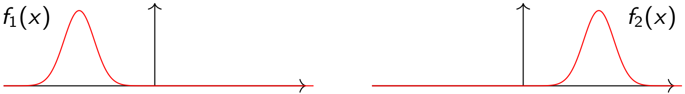




Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero. For $x \geq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} 0 & f_2(x) \\ 0 & f_2'(x) \end{bmatrix}$, so the determinant is again zero. Thus $W(f_1, f_2)(x) = 0$ for all x, but f_1 and f_2 are not linearly dependent. This shows that the test in Proposition **??** is not reversible: if the functions are dependent then the Wronskian vanishes, but if the Wronskian vanishes then the functions need not be dependent.

A Wronskian counterexample

Remark ??: Consider a pair of smooth functions like this:



Suppose that $f_1(x)$ is zero (not just small) for $x \geq 0$, and that $f_2(x)$ is zero for $x \leq 0$. (It is not easy to write down formulae for such functions, but it can be done; we will not discuss this further here.) For $x \leq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} f_1(x) & 0 \\ f_1'(x) & 0 \end{bmatrix}$, so the determinant is zero. For $x \geq 0$, the matrix $WM(f_1, f_2)(x)$ has the form $\begin{bmatrix} 0 & f_2(x) \\ 0 & f_2'(x) \end{bmatrix}$, so the determinant is again zero. Thus $W(f_1, f_2)(x) = 0$ for all x, but f_1 and f_2 are not linearly dependent. This shows that the test in Proposition **??** is not reversible: if the functions are dependent then the Wronskian vanishes, but if the Wronskian vanishes then the functions need not be dependent. In practice it is rare to find such counterexamples, however.

Spanning

Spanning

Definition ??: Given a list $\mathcal{V} = v_1, \ldots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ for some $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Spanning

Definition ??: Given a list $\mathcal{V} = v_1, \ldots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ for some $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Equivalently, span($\mathcal{V}$) is the image of the map $\mu_{\mathcal{V}} : \mathbb{R}^n \to V$

Spanning

Definition ??: Given a list $\mathcal{V} = v_1, \dots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \dots + \lambda_n v_n$ for some $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Equivalently, span($\mathcal{V}$) is the image of the map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ (which shows that span($\mathcal{V}$) is a subspace of V).

Spanning

Definition ??: Given a list $\mathcal{V} = v_1, \dots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \dots + \lambda_n v_n$ for some $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Equivalently, span($\mathcal{V}$) is the image of the map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ (which shows that span($\mathcal{V}$) is a subspace of V). We say that $\mathcal{V}$ *spans* V if span($\mathcal{V}$) $= V$

Definition ??: Given a list $\mathcal{V} = v_1, \ldots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ for some $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Equivalently, span($\mathcal{V}$) is the image of the map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ (which shows that span($\mathcal{V}$) is a subspace of V). We say that $\mathcal{V}$ *spans* V if span($\mathcal{V}$) $= V$, or equivalently, if $\mu_{\mathcal{V}}$ is surjective.

Spanning

Definition ??: Given a list $\mathcal{V} = v_1, \dots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \dots + \lambda_n v_n$ for some $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Equivalently, span($\mathcal{V}$) is the image of the map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ (which shows that span($\mathcal{V}$) is a subspace of V). We say that $\mathcal{V}$ *spans* V if span($\mathcal{V}$) $= V$, or equivalently, if $\mu_{\mathcal{V}}$ is surjective.

Remark ??: Often V will be a subspace of some larger space U.

Spanning

Definition ??: Given a list $\mathcal{V} = v_1, \dots, v_n$ of elements of a vector space V, we write span($\mathcal{V}$) for the set of all vectors $w \in V$ that can be written in the form $w = \lambda_1 v_1 + \dots + \lambda_n v_n$ for some $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Equivalently, span($\mathcal{V}$) is the image of the map $\mu_{\mathcal{V}} : \mathbb{R}^n \to V$ (which shows that span($\mathcal{V}$) is a subspace of V). We say that $\mathcal{V}$ *spans* V if span($\mathcal{V}$) $= V$, or equivalently, if $\mu_{\mathcal{V}}$ is surjective.

Remark ??: Often V will be a subspace of some larger space U. If you are asked whether certain vectors $v_1, \dots, v_n$ span V, the *first* thing that you have to check is that they are actually elements of V. ○

The standard basis spans

The standard basis spans

Definition ??: Let $\mathbf{e}_i$ be the vector in $\mathbb{R}^n$ whose i'th entry is 1, with all other entries being zero.

The standard basis spans

Definition ??: Let $\mathbf{e}_i$ be the vector in $\mathbb{R}^n$ whose i'th entry is 1, with all other entries being zero. For example, in $\mathbb{R}^3$ we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The standard basis spans

Definition ??: Let $\mathbf{e}_i$ be the vector in $\mathbb{R}^n$ whose i'th entry is 1, with all other entries being zero. For example, in $\mathbb{R}^3$ we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Example ??: The list $\mathbf{e}_1, \dots, \mathbf{e}_n$ spans $\mathbb{R}^n$.

The standard basis spans

Definition ??: Let $\mathbf{e}_i$ be the vector in $\mathbb{R}^n$ whose i'th entry is 1, with all other entries being zero. For example, in $\mathbb{R}^3$ we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Example ??: The list $\mathbf{e}_1, \ldots, \mathbf{e}_n$ spans $\mathbb{R}^n$. Indeed, any vector $\mathbf{x} \in \mathbb{R}^n$ can be written as $x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$, which is a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$, as required.

The standard basis spans

Definition ??: Let $\mathbf{e}_i$ be the vector in $\mathbb{R}^n$ whose i'th entry is 1, with all other entries being zero. For example, in $\mathbb{R}^3$ we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Example ??: The list $\mathbf{e}_1, \dots, \mathbf{e}_n$ spans $\mathbb{R}^n$. Indeed, any vector $\mathbf{x} \in \mathbb{R}^n$ can be written as $x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$, which is a linear combination of $\mathbf{e}_1, \dots, \mathbf{e}_n$, as required. For example, in $\mathbb{R}^3$ we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

Monomials span $\mathbb{R}[x]$

Monomials span $\mathbb{R}[x]$

Example ??: The list $1, x, \ldots, x^n$ spans $\mathbb{R}[x]_{\leq n}$.

Monomials span $\mathbb{R}[x]$

Example ??: The list $1, x, \ldots, x^n$ spans $\mathbb{R}[x]_{\leq n}$.

Indeed, any element of $\mathbb{R}[x]_{\leq n}$ is a polynomial of the form
$f(x) = a_0 + a_1 x + \cdots + a_n x^n$

Monomials span $\mathbb{R}[x]$

Example ??: The list $1, x, \ldots, x^n$ spans $\mathbb{R}[x]_{\leq n}$.

Indeed, any element of $\mathbb{R}[x]_{\leq n}$ is a polynomial of the form $f(x) = a_0 + a_1 x + \cdots + a_n x^n$, and so is visibly a linear combination of $1, x, \ldots, x^n$. $\bigcirc$

A spanning set for $\mathbb{R}^4$

A spanning set for $\mathbb{R}^4$

Example ??: Consider the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

A spanning set for $\mathbb{R}^4$

Example ??: Consider the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

We claim that these span $\mathbb{R}^4$.

A spanning set for $\mathbb{R}^4$

Example ??: Consider the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

We claim that these span $\mathbb{R}^4$. Indeed, consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$.

A spanning set for $\mathbb{R}^4$

Example ??: Consider the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

We claim that these span $\mathbb{R}^4$. Indeed, consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We have

$$(a - c + d)\mathbf{u}_1 + (c - d)\mathbf{u}_2 + (c - a)\mathbf{u}_3 + (b - c)\mathbf{u}_3 = \begin{bmatrix} a-c+d \\ a-c+d \\ a-c+d \\ a-c+d \end{bmatrix} + \begin{bmatrix} c-d \\ c-d \\ c-d \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c-a \\ c-a \\ c-a \end{bmatrix} + \begin{bmatrix} 0 \\ b-c \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \mathbf{v}$$

A spanning set for $\mathbb{R}^4$

Example ??: Consider the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

We claim that these span $\mathbb{R}^4$. Indeed, consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We have

$$(a - c + d)\mathbf{u}_1 + (c - d)\mathbf{u}_2 + (c - a)\mathbf{u}_3 + (b - c)\mathbf{u}_3 = \begin{bmatrix} a-c+d \\ a-c+d \\ a-c+d \\ a-c+d \end{bmatrix} + \begin{bmatrix} c-d \\ c-d \\ c-d \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c-a \\ c-a \\ c-a \end{bmatrix} + \begin{bmatrix} 0 \\ b-c \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \mathbf{v}$$

which shows that $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_4$, as required. ◯

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$.

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}$$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}, \text{ or}$$

$$p+q = a \ (1) \qquad p+q+r+s = b \ (2) \qquad p+q+r = c \ (3) \qquad p+r = d \ (4)$$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}, \text{ or}$$

$p + q = a$ (1) $\qquad p + q + r + s = b$ (2) $\qquad p + q + r = c$ (3) $\qquad p + r = d$ (4)

Subtracting (3) and (4) gives $q = c - d$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}, \text{ or}$$

$$p+q = a \ (1) \qquad p+q+r+s = b \ (2) \qquad p+q+r = c \ (3) \qquad p+r = d \ (4)$$

Subtracting (3) and (4) gives $q = c - d$; Subtracting (1) and(3) gives $r = c - a$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}, \text{ or}$$

$$p + q = a \ (1) \qquad p + q + r + s = b \ (2) \qquad p + q + r = c \ (3) \qquad p + r = d \ (4)$$

Subtracting (3) and (4) gives $q = c - d$; Subtracting (1) and(3) gives $r = c - a$; Subtracting (2) and(3) gives $s = b - c$

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}, \text{ or}$$

$p+q = a$ (1) $\qquad$ $p+q+r+s = b$ (2) $\qquad$ $p+q+r = c$ (3) $\qquad$ $p+r = d$ (4)

Subtracting (3) and (4) gives $q = c - d$; Subtracting (1) and(3) gives $r = c - a$; Subtracting (2) and(3) gives $s = b - c$; putting $q = c - d$ in (1) gives $p = a - c + d$.

A spanning set for $\mathbb{R}^4$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Consider an arbitrary vector $\mathbf{v} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T \in \mathbb{R}^4$. We want to find p, q, r, s such that $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2 + r\mathbf{u}_3 + s\mathbf{u}_4$, or equivalently

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = p \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + q \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p+q \\ p+q+r+s \\ p+q+r \\ p+r \end{bmatrix}, \text{ or}$$

$$p+q = a \ (1) \qquad p+q+r+s = b \ (2) \qquad p+q+r = c \ (3) \qquad p+r = d \ (4)$$

Subtracting (3) and (4) gives $q = c - d$; Subtracting (1) and(3) gives $r = c - a$; Subtracting (2) and(3) gives $s = b - c$; putting $q = c - d$ in (1) gives $p = a - c + d$.

$$(a-c+d)\mathbf{u}_1 + (c-d)\mathbf{u}_2 + (c-a)\mathbf{u}_3 + (b-c)\mathbf{u}_3 = \begin{bmatrix} a-c+d \\ a-c+d \\ a-c+d \\ a-c+d \end{bmatrix} + \begin{bmatrix} c-d \\ c-d \\ c-d \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c-a \\ c-a \\ c-a \end{bmatrix} + \begin{bmatrix} 0 \\ b-c \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \mathbf{v}$$

A spanning set for quadratic polynomials

A spanning set for quadratic polynomials

Example ??: Consider the polynomials $p_i(x) = (x+i)^2$.

A spanning set for quadratic polynomials

Example ??: Consider the polynomials $p_i(x) = (x + i)^2$.

We claim that the list $p_{-2}, p_{-1}, p_0, p_1, p_2$ spans $\mathbb{R}[x]_{\leq 2}$.

A spanning set for quadratic polynomials

Example ??: Consider the polynomials $p_i(x) = (x+i)^2$.

We claim that the list $p_{-2}, p_{-1}, p_0, p_1, p_2$ spans $\mathbb{R}[x]_{\leq 2}$. Indeed, we have

$$\begin{aligned}
p_0(x) &= x^2 \\
p_1(x) - p_{-1}(x) &= (x+1)^2 - (x-1)^2 = 4x \\
p_2(x) + p_{-2}(x) - 2p_0(x) &= (x+2)^2 + (x-2)^2 - 2x^2 = 8.
\end{aligned}$$

A spanning set for quadratic polynomials

Example ??: Consider the polynomials $p_i(x) = (x+i)^2$.

We claim that the list $p_{-2}, p_{-1}, p_0, p_1, p_2$ spans $\mathbb{R}[x]_{\leq 2}$. Indeed, we have

$$\begin{aligned} p_0(x) &= x^2 \\ p_1(x) - p_{-1}(x) &= (x+1)^2 - (x-1)^2 = 4x \\ p_2(x) + p_{-2}(x) - 2p_0(x) &= (x+2)^2 + (x-2)^2 - 2x^2 = 8. \end{aligned}$$

Thus for an arbitrary quadratic polynomial $f(x) = ax^2 + bx + c$, we have

$$\begin{aligned} f(x) &= ap_0(x) + \tfrac{1}{4}b(p_1(x) - p_{-1}(x)) + \tfrac{1}{8}c(p_2(x) + p_{-2}(x) - 2p_0(x)) \\ &= \tfrac{c}{8}p_{-2}(x) - \tfrac{b}{4}p_{-1}(x) + (a - \tfrac{c}{4})p_0(x) + \tfrac{b}{4}p_1(x) + \tfrac{c}{8}p_2(x). \bigcirc \end{aligned}$$

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$.

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

Simple harmonic motion

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$.

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$.

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$.

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$. First, we have

$$g(0) = f(0) - a\sin(0) - b\cos(0) = b - a.0 - b.1 = 0$$

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$. First, we have

$$g(0) = f(0) - a\sin(0) - b\cos(0) = b - a.0 - b.1 = 0$$
$$g'(0) = f'(0) - a\sin'(0) - b\cos'(0)$$

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$. First, we have

$$g(0) = f(0) - a\sin(0) - b\cos(0) = b - a.0 - b.1 = 0$$
$$g'(0) = f'(0) - a\sin'(0) - b\cos'(0) = a - a\cos(0) + b\sin(0)$$

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$. First, we have

$$g(0) = f(0) - a\sin(0) - b\cos(0) = b - a.0 - b.1 = 0$$
$$g'(0) = f'(0) - a\sin'(0) - b\cos'(0) = a - a\cos(0) + b\sin(0) = a - a.1 - b.0 = 0.$$

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$. First, we have

$$g(0) = f(0) - a\sin(0) - b\cos(0) = b - a.0 - b.1 = 0$$
$$g'(0) = f'(0) - a\sin'(0) - b\cos'(0) = a - a\cos(0) + b\sin(0) = a - a.1 - b.0 = 0.$$

Now put $h(x) = g(x)^2 + g'(x)^2$

Simple harmonic motion

Example ??: Put $V = \{f \in C^\infty(\mathbb{R}) \mid f'' + f = 0\}$. **Claim:** the functions sin and cos span V.

In other words, if f has $f''(x) = -f(x)$ for all x, then there are constants a and b such that $f(x) = a\sin(x) + b\cos(x)$ for all x.

Proof: Firstly, we have $\sin' = \cos$ and $\cos' = -\sin$, so $\sin'' = -\sin$ and $\cos'' = -\cos$, so sin and cos are indeed elements of V.

Consider an arbitrary element $f \in V$. Put $a = f'(0)$ and $b = f(0)$, and put $g(x) = f(x) - a\sin(x) - b\cos(x)$. We claim that $g = 0$. First, we have

$$g(0) = f(0) - a\sin(0) - b\cos(0) = b - a.0 - b.1 = 0$$
$$g'(0) = f'(0) - a\sin'(0) - b\cos'(0) = a - a\cos(0) + b\sin(0) = a - a.1 - b.0 = 0.$$

Now put $h(x) = g(x)^2 + g'(x)^2$; the above shows that $h(0) = 0^2 + 0^2 = 0$.

Simple harmonic motion

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Simple harmonic motion

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant, but $h(0) = 0$

Simple harmonic motion

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant, but $h(0) = 0$, so $h(x) = 0$ for all x.

Simple harmonic motion

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant, but $h(0) = 0$, so $h(x) = 0$ for all x.

However, $h(x) = g(x)^2 + g'(x)^2$, which is the sum of two nonnegative quantities; the only way we can have $h(x) = 0$ is if $g(x) = 0 = g'(x)$.

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant, but $h(0) = 0$, so $h(x) = 0$ for all x.

However, $h(x) = g(x)^2 + g'(x)^2$, which is the sum of two nonnegative quantities; the only way we can have $h(x) = 0$ is if $g(x) = 0 = g'(x)$. This means that $g = 0$

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant, but $h(0) = 0$, so $h(x) = 0$ for all x.

However, $h(x) = g(x)^2 + g'(x)^2$, which is the sum of two nonnegative quantities; the only way we can have $h(x) = 0$ is if $g(x) = 0 = g'(x)$. This means that $g = 0$, so $f(x) - a\sin(x) - b\cos(x) = 0$

Simple harmonic motion

$g(x) = f(x) - a\sin(x) - b\cos(x)$; $g \in V$ so $g''(x) + g(x) = 0$;
$g(0) = g'(0) = 0$;

$h(x) = g(x)^2 + g'(x)^2$; $h(0) = 0$

Next, we have $g \in V$, so $g'' = -g$, so

$$h'(x) = 2g(x)g'(x) + 2g'(x)g''(x) = 2g'(x)(g(x) + g''(x)) = 0.$$

This means that h is constant, but $h(0) = 0$, so $h(x) = 0$ for all x.

However, $h(x) = g(x)^2 + g'(x)^2$, which is the sum of two nonnegative quantities; the only way we can have $h(x) = 0$ is if $g(x) = 0 = g'(x)$. This means that $g = 0$, so $f(x) - a\sin(x) - b\cos(x) = 0$, so $f(x) = a\sin(x) + b\cos(x)$, as required. ○

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V that spans V.

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional.

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional. To see this, consider a list $\mathcal{P} = p_1, \dots, p_n$ of polynomials.

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional. To see this, consider a list $\mathcal{P} = p_1, \dots, p_n$ of polynomials. Let d be the maximum of the degrees of $p_1, \dots, p_n$.

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional. To see this, consider a list $\mathcal{P} = p_1, \dots, p_n$ of polynomials. Let d be the maximum of the degrees of $p_1, \dots, p_n$. Then p_i lies in $\mathbb{R}[x]_{\leq d}$ for all i

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional. To see this, consider a list $\mathcal{P} = p_1, \dots, p_n$ of polynomials. Let d be the maximum of the degrees of $p_1, \dots, p_n$. Then p_i lies in $\mathbb{R}[x]_{\leq d}$ for all i, so the span of $\mathcal{P}$ is contained in $\mathbb{R}[x]_{\leq d}$.

Finite-dimensional spaces

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional. To see this, consider a list $\mathcal{P} = p_1, \dots, p_n$ of polynomials. Let d be the maximum of the degrees of $p_1, \dots, p_n$. Then p_i lies in $\mathbb{R}[x]_{\leq d}$ for all i, so the span of $\mathcal{P}$ is contained in $\mathbb{R}[x]_{\leq d}$. In particular, the polynomial x^{d+1} does not lie in $\text{span}(\mathcal{P})$

Definition ??: A vector space V is *finite-dimensional* if there is a (finite) list $\mathcal{V} = v_1, \dots, v_n$ of elements of V that spans V.

Example ??: Using our earlier examples of spanning sets, we see that the spaces $\mathbb{R}^n$, $M_{n,m}\mathbb{R}$ and $\mathbb{R}[x]_{\leq n}$ are finite-dimensional.

Example ??: The space $\mathbb{R}[x]$ is not finite-dimensional. To see this, consider a list $\mathcal{P} = p_1, \dots, p_n$ of polynomials. Let d be the maximum of the degrees of $p_1, \dots, p_n$. Then p_i lies in $\mathbb{R}[x]_{\leq d}$ for all i, so the span of $\mathcal{P}$ is contained in $\mathbb{R}[x]_{\leq d}$. In particular, the polynomial x^{d+1} does not lie in span$(\mathcal{P})$, so $\mathcal{P}$ does not span all of $\mathbb{R}[x]$.○

Bases

Definition ??: A *basis* for a vector space V is a list $\mathcal{V}$ of elements of V that is linearly independent and also spans V.

Bases

Definition ??: A *basis* for a vector space V is a list $\mathcal{V}$ of elements of V that is linearly independent and also spans V. Equivalently, a list $\mathcal{V} = v_1, \ldots, v_n$ is a basis iff the map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ is a bijection.

Antisymmetric matrices

Example ??: We will find a basis for the space V of antisymmetric 3×3 matrices.

Antisymmetric matrices

Example ??: We will find a basis for the space V of antisymmetric 3×3 matrices. Such a matrix has the form

$$X = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

Example ??: We will find a basis for the space V of antisymmetric 3×3 matrices. Such a matrix has the form

$$X = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

In other words, if we put

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

then any antisymmetric matrix X can be written in the form $X = aA + bB + cC$.

Antisymmetric matrices

Example ??: We will find a basis for the space V of antisymmetric 3×3 matrices. Such a matrix has the form

$$X = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

In other words, if we put

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

then any antisymmetric matrix X can be written in the form $X = aA + bB + cC$. This means that the matrices A, B and C span V

Antisymmetric matrices

Example ??: We will find a basis for the space V of antisymmetric 3×3 matrices. Such a matrix has the form

$$X = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

In other words, if we put

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

then any antisymmetric matrix X can be written in the form $X = aA + bB + cC$. This means that the matrices A, B and C span V, and they are clearly independent

Example ??: We will find a basis for the space V of antisymmetric 3×3 matrices. Such a matrix has the form

$$X = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

In other words, if we put

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

then any antisymmetric matrix X can be written in the form $X = aA + bB + cC$. This means that the matrices A, B and C span V, and they are clearly independent, so they form a basis.○

Trace-free symmetric matrices

Trace-free symmetric matrices

Put $V = \{A \in M_3R \mid A^T = A \text{ and } \operatorname{trace}(A) = 0\}$.

Trace-free symmetric matrices

Put $V = \{A \in M_3R \mid A^T = A \text{ and } \operatorname{trace}(A) = 0\}$.

Any matrix $X \in V$ has the form

$$X = \left[\begin{smallmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{smallmatrix}\right]$$

for some $a, b, c, d, e \in \mathbb{R}$.

Trace-free symmetric matrices

Put $V = \{A \in M_3R \mid A^T = A \text{ and } \operatorname{trace}(A) = 0\}$.

Any matrix $X \in V$ has the form

$$X = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$$

for some $a, b, c, d, e \in \mathbb{R}$. In other words, if we put

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Trace-free symmetric matrices

Put $V = \{A \in M_3R \mid A^T = A \text{ and } \operatorname{trace}(A) = 0\}$.

Any matrix $X \in V$ has the form

$$X = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$$

for some $a, b, c, d, e \in \mathbb{R}$. In other words, if we put

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

then any matrix $X \in V$ can be written in the form

$$X = aA + bB + cC + dD + eE.$$

Trace-free symmetric matrices

Put $V = \{A \in M_3R \mid A^T = A \text{ and } \operatorname{trace}(A) = 0\}$.

Any matrix $X \in V$ has the form

$$X = \left[\begin{smallmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{smallmatrix}\right]$$

for some $a, b, c, d, e \in \mathbb{R}$. In other words, if we put

$$A = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{smallmatrix}\right] \quad B = \left[\begin{smallmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right] \quad C = \left[\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{smallmatrix}\right] \quad D = \left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{smallmatrix}\right] \quad E = \left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{smallmatrix}\right]$$

then any matrix $X \in V$ can be written in the form

$$X = aA + bB + cC + dD + eE.$$

This means that the matrices $A, \ldots, E$ span V

Trace-free symmetric matrices

Put $V = \{A \in M_3R \mid A^T = A \text{ and } \operatorname{trace}(A) = 0\}$.

Any matrix $X \in V$ has the form

$$X = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$$

for some $a, b, c, d, e \in \mathbb{R}$. In other words, if we put

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

then any matrix $X \in V$ can be written in the form

$$X = aA + bB + cC + dD + eE.$$

This means that the matrices $A, \dots, E$ span V, and they are also linearly independent, so they form a basis for V. ◯

Bases for the space of quadratic polynomials

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two.

Bases for the space of quadratic polynomials

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two. A typical element $f \in Q$ has $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$.

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two. A typical element $f \in Q$ has $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$.

- The list p_0, p_1, p_2, where $p_i(x) = x^i$. This is the most obvious basis.

Bases for the space of quadratic polynomials

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two. A typical element $f \in Q$ has $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$.

- The list p_0, p_1, p_2, where $p_i(x) = x^i$. This is the most obvious basis. For f as above we have

$$f = c\,p_0 + b\,p_1 + a\,p_2 = f(0)\,p_0 + f'(0)\,p_1 + \tfrac{1}{2}f''(0)p_2.$$

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two. A typical element $f \in Q$ has $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$.

- The list p_0, p_1, p_2, where $p_i(x) = x^i$. This is the most obvious basis. For f as above we have

$$f = c\,p_0 + b\,p_1 + a\,p_2 = f(0)\,p_0 + f'(0)\,p_1 + \tfrac{1}{2}f''(0)p_2.$$

- The list q_0, q_1, q_2, where $q_i(x) = (x+1)^i$, is another basis.

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two. A typical element $f \in Q$ has $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$.

- The list p_0, p_1, p_2, where $p_i(x) = x^i$. This is the most obvious basis. For f as above we have

$$f = c\, p_0 + b\, p_1 + a\, p_2 = f(0)\, p_0 + f'(0)\, p_1 + \tfrac{1}{2} f''(0) p_2.$$

- The list q_0, q_1, q_2, where $q_i(x) = (x+1)^i$, is another basis. For f as above, one checks that

$$ax^2 + bx + c = a(x+1)^2 + (b-2a)(x+1) + (a-b+c)$$

Bases for the space of quadratic polynomials

Example ??: There are several interesting bases for the space $Q = \mathbb{R}[x]_{\leq 2}$ of polynomials of degree at most two. A typical element $f \in Q$ has $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$.

- The list p_0, p_1, p_2, where $p_i(x) = x^i$. This is the most obvious basis. For f as above we have

$$f = c\, p_0 + b\, p_1 + a\, p_2 = f(0)\, p_0 + f'(0)\, p_1 + \tfrac{1}{2} f''(0) p_2.$$

- The list q_0, q_1, q_2, where $q_i(x) = (x+1)^i$, is another basis. For f as above, one checks that

$$ax^2 + bx + c = a(x+1)^2 + (b-2a)(x+1) + (a-b+c)$$

so

$$f = (a-b+c)q_0 + (b-2a)q_1 + a\, q_2 = f(-1)q_0 + f'(-1)q_1 + \tfrac{1}{2} f''(-1)q_2.$$

Bases for the space of quadratic polynomials

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis.

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$p_0(x) = 1 = \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2)$$

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$\begin{aligned} p_0(x) = 1 &= \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2) \\ &= \tfrac{1}{2}(r_2(x) - 2r_1(x) + r_0(x)) \end{aligned}$$

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$\begin{aligned} p_0(x) = 1 &= \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2) \\ &= \tfrac{1}{2}(r_2(x) - 2r_1(x) + r_0(x)) \\ p_1(x) = x &= -\tfrac{1}{4}((x+2)^2 - 4(x+1)^2 + 3x^2) \end{aligned}$$

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$\begin{aligned} p_0(x) &= 1 = \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2) \\ &= \tfrac{1}{2}(r_2(x) - 2r_1(x) + r_0(x)) \\ p_1(x) &= x = -\tfrac{1}{4}((x+2)^2 - 4(x+1)^2 + 3x^2) \\ &= -\tfrac{1}{4}(r_2(x) - 4r_1(x) + 3r_0(x)) \end{aligned}$$

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$\begin{aligned} p_0(x) &= 1 = \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2) \\ &= \tfrac{1}{2}(r_2(x) - 2r_1(x) + r_0(x)) \\ p_1(x) &= x = -\tfrac{1}{4}((x+2)^2 - 4(x+1)^2 + 3x^2) \\ &= -\tfrac{1}{4}(r_2(x) - 4r_1(x) + 3r_0(x)) \\ p_2(x) &= x^2 = r_0(x). \end{aligned}$$

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$\begin{aligned} p_0(x) &= 1 = \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2) \\ &= \tfrac{1}{2}(r_2(x) - 2r_1(x) + r_0(x)) \\ p_1(x) &= x = -\tfrac{1}{4}((x+2)^2 - 4(x+1)^2 + 3x^2) \\ &= -\tfrac{1}{4}(r_2(x) - 4r_1(x) + 3r_0(x)) \\ p_2(x) &= x^2 = r_0(x). \end{aligned}$$

  This implies that $p_0, p_1, p_2 \in \mathrm{span}(r_0, r_1, r_2)$

Bases for the space of quadratic polynomials

- The list r_0, r_1, r_2, where $r_i(x) = (x+i)^2$, is another basis. Indeed, we have

$$\begin{aligned} p_0(x) &= 1 = \tfrac{1}{2}((x+2)^2 - 2(x+1)^2 + x^2) \\ &= \tfrac{1}{2}(r_2(x) - 2r_1(x) + r_0(x)) \\ p_1(x) &= x = -\tfrac{1}{4}((x+2)^2 - 4(x+1)^2 + 3x^2) \\ &= -\tfrac{1}{4}(r_2(x) - 4r_1(x) + 3r_0(x)) \\ p_2(x) &= x^2 = r_0(x). \end{aligned}$$

This implies that $p_0, p_1, p_2 \in \operatorname{span}(r_0, r_1, r_2)$ and thus that $\operatorname{span}(r_0, r_1, r_2) = Q$. ◯

Bases for the space of quadratic polynomials

Bases for the space of quadratic polynomials

- The list

$$\begin{aligned} s_0(x) &= (x^2 - 3x + 2)/2 \\ s_1(x) &= -x^2 + 2x \\ s_2(x) &= (x^2 - x)/2. \end{aligned}$$

  gives another basis.

Bases for the space of quadratic polynomials

- The list

$$\begin{aligned} s_0(x) &= (x^2 - 3x + 2)/2 \\ s_1(x) &= -x^2 + 2x \\ s_2(x) &= (x^2 - x)/2. \end{aligned}$$

gives another basis. These functions have the property that

$$\begin{array}{lll} s_0(0) = 1 & s_0(1) = 0 & s_0(2) = 0 \\ s_1(0) = 0 & s_1(1) = 1 & s_1(2) = 0 \\ s_2(0) = 0 & s_2(1) = 0 & s_2(2) = 1 \end{array}$$

Bases for the space of quadratic polynomials

- The list

$$\begin{aligned} s_0(x) &= (x^2 - 3x + 2)/2 \\ s_1(x) &= -x^2 + 2x \\ s_2(x) &= (x^2 - x)/2. \end{aligned}$$

gives another basis. These functions have the property that

$$\begin{array}{lll} s_0(0) = 1 & s_0(1) = 0 & s_0(2) = 0 \\ s_1(0) = 0 & s_1(1) = 1 & s_1(2) = 0 \\ s_2(0) = 0 & s_2(1) = 0 & s_2(2) = 1 \end{array}$$

Given $f \in Q$ we claim that $f = f(0).s_0 + f(1).s_1 + f(2).s_2$.

Bases for the space of quadratic polynomials

- The list

$$s_0(x) = (x^2 - 3x + 2)/2$$
$$s_1(x) = -x^2 + 2x$$
$$s_2(x) = (x^2 - x)/2.$$

gives another basis. These functions have the property that

$$\begin{array}{lll} s_0(0) = 1 & s_0(1) = 0 & s_0(2) = 0 \\ s_1(0) = 0 & s_1(1) = 1 & s_1(2) = 0 \\ s_2(0) = 0 & s_2(1) = 0 & s_2(2) = 1 \end{array}$$

Given $f \in Q$ we claim that $f = f(0).s_0 + f(1).s_1 + f(2).s_2$. Indeed, if we put $g(x) = f(x) - f(0)s_0(x) - f(1)s_1(x) - f(2).s_2(x)$, we find that $g \in Q$ and $g(0) = g(1) = g(2) = 0$.

Bases for the space of quadratic polynomials

- The list

$$
\begin{aligned}
s_0(x) &= (x^2 - 3x + 2)/2 \\
s_1(x) &= -x^2 + 2x \\
s_2(x) &= (x^2 - x)/2.
\end{aligned}
$$

gives another basis. These functions have the property that

$$
\begin{array}{lll}
s_0(0) = 1 & s_0(1) = 0 & s_0(2) = 0 \\
s_1(0) = 0 & s_1(1) = 1 & s_1(2) = 0 \\
s_2(0) = 0 & s_2(1) = 0 & s_2(2) = 1
\end{array}
$$

Given $f \in Q$ we claim that $f = f(0).s_0 + f(1).s_1 + f(2).s_2$. Indeed, if we put $g(x) = f(x) - f(0)s_0(x) - f(1)s_1(x) - f(2).s_2(x)$, we find that $g \in Q$ and $g(0) = g(1) = g(2) = 0$. A quadratic polynomial with three different roots must be zero

Bases for the space of quadratic polynomials

- The list

$$s_0(x) = (x^2 - 3x + 2)/2$$
$$s_1(x) = -x^2 + 2x$$
$$s_2(x) = (x^2 - x)/2.$$

gives another basis. These functions have the property that

$$\begin{array}{lll} s_0(0) = 1 & s_0(1) = 0 & s_0(2) = 0 \\ s_1(0) = 0 & s_1(1) = 1 & s_1(2) = 0 \\ s_2(0) = 0 & s_2(1) = 0 & s_2(2) = 1 \end{array}$$

Given $f \in Q$ we claim that $f = f(0).s_0 + f(1).s_1 + f(2).s_2$. Indeed, if we put $g(x) = f(x) - f(0)s_0(x) - f(1)s_1(x) - f(2).s_2(x)$, we find that $g \in Q$ and $g(0) = g(1) = g(2) = 0$. A quadratic polynomial with three different roots must be zero, so $g = 0$

- The list

$$s_0(x) = (x^2 - 3x + 2)/2$$
$$s_1(x) = -x^2 + 2x$$
$$s_2(x) = (x^2 - x)/2.$$

gives another basis. These functions have the property that

$$\begin{array}{lll} s_0(0) = 1 & s_0(1) = 0 & s_0(2) = 0 \\ s_1(0) = 0 & s_1(1) = 1 & s_1(2) = 0 \\ s_2(0) = 0 & s_2(1) = 0 & s_2(2) = 1 \end{array}$$

Given $f \in Q$ we claim that $f = f(0).s_0 + f(1).s_1 + f(2).s_2$. Indeed, if we put $g(x) = f(x) - f(0)s_0(x) - f(1)s_1(x) - f(2).s_2(x)$, we find that $g \in Q$ and $g(0) = g(1) = g(2) = 0$. A quadratic polynomial with three different roots must be zero, so $g = 0$, so $f = f(0).s_0 + f(1).s_1 + f(2).s_2$. ○

Bases for the space of quadratic polynomials

Bases for the space of quadratic polynomials

- The list

$$t_0(x) = 1$$
$$t_1(x) = \sqrt{3}(2x - 1)$$
$$t_2(x) = \sqrt{5}(6x^2 - 6x + 1).$$

gives another basis.

- The list

$$\begin{aligned} t_0(x) &= 1 \\ t_1(x) &= \sqrt{3}(2x-1) \\ t_2(x) &= \sqrt{5}(6x^2-6x+1). \end{aligned}$$

gives another basis. These functions have the property that

$$\int_0^1 t_i(x)t_j(x)\,dx = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j. \end{cases}$$

- The list

$$\begin{aligned} t_0(x) &= 1 \\ t_1(x) &= \sqrt{3}(2x-1) \\ t_2(x) &= \sqrt{5}(6x^2-6x+1). \end{aligned}$$

gives another basis. These functions have the property that

$$\int_0^1 t_i(x)t_j(x)\,dx = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j. \end{cases}$$

Using this, we find that $f = \lambda_0 t_0 + \lambda_1 t_1 + \lambda_2 t_2$, where $\lambda_i = \int_0^1 f(x)t_i(x)\,dx$.

A space of polynomials

A space of polynomials

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

A space of polynomials

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

Consider a polynomial $f \in \mathbb{R}[x]_{\leq 4}$, so $f(x) = a + bx + cx^2 + dx^3 + ex^4$ for some constants $a, \ldots, e$.

A space of polynomials

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

Consider a polynomial $f \in \mathbb{R}[x]_{\leq 4}$, so $f(x) = a + bx + cx^2 + dx^3 + ex^4$ for some constants $a, \dots, e$. We then have

$$\begin{aligned} f(1) &= a + b + c + d + e \\ f(-1) &= a - b + c - d + e \\ f'(1) - f'(-1) &= (b + 2c + 3d + 4e) - (b - 2c + 3d - 4e) = 4c + 8e \end{aligned}$$

A space of polynomials

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

Consider a polynomial $f \in \mathbb{R}[x]_{\leq 4}$, so $f(x) = a + bx + cx^2 + dx^3 + ex^4$ for some constants $a, \dots, e$. We then have

$$\begin{aligned} f(1) &= a + b + c + d + e \\ f(-1) &= a - b + c - d + e \\ f'(1) - f'(-1) &= (b + 2c + 3d + 4e) - (b - 2c + 3d - 4e) = 4c + 8e \end{aligned}$$

It follows that $f \in V$ iff $a + b + c + d + e = a - b + c - d + e = 4c + 8e = 0$.

A space of polynomials

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

Consider a polynomial $f \in \mathbb{R}[x]_{\leq 4}$, so $f(x) = a + bx + cx^2 + dx^3 + ex^4$ for some constants $a, \dots, e$. We then have

$$\begin{aligned} f(1) &= a + b + c + d + e \\ f(-1) &= a - b + c - d + e \\ f'(1) - f'(-1) &= (b + 2c + 3d + 4e) - (b - 2c + 3d - 4e) = 4c + 8e \end{aligned}$$

It follows that $f \in V$ iff $a + b + c + d + e = a - b + c - d + e = 4c + 8e = 0$. This simplifies to $c = -2e$ and $a = e$ and $b = -d$

A space of polynomials

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

Consider a polynomial $f \in \mathbb{R}[x]_{\leq 4}$, so $f(x) = a + bx + cx^2 + dx^3 + ex^4$ for some constants $a, \dots, e$. We then have

$$\begin{aligned} f(1) &= a + b + c + d + e \\ f(-1) &= a - b + c - d + e \\ f'(1) - f'(-1) &= (b + 2c + 3d + 4e) - (b - 2c + 3d - 4e) = 4c + 8e \end{aligned}$$

It follows that $f \in V$ iff $a + b + c + d + e = a - b + c - d + e = 4c + 8e = 0$. This simplifies to $c = -2e$ and $a = e$ and $b = -d$, so

$$f(x) = e - dx - 2ex^2 + dx^3 + ex^4 = d(x^3 - x) + e(x^4 - 2x^2 + 1).$$

Put $V = \{f \in \mathbb{R}[x]_{\leq 4} \mid f(1) = f(-1) = 0 \text{ and } f'(1) = f'(-1)\}$.

Consider a polynomial $f \in \mathbb{R}[x]_{\leq 4}$, so $f(x) = a + bx + cx^2 + dx^3 + ex^4$ for some constants $a, \dots, e$. We then have

$$\begin{aligned} f(1) &= a + b + c + d + e \\ f(-1) &= a - b + c - d + e \\ f'(1) - f'(-1) &= (b + 2c + 3d + 4e) - (b - 2c + 3d - 4e) = 4c + 8e \end{aligned}$$

It follows that $f \in V$ iff $a + b + c + d + e = a - b + c - d + e = 4c + 8e = 0$. This simplifies to $c = -2e$ and $a = e$ and $b = -d$, so

$$f(x) = e - dx - 2ex^2 + dx^3 + ex^4 = d(x^3 - x) + e(x^4 - 2x^2 + 1).$$

Thus, if we put $p(x) = x^3 - x$ and $q(x) = x^4 - 2x^2 + 1 = (x^2 - 1)^2$, then p, q is a basis for V. $\bigcirc$

Example ??: A *magic square* is a 3×3 matrix in which the sum of every row is the same, and the sum of every column is the same.

Magic squares

Example ??: A *magic square* is a 3×3 matrix in which the sum of every row is the same, and the sum of every column is the same. More explicitly, a matrix

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is a magic square iff we have

$$\begin{aligned} a + b + c &= d + e + f = g + h + i \\ a + d + g &= b + e + h = c + f + i. \end{aligned}$$

Example ??: A *magic square* is a 3×3 matrix in which the sum of every row is the same, and the sum of every column is the same. More explicitly, a matrix

$$X = \left[\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & i \end{smallmatrix}\right]$$

is a magic square iff we have

$$a + b + c = d + e + f = g + h + i$$
$$a + d + g = b + e + h = c + f + i.$$

Let V be the set of magic squares, which is easily seen to be a subspace of $M_3\mathbb{R}$; we will find a basis for V.

Example ??: A *magic square* is a 3×3 matrix in which the sum of every row is the same, and the sum of every column is the same. More explicitly, a matrix

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is a magic square iff we have

$$a + b + c = d + e + f = g + h + i$$
$$a + d + g = b + e + h = c + f + i.$$

Let V be the set of magic squares, which is easily seen to be a subspace of $M_3\mathbb{R}$; we will find a basis for V. First, we write

$$R(X) = a + b + c = d + e + f = g + h + i$$

Magic squares

Example ??: A *magic square* is a 3×3 matrix in which the sum of every row is the same, and the sum of every column is the same. More explicitly, a matrix

$$X = \left[\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & i \end{smallmatrix}\right]$$

is a magic square iff we have

$$\begin{aligned} a+b+c &= d+e+f = g+h+i \\ a+d+g &= b+e+h = c+f+i. \end{aligned}$$

Let V be the set of magic squares, which is easily seen to be a subspace of $M_3\mathbb{R}$; we will find a basis for V. First, we write

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \end{aligned}$$

Magic squares

Example ??: A *magic square* is a 3×3 matrix in which the sum of every row is the same, and the sum of every column is the same. More explicitly, a matrix

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is a magic square iff we have

$$a + b + c = d + e + f = g + h + i$$
$$a + d + g = b + e + h = c + f + i.$$

Let V be the set of magic squares, which is easily seen to be a subspace of $M_3\mathbb{R}$; we will find a basis for V. First, we write

$$R(X) = a + b + c = d + e + f = g + h + i$$
$$C(X) = a + d + g = b + e + h = c + f + i$$
$$T(X) = a + b + c + d + e + f + g + h + i.$$

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in V$$

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \\ T(X) &= a+b+c+d+e+f+g+h+i. \end{aligned}$$

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in V$$

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \\ T(X) &= a+b+c+d+e+f+g+h+i. \end{aligned}$$

On the one hand, we have
$T(X) = a+b+c+d+e+f+g+h+i =$
$(a+b+c)+(d+e+f)+(g+h+i) = 3R(X)$.

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in V$$

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \\ T(X) &= a+b+c+d+e+f+g+h+i. \end{aligned}$$

On the one hand, we have
$T(X) = a+b+c+d+e+f+g+h+i =$
$(a+b+c)+(d+e+f)+(g+h+i) = 3R(X)$. We also have
$T(X) = a+d+g+b+e+h+c+f+i =$
$(a+d+g)+(b+e+h)+(c+f+i) = 3C(X)$.

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in V$$

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \\ T(X) &= a+b+c+d+e+f+g+h+i. \end{aligned}$$

On the one hand, we have $T(X) = a+b+c+d+e+f+g+h+i = (a+b+c)+(d+e+f)+(g+h+i) = 3R(X)$. We also have $T(X) = a+d+g+b+e+h+c+f+i = (a+d+g)+(b+e+h)+(c+f+i) = 3C(X)$.
It follows that $R(X) = C(X) = T(X)/3$.

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in V$$

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \\ T(X) &= a+b+c+d+e+f+g+h+i. \end{aligned}$$

On the one hand, we have
$T(X) = a+b+c+d+e+f+g+h+i =$
$(a+b+c)+(d+e+f)+(g+h+i) = 3R(X)$. We also have
$T(X) = a+d+g+b+e+h+c+f+i =$
$(a+d+g)+(b+e+h)+(c+f+i) = 3C(X)$.
It follows that $R(X) = C(X) = T(X)/3$.

It is now convenient to consider the subspace $W = \{X \in V \mid T(X) = 0\}$

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in V$$

$$\begin{aligned} R(X) &= a+b+c = d+e+f = g+h+i \\ C(X) &= a+d+g = b+e+h = c+f+i \\ T(X) &= a+b+c+d+e+f+g+h+i. \end{aligned}$$

On the one hand, we have
$T(X) = a+b+c+d+e+f+g+h+i =$
$(a+b+c)+(d+e+f)+(g+h+i) = 3R(X)$. We also have
$T(X) = a+d+g+b+e+h+c+f+i =$
$(a+d+g)+(b+e+h)+(c+f+i) = 3C(X)$.
It follows that $R(X) = C(X) = T(X)/3$.

It is now convenient to consider the subspace $W = \{X \in V \mid T(X) = 0\}$, consisting of squares as above for which

$$a+b+c = d+e+f = g+h+i = 0$$
$$a+d+g = b+e+h = c+f+i = 0.$$

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W$$

$$\begin{array}{l} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{array}$$

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W \qquad \begin{array}{c} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{array}$$

For such a square, we certainly have

$$c = -a - b \qquad f = -d - e \qquad g = -a - d \qquad h = -b - e.$$

Magic squares

$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W$ $\qquad$ $\begin{array}{l} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{array}$

For such a square, we certainly have

$$c = -a - b \qquad f = -d - e \qquad g = -a - d \qquad h = -b - e.$$

Substituting this back into the equation $g + h + i = 0$ (or into the equation $c + f + i = 0$) gives $i = a + b + d + e$.

Magic squares

$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W$

$$\begin{aligned} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{aligned}$$

For such a square, we certainly have

$$c = -a - b \qquad f = -d - e \qquad g = -a - d \qquad h = -b - e.$$

Substituting this back into the equation $g + h + i = 0$ (or into the equation $c + f + i = 0$) gives $i = a + b + d + e$. It follows that any element of W can be written in the form

$$X = \begin{bmatrix} a & b & -a-b \\ d & e & -d-e \\ -a-d & -b-e & a+b+d+e \end{bmatrix}.$$

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W \qquad \begin{array}{l} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{array}$$

For such a square, we certainly have

$$c = -a-b \qquad f = -d-e \qquad g = -a-d \qquad h = -b-e.$$

Substituting this back into the equation $g+h+i=0$ (or into the equation $c+f+i=0$) gives $i = a+b+d+e$. It follows that any element of W can be written in the form

$$X = \begin{bmatrix} a & b & -a-b \\ d & e & -d-e \\ -a-d & -b-e & a+b+d+e \end{bmatrix}.$$

Equivalently, if we put

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \qquad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

Magic squares

$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W$ $\qquad$ $\begin{matrix} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{matrix}$

For such a square, we certainly have

$$c = -a - b \qquad f = -d - e \qquad g = -a - d \qquad h = -b - e.$$

Substituting this back into the equation $g + h + i = 0$ (or into the equation $c + f + i = 0$) gives $i = a + b + d + e$. It follows that any element of W can be written in the form

$$X = \begin{bmatrix} a & b & -a-b \\ d & e & -d-e \\ -a-d & -b-e & a+b+d+e \end{bmatrix}.$$

Equivalently, if we put

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \qquad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

then any element of W can be written in the form $X = aA + bB + dD + eE$ for some list a, b, d, e of real numbers.

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W \qquad \begin{array}{l} a+b+c=d+e+f=g+h+i=0 \\ a+d+g=b+e+h=c+f+i=0 \end{array}$$

For such a square, we certainly have

$$c = -a-b \qquad f = -d-e \qquad g = -a-d \qquad h = -b-e.$$

Substituting this back into the equation $g+h+i=0$ (or into the equation $c+f+i=0$) gives $i = a+b+d+e$. It follows that any element of W can be written in the form

$$X = \begin{bmatrix} a & b & -a-b \\ d & e & -d-e \\ -a-d & -b-e & a+b+d+e \end{bmatrix}.$$

Equivalently, if we put

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \qquad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

then any element of W can be written in the form $X = aA + bB + dD + eE$ for some list a, b, d, e of real numbers. This means that A, B, D, E spans W

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W \qquad \begin{matrix} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{matrix}$$

For such a square, we certainly have

$$c = -a-b \qquad f = -d-e \qquad g = -a-d \qquad h = -b-e.$$

Substituting this back into the equation $g+h+i = 0$ (or into the equation $c+f+i = 0$) gives $i = a+b+d+e$. It follows that any element of W can be written in the form

$$X = \begin{bmatrix} a & b & -a-b \\ d & e & -d-e \\ -a-d & -b-e & a+b+d+e \end{bmatrix}.$$

Equivalently, if we put

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \qquad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

then any element of W can be written in the form $X = aA + bB + dD + eE$ for some list a, b, d, e of real numbers. This means that A, B, D, E spans W, and these matrices are clearly linearly independent

Magic squares

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in W \qquad \begin{array}{l} a+b+c = d+e+f = g+h+i = 0 \\ a+d+g = b+e+h = c+f+i = 0 \end{array}$$

For such a square, we certainly have

$$c = -a-b \qquad f = -d-e \qquad g = -a-d \qquad h = -b-e.$$

Substituting this back into the equation $g+h+i=0$ (or into the equation $c+f+i=0$) gives $i = a+b+d+e$. It follows that any element of W can be written in the form

$$X = \begin{bmatrix} a & b & -a-b \\ d & e & -d-e \\ -a-d & -b-e & a+b+d+e \end{bmatrix}.$$

Equivalently, if we put

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \qquad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

then any element of W can be written in the form $X = aA + bB + dD + eE$ for some list a, b, d, e of real numbers. This means that A, B, D, E spans W, and these matrices are clearly linearly independent, so they form a basis for W. $\bigcirc$

Magic squares

Magic squares

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$). We claim that Q, A, B, D, E is a basis for V.

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.

Magic squares

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$.

Magic squares

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$.
This means that Q, A, B, D, E spans V.

Next, the matrix $Q = \left[\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right]$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$.
This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$.

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$.
This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$)

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$). We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$. We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$. As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$. This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$), and so $q = 0$.

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$). We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$. We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$. As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$. This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$), and so $q = 0$. This leaves $aA + bB + dD + eE = 0$

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$.
This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$), and so $q = 0$. This leaves $aA + bB + dD + eE = 0$, and A, B, D and E are linearly independent

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$.
This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$), and so $q = 0$. This leaves $aA + bB + dD + eE = 0$, and A, B, D and E are linearly independent, so $a = b = d = e = 0$ as well.

Next, the matrix $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$). We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$. We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$. As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$. This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$), and so $q = 0$. This leaves $aA + bB + dD + eE = 0$, and A, B, D and E are linearly independent, so $a = b = d = e = 0$ as well. This means that Q, A, B, D and E are linearly independent as well as spanning V, so they form a basis for V.

Next, the matrix $Q = \begin{bmatrix} 1&1&1 \\ 1&1&1 \\ 1&1&1 \end{bmatrix}$ lies in V but not in W (because $T(Q) = 9$).
We claim that Q, A, B, D, E is a basis for V.

Indeed, given $X \in V$ we can put $t = T(X)/9$ and $Y = X - tQ$.
We then have $Y \in V$ and $T(Y) = T(X) - tT(Q) = 0$, so $Y \in W$.
As A, B, D, E is a basis for W, we see that $Y = aA + bB + dD + eE$ for some $a, b, d, e \in \mathbb{R}$. It follows that $X = tQ + Y = tQ + aA + bB + dD + eE$.
This means that Q, A, B, D, E spans V.

Suppose we have a linear relation

$$qQ + aA + bB + dD + eE = 0$$

for some $q, a, b, d, e \in \mathbb{R}$. Applying T gives $9q = 0$ (because $T(A) = T(B) = T(D) = T(E) = 0$ and $T(Q) = 9$), and so $q = 0$. This leaves $aA + bB + dD + eE = 0$, and A, B, D and E are linearly independent, so $a = b = d = e = 0$ as well. This means that Q, A, B, D and E are linearly independent as well as spanning V, so they form a basis for V. Thus $\dim(V) = 5$. ○

Linear maps out of $\mathbb{R}^n$

Linear maps out of $\mathbb{R}^n$

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).

Linear maps out of $\mathbb{R}^n$

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

Definition ??: Let V be a vector space, and let v and w be elements of V. We then define $\mu_{v,w} \colon \mathbb{R}^2 \to V$ by

$$\mu_{v,w}\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = xv + yw.$$

Linear maps out of $\mathbb{R}^n$

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

Definition ??: Let V be a vector space, and let v and w be elements of V. We then define $\mu_{v,w} \colon \mathbb{R}^2 \to V$ by

$$\mu_{v,w}\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = xv + yw.$$

This makes sense because:

Linear maps out of $\mathbb{R}^n$

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

Definition ??: Let V be a vector space, and let v and w be elements of V. We then define $\mu_{v,w}\colon \mathbb{R}^2 \to V$ by

$$\mu_{v,w}\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = xv + yw.$$

This makes sense because:

- x is a number and $v \in V$ and V is a vector space, so $xv \in V$.

Linear maps out of $\mathbb{R}^n$

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

Definition ??: Let V be a vector space, and let v and w be elements of V. We then define $\mu_{v,w} \colon \mathbb{R}^2 \to V$ by

$$\mu_{v,w}\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = xv + yw.$$

This makes sense because:

- x is a number and $v \in V$ and V is a vector space, so $xv \in V$.
- y is a number and $w \in V$ and V is a vector space, so $yw \in V$.

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

Definition ??: Let V be a vector space, and let v and w be elements of V. We then define $\mu_{v,w} \colon \mathbb{R}^2 \to V$ by

$$\mu_{v,w}\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = xv + yw.$$

This makes sense because:

- x is a number and $v \in V$ and V is a vector space, so $xv \in V$.
- y is a number and $w \in V$ and V is a vector space, so $yw \in V$.
- xv and yw lie in the vector space V, so $xv + yw \in V$.

Linear maps out of $\mathbb{R}^n$

We next discuss linear maps $\mathbb{R}^n \to V$ (for any vector space V).
We will do the case $n = 2$ first; the general case is essentially the same, but with more complicated notation.

Definition ??: Let V be a vector space, and let v and w be elements of V. We then define $\mu_{v,w} \colon \mathbb{R}^2 \to V$ by

$$\mu_{v,w}\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right) = xv + yw.$$

This makes sense because:

- x is a number and $v \in V$ and V is a vector space, so $xv \in V$.
- y is a number and $w \in V$ and V is a vector space, so $yw \in V$.
- xv and yw lie in the vector space V, so $xv + yw \in V$.

It is clear that $\mu_{v,w}$ is a linear map. ○

Linear maps out of $\mathbb{R}^2$

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$.

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$. Similarly, the vector $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $w = \phi(\mathbf{e}_2) \in V$.

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$. Similarly, the vector $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $w = \phi(\mathbf{e}_2) \in V$. We claim that $\phi = \mu_{v,w}$.

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$. Similarly, the vector $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $w = \phi(\mathbf{e}_2) \in V$. We claim that $\phi = \mu_{v,w}$. Indeed, as ϕ is linear, we have

$$\phi(x\mathbf{e}_1 + y\mathbf{e}_2) = x\phi(\mathbf{e}_1) + y\phi(\mathbf{e}_2) = xv + yw = \mu_{v,w}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right).$$

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$. Similarly, the vector $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $w = \phi(\mathbf{e}_2) \in V$. We claim that $\phi = \mu_{v,w}$. Indeed, as ϕ is linear, we have

$$\phi(x\mathbf{e}_1 + y\mathbf{e}_2) = x\phi(\mathbf{e}_1) + y\phi(\mathbf{e}_2) = xv + yw = \mu_{v,w}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right).$$

On the other hand, it is clear that

$$x\mathbf{e}_1 + y\mathbf{e}_2 = x\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + y\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]$$

Linear maps out of $\mathbb{R}^2$

Proposition ??: Any linear map $\phi\colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$. Similarly, the vector $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $w = \phi(\mathbf{e}_2) \in V$. We claim that $\phi = \mu_{v,w}$. Indeed, as ϕ is linear, we have

$$\phi(x\mathbf{e}_1 + y\mathbf{e}_2) = x\phi(\mathbf{e}_1) + y\phi(\mathbf{e}_2) = xv + yw = \mu_{v,w}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right).$$

On the other hand, it is clear that

$$x\mathbf{e}_1 + y\mathbf{e}_2 = x\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + y\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}x\\y\end{smallmatrix}\right],$$

so the previous equation reads

$$\phi\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right) = \mu_{v,w}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right).$$

Proposition ??: Any linear map $\phi \colon \mathbb{R}^2 \to V$ has the form $\phi = \mu_{v,w}$ for some $v, w \in V$.

Proof: The vector $\mathbf{e}_1 = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $v = \phi(\mathbf{e}_1) \in V$. Similarly, the vector $\mathbf{e}_2 = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is an element of $\mathbb{R}^2$, and ϕ is a map from $\mathbb{R}^2$ to V, so we have an element $w = \phi(\mathbf{e}_2) \in V$. We claim that $\phi = \mu_{v,w}$. Indeed, as ϕ is linear, we have

$$\phi(x\mathbf{e}_1 + y\mathbf{e}_2) = x\phi(\mathbf{e}_1) + y\phi(\mathbf{e}_2) = xv + yw = \mu_{v,w}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right).$$

On the other hand, it is clear that

$$x\mathbf{e}_1 + y\mathbf{e}_2 = x\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + y\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}x\\y\end{smallmatrix}\right],$$

so the previous equation reads

$$\phi\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right) = \mu_{v,w}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\right).$$

This holds for all x and y, so $\phi = \mu_{v,w}$ as claimed. $\bigcirc$

Linear maps out of $\mathbb{R}^n$

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**).

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}}\colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Proposition ??: Any linear map $\phi\colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**). Thus, a linear map $\mathbb{R}^n \to V$ is essentially the same thing as a list of n elements of V.

For any list $\mathcal{V} = v_1, \dots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \dots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \dots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \dots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**). Thus, a linear map $\mathbb{R}^n \to V$ is essentially the same thing as a list of n elements of V.

Proof:

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**). Thus, a linear map $\mathbb{R}^n \to V$ is essentially the same thing as a list of n elements of V.

Proof: Put $v_i = \phi(\mathbf{e}_i) \in V$.

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**). Thus, a linear map $\mathbb{R}^n \to V$ is essentially the same thing as a list of n elements of V.

Proof: Put $v_i = \phi(\mathbf{e}_i) \in V$. For any $\mathbf{x} \in \mathbb{R}^n$ we have

$$\mathbf{x} = x_1 \mathbf{e}_1 + \ldots + x_n \mathbf{e}_n = \sum_i x_i \mathbf{e}_i$$

Linear maps out of $\mathbb{R}^n$

For any list $\mathcal{V} = v_1, \dots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \dots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \dots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \dots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**). Thus, a linear map $\mathbb{R}^n \to V$ is essentially the same thing as a list of n elements of V.

Proof: Put $v_i = \phi(\mathbf{e}_i) \in V$. For any $\mathbf{x} \in \mathbb{R}^n$ we have

$$\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n = \sum_i x_i\mathbf{e}_i,$$

so

$$\phi(\mathbf{x}) = \sum_i x_i\phi(\mathbf{e}_i) = \sum_i x_i v_i = \mu_{v_1,\dots,v_n}(\mathbf{x}),$$

For any list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V, we can define a linear map $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ by
$\mu_{\mathcal{V}}([x_1, \ldots, x_n]^T) = \sum_i x_i v_i = x_1 v_1 + \ldots + x_n v_n$.

Proposition ??: Any linear map $\phi \colon \mathbb{R}^n \to V$ has the form $\phi = \mu_{\mathcal{V}}$ for some list $\mathcal{V} = v_1, \ldots, v_n$ of elements of V (which are uniquely determined by the formula $v_i = \phi(\mathbf{e}_i)$, where $\mathbf{e}_i$ is as in Definition **??**). Thus, a linear map $\mathbb{R}^n \to V$ is essentially the same thing as a list of n elements of V.

Proof: Put $v_i = \phi(\mathbf{e}_i) \in V$. For any $\mathbf{x} \in \mathbb{R}^n$ we have

$$\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = \sum_i x_i\mathbf{e}_i,$$

so

$$\phi(\mathbf{x}) = \sum_i x_i \phi(\mathbf{e}_i) = \sum_i x_i v_i = \mu_{v_1,\ldots,v_n}(\mathbf{x}),$$

so $\phi = \mu_{v_1,\ldots,v_n}$. $\bigcirc$

An example

An example

Consider the map $\phi\colon \mathbb{R}^3 \to M_3\mathbb{R}$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a & a+b & a \\ a+b & a+b+c & a+b \\ a & a+b & a \end{bmatrix}$$

An example

Consider the map $\phi\colon \mathbb{R}^3 \to M_3\mathbb{R}$ given by

$$\phi\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a & a+b & a \\ a+b & a+b+c & a+b \\ a & a+b & a \end{bmatrix}$$

Put $\mathcal{A} = A_1, A_2, A_3$, where

$$A_1 = \phi(\mathbf{e}_1) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad A_2 = \phi(\mathbf{e}_2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad A_3 = \phi(\mathbf{e}_3) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

An example

Consider the map $\phi\colon \mathbb{R}^3 \to M_3\mathbb{R}$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a & a+b & a \\ a+b & a+b+c & a+b \\ a & a+b & a \end{bmatrix}$$

Put $\mathcal{A} = A_1, A_2, A_3$, where

$$A_1 = \phi(\mathbf{e}_1) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad A_2 = \phi(\mathbf{e}_2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad A_3 = \phi(\mathbf{e}_3) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then

$$\mu_{\mathcal{A}} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & a+b & a \\ a+b & a+b+c & a+b \\ a & a+b & a \end{bmatrix} = \phi \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

An example

Consider the map $\phi\colon \mathbb{R}^3 \to M_3\mathbb{R}$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a & a+b & a \\ a+b & a+b+c & a+b \\ a & a+b & a \end{bmatrix}$$

Put $\mathcal{A} = A_1, A_2, A_3$, where

$$A_1 = \phi(\mathbf{e}_1) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad A_2 = \phi(\mathbf{e}_2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad A_3 = \phi(\mathbf{e}_3) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then

$$\mu_{\mathcal{A}} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & a+b & a \\ a+b & a+b+c & a+b \\ a & a+b & a \end{bmatrix} = \phi \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

so $\phi = \mu_{\mathcal{A}}$. $\bigcirc$

Another example

Another example

Consider the map $\phi\colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Another example

Consider the map $\phi\colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Put $\mathcal{P} = p_1, p_2, p_3$, where

Another example

Consider the map $\phi\colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Put $\mathcal{P} = p_1, p_2, p_3$, where

$$p_1(x) = \phi(\mathbf{e}_1) = x^2 + (x+1)^2 + (x+2)^2 = 3x^2 + 6x + 5$$

Another example

Consider the map $\phi\colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Put $\mathcal{P} = p_1, p_2, p_3$, where

$$p_1(x) = \phi(\mathbf{e}_1) = x^2 + (x+1)^2 + (x+2)^2 = 3x^2 + 6x + 5$$
$$p_2(x) = \phi(\mathbf{e}_2) = x^2 + (x+1)^2 = 2x^2 + 2x + 1$$

Another example

Consider the map $\phi\colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Put $\mathcal{P} = p_1, p_2, p_3$, where

$$\begin{aligned}
p_1(x) &= \phi(\mathbf{e}_1) = x^2 + (x+1)^2 + (x+2)^2 = 3x^2 + 6x + 5 \\
p_2(x) &= \phi(\mathbf{e}_2) = x^2 + (x+1)^2 = 2x^2 + 2x + 1 \\
p_3(x) &= \phi(\mathbf{e}_3) = x^2.
\end{aligned}$$

Another example

Consider the map $\phi \colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Put $\mathcal{P} = p_1, p_2, p_3$, where

$$\begin{aligned} p_1(x) &= \phi(\mathbf{e}_1) = x^2 + (x+1)^2 + (x+2)^2 = 3x^2 + 6x + 5 \\ p_2(x) &= \phi(\mathbf{e}_2) = x^2 + (x+1)^2 = 2x^2 + 2x + 1 \\ p_3(x) &= \phi(\mathbf{e}_3) = x^2. \end{aligned}$$

Then

$$\mu_{\mathcal{P}} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a(3x^2 + 6x + 5) + b(2x^2 + 2x + 1) + cx^2$$

Another example

Consider the map $\phi\colon \mathbb{R}^3 \to \mathbb{R}[x]$ given by

$$\phi \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a+b+c)x^2 + (a+b)(x+1)^2 + a(x+2)^2.$$

Put $\mathcal{P} = p_1, p_2, p_3$, where

$$\begin{aligned} p_1(x) &= \phi(\mathbf{e}_1) = x^2 + (x+1)^2 + (x+2)^2 = 3x^2 + 6x + 5 \\ p_2(x) &= \phi(\mathbf{e}_2) = x^2 + (x+1)^2 = 2x^2 + 2x + 1 \\ p_3(x) &= \phi(\mathbf{e}_3) = x^2. \end{aligned}$$

Then

$$\mu_{\mathcal{P}} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a(3x^2+6x+5) + b(2x^2+2x+1) + cx^2 = \phi \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined).

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined). Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined). Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Proof: A linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as a list $\mathbf{v}_1, \dots, \mathbf{v}_n$ of elements of $\mathbb{R}^m$.

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined). Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Proof: A linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as a list $\mathbf{v}_1, \dots, \mathbf{v}_n$ of elements of $\mathbb{R}^m$. If we write each $\mathbf{v}_i$ as a column vector, then the list can be visualised in an obvious way as an $m \times n$ matrix.

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined). Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Proof: A linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as a list $\mathbf{v}_1, \dots, \mathbf{v}_n$ of elements of $\mathbb{R}^m$. If we write each $\mathbf{v}_i$ as a column vector, then the list can be visualised in an obvious way as an $m \times n$ matrix. For example, the list

$$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right], \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right], \left[\begin{smallmatrix}5\\6\end{smallmatrix}\right], \left[\begin{smallmatrix}7\\8\end{smallmatrix}\right]$$

corresponds to the matrix

$$\left[\begin{smallmatrix}1&3&5&7\\2&4&6&8\end{smallmatrix}\right].$$

Linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined). Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Proof: A linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as a list $\mathbf{v}_1, \dots, \mathbf{v}_n$ of elements of $\mathbb{R}^m$. If we write each $\mathbf{v}_i$ as a column vector, then the list can be visualised in an obvious way as an $m \times n$ matrix. For example, the list

$$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right], \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right], \left[\begin{smallmatrix}5\\6\end{smallmatrix}\right], \left[\begin{smallmatrix}7\\8\end{smallmatrix}\right]$$

corresponds to the matrix

$$\left[\begin{smallmatrix}1&3&5&7\\2&4&6&8\end{smallmatrix}\right].$$

Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Corollary ??: Every linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ has the form ϕ_A (as in Example **??**) for some $m \times n$ matrix A (which is uniquely determined). Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix.

Proof: A linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as a list $\mathbf{v}_1, \dots, \mathbf{v}_n$ of elements of $\mathbb{R}^m$. If we write each $\mathbf{v}_i$ as a column vector, then the list can be visualised in an obvious way as an $m \times n$ matrix. For example, the list

$$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right], \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right], \left[\begin{smallmatrix}5\\6\end{smallmatrix}\right], \left[\begin{smallmatrix}7\\8\end{smallmatrix}\right]$$

corresponds to the matrix

$$\left[\begin{smallmatrix}1&3&5&7\\2&4&6&8\end{smallmatrix}\right].$$

Thus, a linear map $\alpha\colon \mathbb{R}^n \to \mathbb{R}^m$ is essentially the same thing as an $m \times n$ matrix. There are some things to check to see that this is compatible with Example **??**, but we shall not go through the details. $\bigcirc$

A rotation matrix

A rotation matrix

Consider the linear map $\rho\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\rho \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}$$

A rotation matrix

Consider the linear map $\rho\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\rho \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}$$

(so $\rho(\mathbf{v})$ is obtained by rotating $\mathbf{v}$ through $2\pi/3$ around the line $x = y = z$).

A rotation matrix

Consider the linear map $\rho\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\rho \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}$$

(so $\rho(\mathbf{v})$ is obtained by rotating $\mathbf{v}$ through $2\pi/3$ around the line $x = y = z$).

Then

$$\rho(\mathbf{e}_1) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \rho(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \rho(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

A rotation matrix

Consider the linear map $\rho\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\rho \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}$$

(so $\rho(\mathbf{v})$ is obtained by rotating $\mathbf{v}$ through $2\pi/3$ around the line $x = y = z$).

Then

$$\rho(\mathbf{e}_1) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \rho(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \rho(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

This means that $\rho = \phi_R$, where

$$R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Matrices for vector products

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$

Matrices for vector products

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$.

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. This is linear, so it must have the form $\beta = \phi_B$ for some 3×3 matrix B.

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. This is linear, so it must have the form $\beta = \phi_B$ for some 3×3 matrix B. To find B, we note that

$$\beta \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} bz - cy \\ cx - az \\ ay - bx \end{bmatrix},$$

Matrices for vector products

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. This is linear, so it must have the form $\beta = \phi_B$ for some 3×3 matrix B. To find B, we note that

$$\beta \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} bz - cy \\ cx - az \\ ay - bx \end{bmatrix},$$

so

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ c \\ -b \end{bmatrix} \qquad \beta(\mathbf{e}_2) = \begin{bmatrix} -c \\ 0 \\ a \end{bmatrix} \qquad \beta(\mathbf{e}_3) = \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}.$$

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. This is linear, so it must have the form $\beta = \phi_B$ for some 3×3 matrix B. To find B, we note that

$$\beta \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} bz - cy \\ cx - az \\ ay - bx \end{bmatrix},$$

so

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ c \\ -b \end{bmatrix} \qquad \beta(\mathbf{e}_2) = \begin{bmatrix} -c \\ 0 \\ a \end{bmatrix} \qquad \beta(\mathbf{e}_3) = \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}.$$

These three vectors are the columns of B, so

$$B = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}.$$

Example ??: Consider a vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{v}) = \mathbf{a} \times \mathbf{v}$. This is linear, so it must have the form $\beta = \phi_B$ for some 3×3 matrix B. To find B, we note that

$$\beta \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} bz - cy \\ cx - az \\ ay - bx \end{bmatrix},$$

so

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ c \\ -b \end{bmatrix} \qquad \beta(\mathbf{e}_2) = \begin{bmatrix} -c \\ 0 \\ a \end{bmatrix} \qquad \beta(\mathbf{e}_3) = \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}.$$

These three vectors are the columns of B, so

$$B = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}.$$

(Note incidentally that the matrices arising in this way are precisely the 3×3 antisymmetric matrices.) ◯

Example ??:

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$)

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$.

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P.

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P. The formula for this is $\pi(\mathbf{v}) = \mathbf{v} - \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.

Matrices for plane projections

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P. The formula for this is $\pi(\mathbf{v}) = \mathbf{v} - \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.
The map π is linear, so it must have the form $\pi(\mathbf{v}) = A\mathbf{v}$ for some 3×3 matrix A.

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P. The formula for this is $\pi(\mathbf{v}) = \mathbf{v} - \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.
The map π is linear, so it must have the form $\pi(\mathbf{v}) = A\mathbf{v}$ for some 3×3 matrix A. To find A, we observe that

$$\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} - (ax + by + cz) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x - a^2x - aby - acz \\ y - abx - b^2y - bcz \\ z - acx - bcy - c^2z \end{bmatrix}.$$

Matrices for plane projections

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P. The formula for this is $\pi(\mathbf{v}) = \mathbf{v} - \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.
The map π is linear, so it must have the form $\pi(\mathbf{v}) = A\mathbf{v}$ for some 3×3 matrix A. To find A, we observe that

$$\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} - (ax + by + cz) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x - a^2x - aby - acz \\ y - abx - b^2y - bcz \\ z - acx - bcy - c^2z \end{bmatrix}.$$

It follows that

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1 - a^2 \\ -ab \\ -ac \end{bmatrix} \qquad \pi(\mathbf{e}_2) = \begin{bmatrix} -ab \\ 1 - b^2 \\ -bc \end{bmatrix} \qquad \pi(\mathbf{e}_3) = \begin{bmatrix} -ac \\ -bc \\ 1 - c^2 \end{bmatrix}.$$

Matrices for plane projections

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P. The formula for this is $\pi(\mathbf{v}) = \mathbf{v} - \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.
The map π is linear, so it must have the form $\pi(\mathbf{v}) = A\mathbf{v}$ for some 3×3 matrix A. To find A, we observe that

$$\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} - (ax + by + cz) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x - a^2x - aby - acz \\ y - abx - b^2y - bcz \\ z - acx - bcy - c^2z \end{bmatrix}.$$

It follows that

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1 - a^2 \\ -ab \\ -ac \end{bmatrix} \qquad \pi(\mathbf{e}_2) = \begin{bmatrix} -ab \\ 1 - b^2 \\ -bc \end{bmatrix} \qquad \pi(\mathbf{e}_3) = \begin{bmatrix} -ac \\ -bc \\ 1 - c^2 \end{bmatrix}.$$

These three vectors are the columns of A, so $A = \begin{bmatrix} 1 - a^2 & -ab & -ac \\ -ab & 1 - b^2 & -bc \\ -ac & -bc & 1 - c^2 \end{bmatrix}$.

Matrices for plane projections

Example ??: Consider a unit vector $\mathbf{a} = [a, b, c]^T \in \mathbb{R}^3$ (so $a^2 + b^2 + c^2 = 1$) and let P be the plane perpendicular to $\mathbf{a}$. For any $\mathbf{v} \in \mathbb{R}^3$, we let $\pi(\mathbf{v})$ be the projection of $\mathbf{v}$ onto P. The formula for this is $\pi(\mathbf{v}) = \mathbf{v} - \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.
The map π is linear, so it must have the form $\pi(\mathbf{v}) = A\mathbf{v}$ for some 3×3 matrix A. To find A, we observe that

$$\pi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} - (ax + by + cz) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x - a^2x - aby - acz \\ y - abx - b^2y - bcz \\ z - acx - bcy - c^2z \end{bmatrix}.$$

It follows that

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1 - a^2 \\ -ab \\ -ac \end{bmatrix} \qquad \pi(\mathbf{e}_2) = \begin{bmatrix} -ab \\ 1 - b^2 \\ -bc \end{bmatrix} \qquad \pi(\mathbf{e}_3) = \begin{bmatrix} -ac \\ -bc \\ 1 - c^2 \end{bmatrix}.$$

These three vectors are the columns of A, so $A = \begin{bmatrix} 1-a^2 & -ab & -ac \\ -ab & 1-b^2 & -bc \\ -ac & -bc & 1-c^2 \end{bmatrix}$.

It is an exercise to check that $A^2 = A^T = A$ and $\det(A) = 0$. ○

Matrices for linear maps

Matrices for linear maps

Let V and W be finite-dimensional vector spaces

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ say.

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha\colon V \to W$ be a linear map.

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ say. Let $\alpha\colon V \to W$ be a linear map. Then $\alpha(v_j)$ is an element of W, so it can be expressed (uniquely) in terms of the basis $\mathcal{W}$, say

$$\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m.$$

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha \colon V \to W$ be a linear map. Then $\alpha(v_j)$ is an element of W, so it can be expressed (uniquely) in terms of the basis $\mathcal{W}$, say

$$\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m.$$

These numbers a_{ij} form an $n \times m$ matrix A, which we call *the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$*.

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha \colon V \to W$ be a linear map. Then $\alpha(v_j)$ is an element of W, so it can be expressed (uniquely) in terms of the basis $\mathcal{W}$, say

$$\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m.$$

These numbers a_{ij} form an $n \times m$ matrix A, which we call *the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$*.

Remark ??: Often we consider the case where $W = V$ and so we have a map $\alpha \colon V \to V$, and $\mathcal{V}$ and $\mathcal{W}$ are bases for the same space.

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha\colon V \to W$ be a linear map. Then $\alpha(v_j)$ is an element of W, so it can be expressed (uniquely) in terms of the basis $\mathcal{W}$, say

$$\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m.$$

These numbers a_{ij} form an $n \times m$ matrix A, which we call *the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$*.

Remark ??: Often we consider the case where $W = V$ and so we have a map $\alpha\colon V \to V$, and $\mathcal{V}$ and $\mathcal{W}$ are bases for the same space. It is often natural to take $\mathcal{W} = \mathcal{V}$, but everything still makes sense even if $\mathcal{W} \neq \mathcal{V}$. ○

Matrices for linear maps

Matrices for linear maps

Let V and W be finite-dimensional vector spaces

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say.

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha \colon V \to W$ be a linear map.

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha \colon V \to W$ be a linear map. Then $\alpha(v_j)$ is an element of W, so it can be expressed (uniquely) in terms of the basis $\mathcal{W}$, say

$$\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m.$$

Matrices for linear maps

Let V and W be finite-dimensional vector spaces, with bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ say. Let $\alpha \colon V \to W$ be a linear map. Then $\alpha(v_j)$ is an element of W, so it can be expressed (uniquely) in terms of the basis $\mathcal{W}$, say

$$\alpha(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m.$$

These numbers a_{ij} form an $n \times m$ matrix A, which we call *the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$*.

Adapted bases for vector products

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$$

Adapted bases for vector products

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We then have

$$\begin{aligned}
\beta(\mathbf{a}) &= 0 &&= 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} \\
\beta(\mathbf{b}) &= \mathbf{c} &&= 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c} \\
\beta(\mathbf{c}) &= -\mathbf{b} &&= 0\mathbf{a} + (-1)\mathbf{b} + 0\mathbf{c}.
\end{aligned}$$

Adapted bases for vector products

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We then have

$$\begin{aligned} \beta(\mathbf{a}) &= 0 && = 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} \\ \beta(\mathbf{b}) &= \mathbf{c} && = 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c} \\ \beta(\mathbf{c}) &= -\mathbf{b} && = 0\mathbf{a} + (-1)\mathbf{b} + 0\mathbf{c}. \end{aligned}$$

The columns in the matrix we want are the lists of coefficients in the three equations above: the first equation gives the first column, the second equation gives the second column, and the third equation gives the third column.

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We then have

$$\begin{aligned}
\beta(\mathbf{a}) &= 0 &&= 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} \\
\beta(\mathbf{b}) &= \mathbf{c} &&= 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c} \\
\beta(\mathbf{c}) &= -\mathbf{b} &&= 0\mathbf{a} + (-1)\mathbf{b} + 0\mathbf{c}.
\end{aligned}$$

The columns in the matrix we want are the lists of coefficients in the three equations above: the first equation gives the first column, the second equation gives the second column, and the third equation gives the third column. Thus, the the matrix of β with respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Adapted bases for projectors

Adapted bases for projectors

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{a}, \mathbf{x} \rangle \mathbf{a}.$$

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{a}, \mathbf{x} \rangle \mathbf{a}.$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{a}, \mathbf{x} \rangle \mathbf{a}.$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We then have

$$\begin{aligned} \pi(\mathbf{a}) &= 0 &&= 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} \\ \pi(\mathbf{b}) &= \mathbf{b} &&= 0\mathbf{a} + 1\mathbf{b} + 0\mathbf{c} \\ \pi(\mathbf{c}) &= \mathbf{c} &&= 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c}. \end{aligned}$$

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{a}, \mathbf{x} \rangle \mathbf{a}.$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We then have

$$\begin{aligned} \pi(\mathbf{a}) &= 0 &&= 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} \\ \pi(\mathbf{b}) &= \mathbf{b} &&= 0\mathbf{a} + 1\mathbf{b} + 0\mathbf{c} \\ \pi(\mathbf{c}) &= \mathbf{c} &&= 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c}. \end{aligned}$$

The columns in the matrix we want are the lists of coefficients in the three equations above: the first equation gives the first column, the second equation gives the second column, and the third equation gives the third column.

Example ??: Let $\mathbf{a}$ be a unit vector in $\mathbb{R}^3$, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{a}, \mathbf{x} \rangle \mathbf{a}.$$

Choose any unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$, and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, so $\mathbf{c}$ is another unit vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We then have

$$\begin{aligned}
\pi(\mathbf{a}) &= 0 &&= 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} \\
\pi(\mathbf{b}) &= \mathbf{b} &&= 0\mathbf{a} + 1\mathbf{b} + 0\mathbf{c} \\
\pi(\mathbf{c}) &= \mathbf{c} &&= 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c}.
\end{aligned}$$

The columns in the matrix we want are the lists of coefficients in the three equations above: the first equation gives the first column, the second equation gives the second column, and the third equation gives the third column. Thus, the the matrix of π with respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Shifting polynomials

Shifting polynomials

Example ??: Define $\phi\colon \mathbb{R}[x]_{<4} \to \mathbb{R}[x]_{<4}$ by $\phi(x^k) = (x+1)^k$.

Shifting polynomials

Example ??: Define $\phi\colon \mathbb{R}[x]_{<4} \to \mathbb{R}[x]_{<4}$ by $\phi(x^k) = (x+1)^k$. Let A be the matrix of ϕ with respect to the basis $1, x, x^2, x^3$.

Shifting polynomials

Example ??: Define $\phi\colon \mathbb{R}[x]_{<4} \to \mathbb{R}[x]_{<4}$ by $\phi(x^k) = (x+1)^k$. Let A be the matrix of ϕ with respect to the basis $1, x, x^2, x^3$. We then have

$$\begin{aligned} \phi(1) &= 1 \\ \phi(x) &= 1 + x \\ \phi(x^2) &= 1 + 2x + x^2 \\ \phi(x^3) &= 1 + 3x + 3x^2 + x^3, \end{aligned}$$

Example ??: Define $\phi\colon \mathbb{R}[x]_{<4} \to \mathbb{R}[x]_{<4}$ by $\phi(x^k) = (x+1)^k$. Let A be the matrix of ϕ with respect to the basis $1, x, x^2, x^3$. We then have

$$\begin{aligned}
\phi(1) &= 1 \\
\phi(x) &= 1 + x \\
\phi(x^2) &= 1 + 2x + x^2 \\
\phi(x^3) &= 1 + 3x + 3x^2 + x^3,
\end{aligned}$$

or in other words

$$\begin{aligned}
\phi(x^0) &= 1.x^0 + 0.x^1 + 0.x^2 + 0.x^3 \\
\phi(x^1) &= 1.x^0 + 1.x^1 + 0.x^2 + 0.x^3 \\
\phi(x^2) &= 1.x^0 + 2.x^1 + 1.x^2 + 0.x^3 \\
\phi(x^3) &= 1.x^0 + 3.x^1 + 3.x^2 + 1.x^3.
\end{aligned}$$

Shifting polynomials

Example ??: Define $\phi\colon \mathbb{R}[x]_{<4} \to \mathbb{R}[x]_{<4}$ by $\phi(x^k) = (x+1)^k$. Let A be the matrix of ϕ with respect to the basis $1, x, x^2, x^3$. We then have

$$\begin{aligned}
\phi(1) &= 1 \\
\phi(x) &= 1 + x \\
\phi(x^2) &= 1 + 2x + x^2 \\
\phi(x^3) &= 1 + 3x + 3x^2 + x^3,
\end{aligned}$$

or in other words

$$\begin{aligned}
\phi(x^0) &= 1.x^0 + 0.x^1 + 0.x^2 + 0.x^3 \\
\phi(x^1) &= 1.x^0 + 1.x^1 + 0.x^2 + 0.x^3 \\
\phi(x^2) &= 1.x^0 + 2.x^1 + 1.x^2 + 0.x^3 \\
\phi(x^3) &= 1.x^0 + 3.x^1 + 3.x^2 + 1.x^3.
\end{aligned}
\qquad
A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The Vandermonde matrix

Example ??:

The Vandermonde matrix

Example ??: Define $\phi\colon \mathbb{R}[x]_{<5} \to \mathbb{R}^4$ by

$$\phi(f) = [f(1), f(2), f(3), f(4)]^T.$$

The Vandermonde matrix

Example ??: Define $\phi\colon \mathbb{R}[x]_{<5} \to \mathbb{R}^4$ by

$$\phi(f) = [f(1), f(2), f(3), f(4)]^T.$$

Then

$$\phi(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \phi(x) = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \phi(x^2) = \begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \end{bmatrix} \quad \phi(x^3) = \begin{bmatrix} 1 \\ 8 \\ 27 \\ 64 \end{bmatrix} \quad \phi(x^4) = \begin{bmatrix} 1 \\ 16 \\ 81 \\ 256 \end{bmatrix}$$

The Vandermonde matrix

Example ??: Define $\phi\colon \mathbb{R}[x]_{<5} \to \mathbb{R}^4$ by

$$\phi(f) = [f(1), f(2), f(3), f(4)]^T.$$

Then

$$\phi(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \phi(x) = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \phi(x^2) = \begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \end{bmatrix} \quad \phi(x^3) = \begin{bmatrix} 1 \\ 8 \\ 27 \\ 64 \end{bmatrix} \quad \phi(x^4) = \begin{bmatrix} 1 \\ 16 \\ 81 \\ 256 \end{bmatrix}$$

so the matrix of ϕ with respect to the usual bases is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 3 & 9 & 27 & 81 \\ 1 & 4 & 16 & 64 & 256 \end{bmatrix}.$$

The reverse maps

The reverse maps

Example ??: Define $\phi\colon \mathbb{R}^4 \to \mathbb{R}^4$ by

$$\phi \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_4 \\ x_3 \\ x_2 \\ x_1 \end{bmatrix}.$$

Example ??: Define $\phi\colon \mathbb{R}^4 \to \mathbb{R}^4$ by

$$\phi \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_4 \\ x_3 \\ x_2 \\ x_1 \end{bmatrix}.$$

The associated matrix (with respect to the standard basis) is

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Shifting waves

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x + \pi/4)$.

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x + \pi/4)$. As

$$\sin(x + \pi/4) = \sin(x)\cos(\pi/4) + \cos(x)\sin(\pi/4) = \tfrac{1}{\sqrt{2}}\sin(x) + \tfrac{1}{\sqrt{2}}\cos(x)$$

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x + \pi/4)$. As

$$\sin(x + \pi/4) = \sin(x)\cos(\pi/4) + \cos(x)\sin(\pi/4) = \frac{1}{\sqrt{2}}\sin(x) + \frac{1}{\sqrt{2}}\cos(x),$$

we have $\phi(\sin) = \frac{1}{\sqrt{2}}\sin + \frac{1}{\sqrt{2}}\cos$.

Shifting waves

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x + \pi/4)$. As

$$\sin(x + \pi/4) = \sin(x)\cos(\pi/4) + \cos(x)\sin(\pi/4) = \frac{1}{\sqrt{2}}\sin(x) + \frac{1}{\sqrt{2}}\cos(x),$$

we have $\phi(\sin) = \frac{1}{\sqrt{2}}\sin + \frac{1}{\sqrt{2}}\cos$. As

$$\cos(x + \pi/4) = \cos(x)\cos(\pi/4) - \sin(x)\sin(\pi/4) = \frac{1}{\sqrt{2}}\cos(x) + (-\frac{1}{\sqrt{2}})\sin(x)$$

Shifting waves

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x + \pi/4)$. As

$$\sin(x + \pi/4) = \sin(x)\cos(\pi/4) + \cos(x)\sin(\pi/4) = \tfrac{1}{\sqrt{2}}\sin(x) + \tfrac{1}{\sqrt{2}}\cos(x),$$

we have $\phi(\sin) = \frac{1}{\sqrt{2}}\sin + \frac{1}{\sqrt{2}}\cos$. As

$$\cos(x + \pi/4) = \cos(x)\cos(\pi/4) - \sin(x)\sin(\pi/4) = \tfrac{1}{\sqrt{2}}\cos(x) + (-\tfrac{1}{\sqrt{2}})\sin(x),$$

we have $\phi(\cos) = -\frac{1}{\sqrt{2}}\sin + \frac{1}{\sqrt{2}}\cos$.

Shifting waves

Example ??: Let V be the space of solutions of the differential equation $f'' + f = 0$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x + \pi/4)$. As

$$\sin(x+\pi/4) = \sin(x)\cos(\pi/4) + \cos(x)\sin(\pi/4) = \frac{1}{\sqrt{2}}\sin(x) + \frac{1}{\sqrt{2}}\cos(x),$$

we have $\phi(\sin) = \frac{1}{\sqrt{2}}\sin + \frac{1}{\sqrt{2}}\cos$. As

$$\cos(x+\pi/4) = \cos(x)\cos(\pi/4) - \sin(x)\sin(\pi/4) = \frac{1}{\sqrt{2}}\cos(x) + (-\frac{1}{\sqrt{2}})\sin(x),$$

we have $\phi(\cos) = -\frac{1}{\sqrt{2}}\sin + \frac{1}{\sqrt{2}}\cos$. It follows that the matrix of ϕ with respect to the basis $\{\sin, \cos\}$ is

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Matrix examples

Matrix examples

Example ??: Define $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^T$.

Example ??: Define $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^T$. In terms of the usual basis

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$$

Matrix examples

Example ??: Define $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^T$. In terms of the usual basis

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$$

we have

$$\begin{aligned}
\phi(E_1) &= E_1 &&= 1.E_1 + 0.E_2 + 0.E_3 + 0.E_4 \\
\phi(E_2) &= E_3 &&= 0.E_1 + 0.E_2 + 1.E_3 + 0.E_4 \\
\phi(E_3) &= E_2 &&= 0.E_1 + 1.E_2 + 0.E_3 + 0.E_4 \\
\phi(E_4) &= E_4 &&= 0.E_1 + 0.E_2 + 0.E_3 + 1.E_4
\end{aligned}$$

Matrix examples

Example ??: Define $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\phi(A) = A^T$. In terms of the usual basis

$$E_1 = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$$

we have

$$\begin{aligned}
\phi(E_1) &= E_1 &&= 1.E_1 + 0.E_2 + 0.E_3 + 0.E_4 \\
\phi(E_2) &= E_3 &&= 0.E_1 + 0.E_2 + 1.E_3 + 0.E_4 \\
\phi(E_3) &= E_2 &&= 0.E_1 + 1.E_2 + 0.E_3 + 0.E_4 \\
\phi(E_4) &= E_4 &&= 0.E_1 + 0.E_2 + 0.E_3 + 1.E_4
\end{aligned}$$

The matrix of ϕ is thus

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Matrix examples

Matrix examples

Example ??: Define $\psi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\psi(A) = A - \text{trace}(A)I/2$.

Matrix examples

Example ??: Define $\psi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\psi(A) = A - \text{trace}(A)I/2$. In terms of the usual basis

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$$

Matrix examples

Example ??: Define $\psi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\psi(A) = A - \mathrm{trace}(A)I/2$. In terms of the usual basis

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$$

we have

$$\begin{aligned}
\psi(E_1) &= E_1 - I/2 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} && = \tfrac{1}{2}E_1 + 0.E_2 + 0.E_3 + (-\tfrac{1}{2})E_4 \\
\psi(E_2) &= E_2 && = 0.E_1 + 1.E_2 + 0.E_3 + 0.E_4 \\
\psi(E_3) &= E_3 && = 0.E_1 + 0.E_2 + 1.E_3 + 0.E_4 \\
\psi(E_4) &= E_4 - I/2 = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} && = (-\tfrac{1}{2}).E_1 + 0.E_2 + 0.E_3 + \tfrac{1}{2}.E_4
\end{aligned}$$

Matrix examples

Example ??: Define $\psi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ by $\psi(A) = A - \mathrm{trace}(A)I/2$. In terms of the usual basis

$$E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \qquad E_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] \qquad E_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$$

we have

$$\begin{aligned}
\psi(E_1) &= E_1 - I/2 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} & &= \tfrac{1}{2}E_1 + 0.E_2 + 0.E_3 + (-\tfrac{1}{2})E_4 \\
\psi(E_2) &= E_2 & &= 0.E_1 + 1.E_2 + 0.E_3 + 0.E_4 \\
\psi(E_3) &= E_3 & &= 0.E_1 + 0.E_2 + 1.E_3 + 0.E_4 \\
\psi(E_4) &= E_4 - I/2 = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} & &= (-\tfrac{1}{2}).E_1 + 0.E_2 + 0.E_3 + \tfrac{1}{2}.E_4
\end{aligned}$$

The matrix is thus

$$\begin{bmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}. \bigcirc$$

Reminders

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.
- Every linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ is ϕ_A for some A.

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.
- Every linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ is ϕ_A for some A.
- Given a vector space V and a list $\mathcal{V} = v_1, \ldots, v_m$ of elements of V we define $\mu_{\mathcal{V}} \colon \mathbb{R}^m \to V$ by $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = \sum_i \lambda_i v_i$.

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.
- Every linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ is ϕ_A for some A.
- Given a vector space V and a list $\mathcal{V} = v_1, \dots, v_m$ of elements of V we define $\mu_{\mathcal{V}} \colon \mathbb{R}^m \to V$ by $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = \sum_i \lambda_i v_i$.
- If $\mathcal{V}$ is a basis then $\mu_{\mathcal{V}}$ is an isomorphism.

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.
- Every linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ is ϕ_A for some A.
- Given a vector space V and a list $\mathcal{V} = v_1, \dots, v_m$ of elements of V we define $\mu_{\mathcal{V}} \colon \mathbb{R}^m \to V$ by $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = \sum_i \lambda_i v_i$.
- If $\mathcal{V}$ is a basis then $\mu_{\mathcal{V}}$ is an isomorphism.
- Suppose we have a linear map $\alpha \colon V \to W$, a basis $\mathcal{V} = v_1, \dots, v_m$ for V and a basis $\mathcal{W} = w_1, \dots, w_n$ for W.

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.
- Every linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ is ϕ_A for some A.
- Given a vector space V and a list $\mathcal{V} = v_1, \dots, v_m$ of elements of V we define $\mu_{\mathcal{V}} \colon \mathbb{R}^m \to V$ by $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = \sum_i \lambda_i v_i$.
- If $\mathcal{V}$ is a basis then $\mu_{\mathcal{V}}$ is an isomorphism.
- Suppose we have a linear map $\alpha \colon V \to W$, a basis $\mathcal{V} = v_1, \dots, v_m$ for V and a basis $\mathcal{W} = w_1, \dots, w_n$ for W. Then there is a unique matrix $A = (a_{ij})$ such that $\alpha(v_j) = \sum_i a_{ij} w_i$.

Reminders

- Given an $n \times m$ matrix A, we define a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{x}) = A\mathbf{x}$.
- Every linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ is ϕ_A for some A.
- Given a vector space V and a list $\mathcal{V} = v_1, \ldots, v_m$ of elements of V we define $\mu_{\mathcal{V}} \colon \mathbb{R}^m \to V$ by $\mu_{\mathcal{V}}(\boldsymbol{\lambda}) = \sum_i \lambda_i v_i$.
- If $\mathcal{V}$ is a basis then $\mu_{\mathcal{V}}$ is an isomorphism.
- Suppose we have a linear map $\alpha \colon V \to W$, a basis $\mathcal{V} = v_1, \ldots, v_m$ for V and a basis $\mathcal{W} = w_1, \ldots, w_n$ for W. Then there is a unique matrix $A = (a_{ij})$ such that $\alpha(v_j) = \sum_i a_{ij} w_i$. This is called *the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$.*

Matrices for linear maps

Matrices for linear maps

Proposition ??: For any $\mathbf{x} \in \mathbb{R}^m$, we have $\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \alpha(\mu_{\mathcal{V}}(\mathbf{x}))$

Matrices for linear maps

Proposition ??: For any $\mathbf{x} \in \mathbb{R}^m$, we have $\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \alpha(\mu_{\mathcal{V}}(\mathbf{x}))$, so the two routes around the square below are the same:

$$\begin{tikzcd} \mathbb{R}^m \arrow[r, "\phi_A"] \arrow[d, "\mu_{\mathcal{V}}"'] & \mathbb{R}^n \arrow[d, "\mu_{\mathcal{W}}"] \\ V \arrow[r, "\alpha"'] & W \end{tikzcd}$$

Matrices for linear maps

Proposition ??: For any $\mathbf{x} \in \mathbb{R}^m$, we have $\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \alpha(\mu_{\mathcal{V}}(\mathbf{x}))$, so the two routes around the square below are the same:

$$\begin{tikzcd}
\mathbb{R}^m \arrow[r, "\phi_A"] \arrow[d, "\mu_{\mathcal{V}}"'] & \mathbb{R}^n \arrow[d, "\mu_{\mathcal{W}}"] \\
V \arrow[r, "\alpha"'] & W
\end{tikzcd}$$

(This is often expressed by saying that the square *commutes*.)

Matrices for linear maps

Proposition ??: For any $\mathbf{x} \in \mathbb{R}^m$, we have $\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \alpha(\mu_{\mathcal{V}}(\mathbf{x}))$, so the two routes around the square below are the same:

$$\begin{tikzcd} \mathbb{R}^m \arrow[r, "\phi_A"] \arrow[d, "\mu_{\mathcal{V}}"'] & \mathbb{R}^n \arrow[d, "\mu_{\mathcal{W}}"] \\ V \arrow[r, "\alpha"'] & W \end{tikzcd}$$

(This is often expressed by saying that the square *commutes*.)

Proof: We will do the case where $m = 2$ and $n = 3$; the general case is essentially the same, but with more complicated notation.

Proposition ??: For any $\mathbf{x} \in \mathbb{R}^m$, we have $\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \alpha(\mu_{\mathcal{V}}(\mathbf{x}))$, so the two routes around the square below are the same:

$$\begin{array}{ccc} \mathbb{R}^m & \xrightarrow{\phi_A} & \mathbb{R}^n \\ {\scriptstyle \mu_{\mathcal{V}}}\downarrow & & \downarrow{\scriptstyle \mu_{\mathcal{W}}} \\ V & \xrightarrow[\alpha]{} & W \end{array}$$

(This is often expressed by saying that the square *commutes*.)

Proof: We will do the case where $m = 2$ and $n = 3$; the general case is essentially the same, but with more complicated notation. In our case, v_1, v_2 is a basis for V, and w_1, w_2, w_3 is a basis for W.

Proposition ??: For any $\mathbf{x} \in \mathbb{R}^m$, we have $\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \alpha(\mu_{\mathcal{V}}(\mathbf{x}))$, so the two routes around the square below are the same:

$$\begin{tikzcd} \mathbb{R}^m \arrow[r, "\phi_A"] \arrow[d, "\mu_{\mathcal{V}}"'] & \mathbb{R}^n \arrow[d, "\mu_{\mathcal{W}}"] \\ V \arrow[r, "\alpha"'] & W \end{tikzcd}$$

(This is often expressed by saying that the square *commutes*.)

Proof: We will do the case where $m = 2$ and $n = 3$; the general case is essentially the same, but with more complicated notation. In our case, v_1, v_2 is a basis for V, and w_1, w_2, w_3 is a basis for W. From the definitions of a_{ij} and A, we have

$$\begin{aligned} \alpha(v_1) &= a_{11} w_1 + a_{21} w_2 + a_{31} w_3 \\ \alpha(v_2) &= a_{12} w_1 + a_{22} w_2 + a_{32} w_3 \end{aligned} \qquad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Matrices for linear maps

Matrices for linear maps

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$.

Matrices for linear maps

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$).

Matrices for linear maps

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$). It follows that

$$\alpha(\mu_{\mathcal{V}}(\mathbf{x})) = \alpha(x_1 v_1 + x_2 v_2)$$

Matrices for linear maps

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$). It follows that

$$\alpha(\mu_{\mathcal{V}}(\mathbf{x})) = \alpha(x_1 v_1 + x_2 v_2) = x_1 \alpha(v_1) + x_2 \alpha(v_2)$$

Matrices for linear maps

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$). It follows that

$$\begin{aligned} \alpha(\mu_{\mathcal{V}}(\mathbf{x})) &= \alpha(x_1 v_1 + x_2 v_2) = x_1 \alpha(v_1) + x_2 \alpha(v_2) \\ &= x_1(a_{11} w_1 + a_{21} w_2 + a_{31} w_3) + x_2(a_{12} w_1 + a_{22} w_2 + a_{32} w_3) \end{aligned}$$

Matrices for linear maps

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$). It follows that

$$\begin{aligned}
\alpha(\mu_{\mathcal{V}}(\mathbf{x})) &= \alpha(x_1 v_1 + x_2 v_2) = x_1 \alpha(v_1) + x_2 \alpha(v_2) \\
&= x_1(a_{11} w_1 + a_{21} w_2 + a_{31} w_3) + x_2(a_{12} w_1 + a_{22} w_2 + a_{32} w_3) \\
&= (a_{11} x_1 + a_{12} x_2) w_1 + (a_{21} x_1 + a_{22} x_2) w_2 + (a_{31} x_1 + a_{32} x_2) w_3
\end{aligned}$$

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$). It follows that

$$\begin{aligned} \alpha(\mu_{\mathcal{V}}(\mathbf{x})) &= \alpha(x_1 v_1 + x_2 v_2) = x_1 \alpha(v_1) + x_2 \alpha(v_2) \\ &= x_1(a_{11} w_1 + a_{21} w_2 + a_{31} w_3) + x_2(a_{12} w_1 + a_{22} w_2 + a_{32} w_3) \\ &= (a_{11} x_1 + a_{12} x_2) w_1 + (a_{21} x_1 + a_{22} x_2) w_2 + (a_{31} x_1 + a_{32} x_2) w_3 \end{aligned}$$

On the other hand, we have

$$\phi_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} x_1 + a_{12} x_2 \\ a_{21} x_1 + a_{22} x_2 \\ a_{31} x_1 + a_{32} x_2 \end{bmatrix},$$

Now consider a vector $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] \in \mathbb{R}^2$. We have $\mu_{\mathcal{V}}(\mathbf{x}) = x_1 v_1 + x_2 v_2$ (by the definition of $\mu_{\mathcal{V}}$). It follows that

$$\begin{aligned}
\alpha(\mu_{\mathcal{V}}(\mathbf{x})) &= \alpha(x_1 v_1 + x_2 v_2) = x_1 \alpha(v_1) + x_2 \alpha(v_2) \\
&= x_1(a_{11} w_1 + a_{21} w_2 + a_{31} w_3) + x_2(a_{12} w_1 + a_{22} w_2 + a_{32} w_3) \\
&= (a_{11} x_1 + a_{12} x_2) w_1 + (a_{21} x_1 + a_{22} x_2) w_2 + (a_{31} x_1 + a_{32} x_2) w_3
\end{aligned}$$

On the other hand, we have

$$\phi_A(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} x_1 + a_{12} x_2 \\ a_{21} x_1 + a_{22} x_2 \\ a_{31} x_1 + a_{32} x_2 \end{bmatrix},$$

so

$$\mu_{\mathcal{W}}(\phi_A(\mathbf{x})) = \mu_{\mathcal{W}} \begin{bmatrix} a_{11} x_1 + a_{12} x_2 \\ a_{21} x_1 + a_{22} x_2 \\ a_{31} x_1 + a_{32} x_2 \end{bmatrix} = \begin{array}{l} (a_{11} x_1 + a_{12} x_2) w_1 + \\ (a_{21} x_1 + a_{22} x_2) w_2 + \\ (a_{31} x_1 + a_{32} x_2) w_3 \end{array} = \alpha(\mu_{\mathcal{V}}(\mathbf{x})).$$

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta \colon U \to W$).

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta \colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W.

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$.

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta \colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof:

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$.

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad \beta(u_k) = \sum_j b_{jk} v_j$$

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad \beta(u_k) = \sum_j b_{jk} v_j$$

$$\alpha\beta(u_k) = \alpha\left(\textstyle\sum_j b_{jk} v_j\right)$$

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad\qquad \beta(u_k) = \sum_j b_{jk} v_j$$

$$\alpha\beta(u_k) = \alpha\left(\sum_j b_{jk} v_j\right) = \sum_j b_{jk}\alpha(v_j)$$

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad \beta(u_k) = \sum_j b_{jk} v_j$$

$$\alpha\beta(u_k) = \alpha\left(\textstyle\sum_j b_{jk} v_j\right) = \textstyle\sum_j b_{jk}\alpha(v_j) = \sum_j b_{jk} \sum_i a_{ij} w_i$$

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad\qquad \beta(u_k) = \sum_j b_{jk} v_j$$

$$\alpha\beta(u_k) = \alpha\left(\textstyle\sum_j b_{jk} v_j\right) = \textstyle\sum_j b_{jk}\alpha(v_j) = \sum_j b_{jk} \sum_i a_{ij} w_i = \sum_i \left(\sum_j a_{ij} b_{jk}\right) w_i$$

Composition and matrices

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta \colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad\qquad \beta(u_k) = \sum_j b_{jk} v_j$$

$$\alpha\beta(u_k) = \alpha\left(\sum_j b_{jk} v_j\right) = \sum_j b_{jk}\alpha(v_j) = \sum_j b_{jk} \sum_i a_{ij} w_i = \\ \sum_i \left(\sum_j a_{ij} b_{jk}\right) w_i = \sum_i c_{ik} w_i.$$

Proposition ??: Suppose we have linear maps $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ (which can therefore be composed to give a linear map $\alpha\beta\colon U \to W$). Suppose that we have bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ for U, V and W. Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$, and let B be the matrix of β with respect to $\mathcal{U}$ and $\mathcal{V}$. Then the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is AB.

Proof: By the definition of matrix multiplication, the matrix $C = AB$ has entries $c_{ik} = \sum_j a_{ij} b_{jk}$. By the definitions of A and B, we have

$$\alpha(v_j) = \sum_i a_{ij} w_i \qquad\qquad \beta(u_k) = \sum_j b_{jk} v_j$$

$$\alpha\beta(u_k) = \alpha\left(\textstyle\sum_j b_{jk} v_j\right) = \textstyle\sum_j b_{jk}\alpha(v_j) = \sum_j b_{jk} \sum_i a_{ij} w_i = \\ \textstyle\sum_i \left(\sum_j a_{ij} b_{jk}\right) w_i = \sum_i c_{ik} w_i.$$

This means precisely that C is the matrix of $\alpha\beta$ with respect to $\mathcal{U}$ and $\mathcal{W}$. □

Change of basis

Change of basis

Definition ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{V}' = v'_1, \ldots, v'_n$.

Change of basis

Definition ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v_1', \dots, v_n'$. We then have

$$v_j' = p_{1j} v_1 + \cdots + p_{nj} v_n$$

for some scalars p_{ij}.

Change of basis

Definition ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v'_1, \dots, v'_n$. We then have

$$v'_j = p_{1j} v_1 + \cdots + p_{nj} v_n$$

for some scalars p_{ij}. Let P be the $n \times n$ matrix with entries p_{ij}.

Change of basis

Definition ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{V}' = v'_1, \ldots, v'_n$. We then have

$$v'_j = p_{1j} v_1 + \cdots + p_{nj} v_n$$

for some scalars p_{ij}. Let P be the $n \times n$ matrix with entries p_{ij}. This is called the *change-of-basis* matrix from $\mathcal{V}$ to $\mathcal{V}'$.

Definition ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v'_1, \dots, v'_n$. We then have

$$v'_j = p_{1j} v_1 + \cdots + p_{nj} v_n$$

for some scalars p_{ij}. Let P be the $n \times n$ matrix with entries p_{ij}. This is called the *change-of-basis* matrix from $\mathcal{V}$ to $\mathcal{V}'$. One can check that it is invertible, and that P^{-1} is the change of basis matrix from $\mathcal{V}'$ to $\mathcal{V}$.

An example

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$v_1 = x^3 \qquad v_2 = x^2 \qquad v_3 = x \qquad v_4 = 1$$

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$\begin{aligned} v_1 &= x^3 & v_2 &= x^2 & v_3 &= x & v_4 &= 1 \\ v_1' &= x^3 + x^2 + x + 1 & v_2' &= x^3 + x^2 + x & v_3' &= x^3 + x^2 & v_4' &= x^3 \end{aligned}$$

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$\begin{aligned} v_1 &= x^3 \\ v_1' &= x^3 + x^2 + x + 1 \end{aligned} \qquad \begin{aligned} v_2 &= x^2 \\ v_2' &= x^3 + x^2 + x \end{aligned} \qquad \begin{aligned} v_3 &= x \\ v_3' &= x^3 + x^2 \end{aligned} \qquad \begin{aligned} v_4 &= 1 \\ v_4' &= x^3 \end{aligned}$$

Then

$$v_1' = \color{red}{1}.v_1 + \color{red}{1}.v_2 + \color{red}{1}.v_3 + \color{red}{1}.v_4$$

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$\begin{aligned} v_1 &= x^3 \\ v_1' &= x^3 + x^2 + x + 1 \end{aligned} \qquad \begin{aligned} v_2 &= x^2 \\ v_2' &= x^3 + x^2 + x \end{aligned} \qquad \begin{aligned} v_3 &= x \\ v_3' &= x^3 + x^2 \end{aligned} \qquad \begin{aligned} v_4 &= 1 \\ v_4' &= x^3 \end{aligned}$$

Then

$$v_1' = 1.v_1 + 1.v_2 + 1.v_3 + 1.v_4$$
$$v_2' = 1.v_1 + 1.v_2 + 1.v_3 + 0.v_4$$

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$\begin{aligned} v_1 &= x^3 \\ v_1' &= x^3 + x^2 + x + 1 \end{aligned} \qquad \begin{aligned} v_2 &= x^2 \\ v_2' &= x^3 + x^2 + x \end{aligned} \qquad \begin{aligned} v_3 &= x \\ v_3' &= x^3 + x^2 \end{aligned} \qquad \begin{aligned} v_4 &= 1 \\ v_4' &= x^3 \end{aligned}$$

Then

$$\begin{aligned} v_1' &= 1.v_1 + 1.v_2 + 1.v_3 + 1.v_4 \\ v_2' &= 1.v_1 + 1.v_2 + 1.v_3 + 0.v_4 \\ v_3' &= 1.v_1 + 1.v_2 + 0.v_3 + 0.v_4 \end{aligned}$$

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$\begin{array}{llllllll} v_1 & = x^3 & v_2 & = x^2 & v_3 & = x & v_4 & = 1 \\ v_1' & = x^3 + x^2 + x + 1 & v_2' & = x^3 + x^2 + x & v_3' & = x^3 + x^2 & v_4' & = x^3 \end{array}$$

Then

$$\begin{aligned} v_1' &= 1.v_1 + 1.v_2 + 1.v_3 + 1.v_4 \\ v_2' &= 1.v_1 + 1.v_2 + 1.v_3 + 0.v_4 \\ v_3' &= 1.v_1 + 1.v_2 + 0.v_3 + 0.v_4 \\ v_4' &= 1.v_1 + 0.v_2 + 0.v_3 + 0.v_4 \end{aligned}$$

An example

Consider the following bases of $\mathbb{R}[x]_{\leq 3}$:

$$\begin{aligned} v_1 &= x^3 & v_2 &= x^2 & v_3 &= x & v_4 &= 1 \\ v_1' &= x^3 + x^2 + x + 1 & v_2' &= x^3 + x^2 + x & v_3' &= x^3 + x^2 & v_4' &= x^3 \end{aligned}$$

Then

$$\begin{aligned} v_1' &= 1.v_1 + 1.v_2 + 1.v_3 + 1.v_4 \\ v_2' &= 1.v_1 + 1.v_2 + 1.v_3 + 0.v_4 \\ v_3' &= 1.v_1 + 1.v_2 + 0.v_3 + 0.v_4 \\ v_4' &= 1.v_1 + 0.v_2 + 0.v_3 + 0.v_4 \end{aligned}$$

so the change of basis matrix is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Another example

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$\begin{aligned} A_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & A_2 &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} & A_3 &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} & A_4 &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ A_1' &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} & A_2' &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & A_3' &= \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} & A_4' &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$\begin{aligned} A_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & A_2 &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} & A_3 &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} & A_4 &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ A'_1 &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} & A'_2 &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & A'_3 &= \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} & A'_4 &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Then

$$A'_1 = 2.A_1 + (-2).A_2 + 0.A_3 + 1.A_4$$

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$A'_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad A'_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad A'_3 = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \quad A'_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Then

$$A'_1 = 2.A_1 + (-2).A_2 + 0.A_3 + 1.A_4$$
$$A'_2 = 0.A_1 + 0.A_2 + 2.A_3 + (-1).A_4$$

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$
\begin{array}{llll}
A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} & A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} & A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\
A_1' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} & A_2' = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & A_3' = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} & A_4' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\end{array}
$$

Then

$$
\begin{aligned}
A_1' &= 2.A_1 + (-2).A_2 + 0.A_3 + 1.A_4 \\
A_2' &= 0.A_1 + 0.A_2 + 2.A_3 + (-1).A_4 \\
A_3' &= 0.A_1 + 2.A_2 + 0.A_3 + (-1).A_4
\end{aligned}
$$

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$A_1' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad A_2' = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad A_3' = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \quad A_4' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Then

$$\begin{aligned}
A_1' &= 2.A_1 + (-2).A_2 + 0.A_3 + 1.A_4 \\
A_2' &= 0.A_1 + 0.A_2 + 2.A_3 + (-1).A_4 \\
A_3' &= 0.A_1 + 2.A_2 + 0.A_3 + (-1).A_4 \\
A_4' &= 0.A_1 + 0.A_2 + 0.A_3 + 1.A_4
\end{aligned}$$

Another example

Consider the following bases of $M_2\mathbb{R}$:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$A_1' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad A_2' = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad A_3' = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \qquad A_4' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Then

$$\begin{aligned} A_1' &= 2.A_1 + (-2).A_2 + 0.A_3 + 1.A_4 \\ A_2' &= 0.A_1 + 0.A_2 + 2.A_3 + (-1).A_4 \\ A_3' &= 0.A_1 + 2.A_2 + 0.A_3 + (-1).A_4 \\ A_4' &= 0.A_1 + 0.A_2 + 0.A_3 + 1.A_4 \end{aligned}$$

so the change of basis matrix is

$$P = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Change of basis

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{V}' = v'_1, \ldots, v'_n$.

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v'_1, \dots, v'_n$. Let P be the change of basis matrix, so

$$v'_j = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{V}' = v_1', \ldots, v_n'$. Let P be the change of basis matrix, so

$$v_j' = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v_1', \dots, v_n'$. Let P be the change of basis matrix, so

$$v_j' = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\
& V &
\end{tikzcd}$$

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v_1', \dots, v_n'$. Let P be the change of basis matrix, so

$$v_j' = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\
& V &
\end{tikzcd}$$

Proof: We have $P\mathbf{x} = \mathbf{y}$, where $y_i = \sum_j p_{ij} x_j$.

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v_1', \dots, v_n'$. Let P be the change of basis matrix, so

$$v_j' = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\
& V &
\end{tikzcd}$$

Proof: We have $P\mathbf{x} = \mathbf{y}$, where $y_i = \sum_j p_{ij} x_j$. Thus

$$\mu_{\mathcal{V}}(P\mathbf{x}) = \sum_i y_i v_i$$

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v'_1, \dots, v'_n$. Let P be the change of basis matrix, so

$$v'_j = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\
& V &
\end{tikzcd}$$

Proof: We have $P\mathbf{x} = \mathbf{y}$, where $y_i = \sum_j p_{ij} x_j$. Thus

$$\mu_{\mathcal{V}}(P\mathbf{x}) = \sum_i y_i v_i = \sum_{i,j} p_{ij} x_j v_i$$

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{V}' = v_1', \ldots, v_n'$. Let P be the change of basis matrix, so

$$v_j' = p_{1j} v_1 + \cdots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd} \mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\ & V & \end{tikzcd}$$

Proof: We have $P\mathbf{x} = \mathbf{y}$, where $y_i = \sum_j p_{ij} x_j$. Thus

$$\mu_{\mathcal{V}}(P\mathbf{x}) = \sum_i y_i v_i = \sum_{i,j} p_{ij} x_j v_i = \sum_j x_j \left(\sum_i p_{ij} v_i \right)$$

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{V}' = v'_1, \ldots, v'_n$. Let P be the change of basis matrix, so

$$v'_j = p_{1j}v_1 + \cdots + p_{nj}v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd} \mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\ & V & \end{tikzcd}$$

Proof: We have $P\mathbf{x} = \mathbf{y}$, where $y_i = \sum_j p_{ij}x_j$. Thus

$$\mu_{\mathcal{V}}(P\mathbf{x}) = \sum_i y_i v_i = \sum_{i,j} p_{ij}x_j v_i = \sum_j x_j \left(\sum_i p_{ij}v_i\right) = \sum_j x_j v'_j$$

Change of basis

Lemma ??: Let V be a finite-dimensional vector space, with two different bases $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{V}' = v_1', \dots, v_n'$. Let P be the change of basis matrix, so

$$v_j' = p_{1j} v_1 + \dots + p_{nj} v_n.$$

Then for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mu_{\mathcal{V}}(\phi_P(\mathbf{x})) = \mu_{\mathcal{V}}(P\mathbf{x}) = \mu_{\mathcal{V}'}(\mathbf{x})$, so the following diagram commutes:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[rr, "\phi_P"] \arrow[dr, "\mu_{\mathcal{V}'}"'] & & \mathbb{R}^n \arrow[dl, "\mu_{\mathcal{V}}"] \\
& V &
\end{tikzcd}$$

Proof: We have $P\mathbf{x} = \mathbf{y}$, where $y_i = \sum_j p_{ij} x_j$. Thus

$$\mu_{\mathcal{V}}(P\mathbf{x}) = \sum_i y_i v_i = \sum_{i,j} p_{ij} x_j v_i = \sum_j x_j \left(\sum_i p_{ij} v_i \right) = \sum_j x_j v_j' = \mu_{\mathcal{V}'}(\mathbf{x}).$$

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$

Change of basis

Proposition ??: Let $\alpha \colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P
and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$.

Change of basis

Proposition ??: Let $\alpha \colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof:

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.

Change of basis

Proposition ??: Let $\alpha \colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\mu_{\mathcal{W}}(QA'\mathbf{x})$$

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\mu_{\mathcal{W}}(QA'\mathbf{x}) = \mu_{\mathcal{W}'}(A'\mathbf{x}) \qquad \text{(Lemma \textbf{??})}$$

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned} \mu_{\mathcal{W}}(QA'\mathbf{x}) &= \mu_{\mathcal{W}'}(A'\mathbf{x}) && \text{(Lemma \textbf{??})} \\ &= \alpha(\mu_{\mathcal{V}'}(\mathbf{x})) && \text{(Proposition \textbf{??})} \end{aligned}$$

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\mu_{\mathcal{W}}(QA'\mathbf{x}) &= \mu_{\mathcal{W}'}(A'\mathbf{x}) && \text{(Lemma \textbf{??})} \\
&= \alpha(\mu_{\mathcal{V}'}(\mathbf{x})) && \text{(Proposition \textbf{??})} \\
&= \alpha(\mu_{\mathcal{V}}(P\mathbf{x})) && \text{(Lemma \textbf{??})}
\end{aligned}$$

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\mu_{\mathcal{W}}(QA'\mathbf{x}) &= \mu_{\mathcal{W}'}(A'\mathbf{x}) && \text{(Lemma \textbf{??})}\\
&= \alpha(\mu_{\mathcal{V}'}(\mathbf{x})) && \text{(Proposition \textbf{??})}\\
&= \alpha(\mu_{\mathcal{V}}(P\mathbf{x})) && \text{(Lemma \textbf{??})}\\
&= \mu_{\mathcal{W}}(AP\mathbf{x}) && \text{(Proposition \textbf{??})}.
\end{aligned}$$

Change of basis

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\mu_{\mathcal{W}}(QA'\mathbf{x}) &= \mu_{\mathcal{W}'}(A'\mathbf{x}) && \text{(Lemma \textbf{??})} \\
&= \alpha(\mu_{\mathcal{V}'}(\mathbf{x})) && \text{(Proposition \textbf{??})} \\
&= \alpha(\mu_{\mathcal{V}}(P\mathbf{x})) && \text{(Lemma \textbf{??})} \\
&= \mu_{\mathcal{W}}(AP\mathbf{x}) && \text{(Proposition \textbf{??})}.
\end{aligned}$$

This shows that $\mu_{\mathcal{W}}((QA' - AP)\mathbf{x}) = 0$.

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\mu_{\mathcal{W}}(QA'\mathbf{x}) &= \mu_{\mathcal{W}'}(A'\mathbf{x}) && \text{(Lemma ??)} \\
&= \alpha(\mu_{\mathcal{V}'}(\mathbf{x})) && \text{(Proposition ??)} \\
&= \alpha(\mu_{\mathcal{V}}(P\mathbf{x})) && \text{(Lemma ??)} \\
&= \mu_{\mathcal{W}}(AP\mathbf{x}) && \text{(Proposition ??).}
\end{aligned}$$

This shows that $\mu_{\mathcal{W}}((QA' - AP)\mathbf{x}) = 0$. Moreover, $\mathcal{W}$ is linearly independent, so $\mu_{\mathcal{W}}$ is injective and has trivial kernel, so $(QA' - AP)\mathbf{x} = 0$.

Proposition ??: Let $\alpha\colon V \to W$ be a linear map.
Suppose we have two bases $\mathcal{V}$ and $\mathcal{V}'$ for V, with change-of basis matrix P and two bases $\mathcal{W}$ and $\mathcal{W}'$ for W, with change-of-basis matrix Q.
Let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{W}$,
and let A' be the matrix with respect to $\mathcal{V}'$ and $\mathcal{W}'$. Then $A' = Q^{-1}AP$.

Proof: We actually prove that $QA' = AP$, which comes to the same thing.
For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\mu_{\mathcal{W}}(QA'\mathbf{x}) &= \mu_{\mathcal{W}'}(A'\mathbf{x}) && \text{(Lemma \textbf{??})} \\
&= \alpha(\mu_{\mathcal{V}'}(\mathbf{x})) && \text{(Proposition \textbf{??})} \\
&= \alpha(\mu_{\mathcal{V}}(P\mathbf{x})) && \text{(Lemma \textbf{??})} \\
&= \mu_{\mathcal{W}}(AP\mathbf{x}) && \text{(Proposition \textbf{??})}.
\end{aligned}$$

This shows that $\mu_{\mathcal{W}}((QA' - AP)\mathbf{x}) = 0$. Moreover, $\mathcal{W}$ is linearly independent, so $\mu_{\mathcal{W}}$ is injective and has trivial kernel, so $(QA' - AP)\mathbf{x} = 0$. This applies for *any* vector $\mathbf{x}$, so the matrix $QA' - AP$ must be zero, as claimed.

Change of basis

The upshot is that all parts of the following diagram commute:

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[rrr, "\phi_{A'}"] \arrow[dd, "\mu_{\mathcal{V}'}"'] \arrow[dr, "\phi_P"] & & & \mathbb{R}^m \arrow[dl, "\phi_Q"'] \arrow[dd, "\mu_{\mathcal{W}'}"] \\
& \mathbb{R}^n \arrow[r, "\phi_A"'] \arrow[dl, "\mu_{\mathcal{V}}"] & \mathbb{R}^m \arrow[dr, "\mu_{\mathcal{W}}"'] & \\
V \arrow[rrr, "\alpha"'] & & & W
\end{tikzcd}
$$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself.

Traces and determinants

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α.

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\text{trace}(\alpha) = \text{trace}(A) \qquad \det(\alpha) = \det(A)$$
$$\text{char}(\alpha)(t) = \text{char}(A)(t) = \det(tI - A).$$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\mathrm{trace}(\alpha) = \mathrm{trace}(A) \qquad \det(\alpha) = \det(A)$$
$$\mathrm{char}(\alpha)(t) = \mathrm{char}(A)(t) = \det(tI - A).$$

This is not obviously well-defined

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\mathrm{trace}(\alpha) = \mathrm{trace}(A) \qquad \det(\alpha) = \det(A)$$
$$\mathrm{char}(\alpha)(t) = \mathrm{char}(A)(t) = \det(tI - A).$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'?

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\begin{gathered} \operatorname{trace}(\alpha) = \operatorname{trace}(A) \qquad\qquad \det(\alpha) = \det(A) \\ \operatorname{char}(\alpha)(t) = \operatorname{char}(A)(t) = \det(tI - A). \end{gathered}$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\begin{gathered}\operatorname{trace}(\alpha) = \operatorname{trace}(A) \qquad\qquad \det(\alpha) = \det(A) \\ \operatorname{char}(\alpha)(t) = \operatorname{char}(A)(t) = \det(tI - A).\end{gathered}$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$, and it follows that $P^{-1}(tI - A)P = tI - A'$.

Traces and determinants

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\begin{aligned} \operatorname{trace}(\alpha) &= \operatorname{trace}(A) \qquad\qquad \det(\alpha) = \det(A) \\ \operatorname{char}(\alpha)(t) &= \operatorname{char}(A)(t) = \det(tI - A). \end{aligned}$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$, and it follows that $P^{-1}(tI - A)P = tI - A'$. Using the rules $\operatorname{trace}(MN) = \operatorname{trace}(NM)$ and $\det(MN) = \det(M)\det(N)$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\begin{gathered}\operatorname{trace}(\alpha) = \operatorname{trace}(A) \qquad\qquad \det(\alpha) = \det(A) \\ \operatorname{char}(\alpha)(t) = \operatorname{char}(A)(t) = \det(tI - A).\end{gathered}$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$, and it follows that $P^{-1}(tI - A)P = tI - A'$. Using the rules $\operatorname{trace}(MN) = \operatorname{trace}(NM)$ and $\det(MN) = \det(M)\det(N)$ we see that

$$\operatorname{trace}(A') = \operatorname{trace}(P^{-1}(AP)) = \operatorname{trace}((AP)P^{-1}) = \operatorname{trace}(A(PP^{-1})) = \operatorname{trace}(A)$$

Traces and determinants

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\begin{aligned} \operatorname{trace}(\alpha) = \operatorname{trace}(A) \qquad\qquad \det(\alpha) = \det(A) \\ \operatorname{char}(\alpha)(t) = \operatorname{char}(A)(t) = \det(tI - A). \end{aligned}$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$, and it follows that $P^{-1}(tI - A)P = tI - A'$. Using the rules $\operatorname{trace}(MN) = \operatorname{trace}(NM)$ and $\det(MN) = \det(M)\det(N)$ we see that

$$\operatorname{trace}(A') = \operatorname{trace}(P^{-1}(AP)) = \operatorname{trace}((AP)P^{-1}) = \operatorname{trace}(A(PP^{-1})) = \operatorname{trace}(A)$$

$$\det(A') = \det(P)^{-1}\det(A)\det(P) = \det(A)$$

Traces and determinants

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\mathrm{trace}(\alpha) = \mathrm{trace}(A) \qquad \det(\alpha) = \det(A)$$
$$\mathrm{char}(\alpha)(t) = \mathrm{char}(A)(t) = \det(tI - A).$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$, and it follows that $P^{-1}(tI - A)P = tI - A'$. Using the rules $\mathrm{trace}(MN) = \mathrm{trace}(NM)$ and $\det(MN) = \det(M)\det(N)$ we see that

$$\begin{aligned}
\mathrm{trace}(A') &= \mathrm{trace}(P^{-1}(AP)) = \mathrm{trace}((AP)P^{-1}) = \mathrm{trace}(A(PP^{-1})) = \mathrm{trace}(A) \\
\det(A') &= \det(P)^{-1}\det(A)\det(P) = \det(A) \\
\mathrm{char}(A')(t) &= \det(P)^{-1}\det(tI - A)\det(P) = \mathrm{char}(A)(t).
\end{aligned}$$

Remark ??: Suppose we have a finite-dimensional vector space V and a linear map α from V to itself. We can now define the trace, determinant and characteristic polynomial of α. We pick any basis $\mathcal{V}$, let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$, and put

$$\operatorname{trace}(\alpha) = \operatorname{trace}(A) \qquad \det(\alpha) = \det(A)$$
$$\operatorname{char}(\alpha)(t) = \operatorname{char}(A)(t) = \det(tI - A).$$

This is not obviously well-defined: what if we used a different basis, say $\mathcal{V}'$, giving a different matrix, say A'? The proposition tells us that $P^{-1}AP = A'$, and it follows that $P^{-1}(tI - A)P = tI - A'$. Using the rules $\operatorname{trace}(MN) = \operatorname{trace}(NM)$ and $\det(MN) = \det(M)\det(N)$ we see that

$$\begin{aligned}
\operatorname{trace}(A') &= \operatorname{trace}(P^{-1}(AP)) = \operatorname{trace}((AP)P^{-1}) = \operatorname{trace}(A(PP^{-1})) = \operatorname{trace}(A)\\
\det(A') &= \det(P)^{-1}\det(A)\det(P) = \det(A)\\
\operatorname{char}(A')(t) &= \det(P)^{-1}\det(tI - A)\det(P) = \operatorname{char}(A)(t).
\end{aligned}$$

This shows that definitions are in fact basis-independent.

An example

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$.

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\operatorname{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$

$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3 a_1 - 0.(-a_2)) = 0$

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\mathrm{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$

$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3 a_1 - 0.(-a_2)) = 0$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$.

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\mathrm{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$
$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3a_1 - 0.(-a_2)) = 0$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map β has matrix $B' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\text{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$
$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3a_1 - 0.(-a_2)) = 0$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map β has matrix $B' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.
It is easy to see that $\text{trace}(B') = 0 = \det(B')$.

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\operatorname{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$
$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3a_1 - 0.(-a_2)) = 0$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$.
With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map β has matrix $B' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.
It is easy to see that $\operatorname{trace}(B') = 0 = \det(B')$.
Either way we have $\operatorname{trace}(\beta) = 0 = \det(\beta)$.

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\operatorname{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$
$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3a_1 - 0.(-a_2)) = 0$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map β has matrix $B' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.
It is easy to see that $\operatorname{trace}(B') = 0 = \det(B')$.
Either way we have $\operatorname{trace}(\beta) = 0 = \det(\beta)$.
We also find that $\operatorname{char}(\beta)(t) = \operatorname{char}(B')(t) = t^3 + t$.

An example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\beta\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\beta(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The matrix B of β with respect to the standard basis is found as follows:

$$\beta(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \quad \beta(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \quad \beta(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

We have $\operatorname{trace}(B) = 0$ and $\det(B) = 0.\det\begin{bmatrix} 0 & -a_1 \\ a_1 & 0 \end{bmatrix} - (-a_3).\det\begin{bmatrix} a_3 & -a_1 \\ -a_2 & 0 \end{bmatrix} + a_2.\det\begin{bmatrix} a_3 & 0 \\ -a_2 & a_1 \end{bmatrix} =$
$0 - (-a_3)(a_3.0 - (-a_2)(-a_1)) + a_2(a_3a_1 - 0.(-a_2)) = 0$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map β has matrix $B' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$.
It is easy to see that $\operatorname{trace}(B') = 0 = \det(B')$.
Either way we have $\operatorname{trace}(\beta) = 0 = \det(\beta)$.
We also find that $\operatorname{char}(\beta)(t) = \operatorname{char}(B')(t) = t^3 + t$.
This is much more complicated using B.

Another example

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$.

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi : \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\operatorname{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2$.

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\mathrm{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2$.

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$.

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\mathrm{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2$.

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map π has matrix $P' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\operatorname{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2$.

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map π has matrix $P' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
It is easy to see that $\operatorname{trace}(P') = 2$.

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\operatorname{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2$.

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$.
With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map π has matrix $P' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
It is easy to see that $\operatorname{trace}(P') = 2$.
Either way we have $\operatorname{trace}(\pi) = 2$.

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a} \rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\operatorname{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2$.

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map π has matrix $P' = \begin{bmatrix} 0&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}$.
It is easy to see that $\operatorname{trace}(P') = 2$.
Either way we have $\operatorname{trace}(\pi) = 2$.
We also find that $\det(\pi) = \det(P') = 0$ and $\operatorname{char}(\pi)(t) = \operatorname{char}(P')(t) = t(t-1)^2$.

Another example

Example ??: Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector, and define $\pi\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\pi(\mathbf{x}) = \mathbf{x} - \langle \mathbf{x}, \mathbf{a}\rangle \mathbf{a}$. The matrix P of π with respect to the standard basis is found as follows:

$$\pi(\mathbf{e}_1) = \begin{bmatrix} 1-a_1^2 \\ -a_1a_2 \\ -a_1a_3 \end{bmatrix} \quad \pi(\mathbf{e}_2) = \begin{bmatrix} -a_2a_1 \\ 1-a_2^2 \\ -a_2a_3 \end{bmatrix} \quad \pi(\mathbf{e}_3) = \begin{bmatrix} -a_3a_1 \\ -a_3a_2 \\ 1-a_3^2 \end{bmatrix} \quad P = \begin{bmatrix} 1-a_1^2 & -a_1a_2 & -a_1a_3 \\ -a_1a_2 & 1-a_2^2 & -a_2a_3 \\ -a_1a_3 & -a_2a_3 & 1-a_3^2 \end{bmatrix}$$

We have $\operatorname{trace}(P) = 1 - a_1^2 + 1 - a_2^2 + 1 - a_3^2 = 3 - (a_1^2 + a_2^2 + a_3^2) = 2.$

We can instead choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and then put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. With respect to the basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$, the map π has matrix $P' = \left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]$.
It is easy to see that $\operatorname{trace}(P') = 2$.
Either way we have $\operatorname{trace}(\pi) = 2$.
We also find that $\det(\pi) = \det(P') = 0$ and $\operatorname{char}(\pi)(t) = \operatorname{char}(P')(t) = t(t-1)^2$.
This is much more complicated using P.

The determinant criterion

Remark ??: Suppose again that we have a finite-dimensional vector space V and a linear map α from V to itself.

The determinant criterion

Remark ??: Suppose again that we have a finite-dimensional vector space V and a linear map α from V to itself. One can show that the following are equivalent:

(a) α is injective

(b) α is surjective

(c) α is an isomorphism

(d) $\det(\alpha) \neq 0$.

The determinant criterion

Remark ??: Suppose again that we have a finite-dimensional vector space V and a linear map α from V to itself. One can show that the following are equivalent:

(a) α is injective

(b) α is surjective

(c) α is an isomorphism

(d) $\det(\alpha) \neq 0$.

(It is important here that α goes from V to itself, not to some other space.)

The determinant criterion

Remark ??: Suppose again that we have a finite-dimensional vector space V and a linear map α from V to itself. One can show that the following are equivalent:

(a) α is injective

(b) α is surjective

(c) α is an isomorphism

(d) $\det(\alpha) \neq 0$.

(It is important here that α goes from V to itself, not to some other space.)

We shall not give proofs, however.

Height of linear relations

Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ be a list of elements in V.

Height of linear relations

Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements in V. We put $V_i = \mathrm{span}(v_1, \dots, v_i)$ (with the convention that $V_0 = 0$).

Height of linear relations

Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ be a list of elements in V. We put $V_i = \text{span}(v_1, \ldots, v_i)$ (with the convention that $V_0 = 0$).

There may or may not be any nontrivial linear relations for $\mathcal{V}$.

Height of linear relations

Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements in V. We put $V_i = \text{span}(v_1, \dots, v_i)$ (with the convention that $V_0 = 0$).

There may or may not be any nontrivial linear relations for $\mathcal{V}$.

If there is a nontrivial relation $\boldsymbol{\lambda}$, so that $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ and $\lambda_k \neq 0$ for some k, then we define the *height* of $\boldsymbol{\lambda}$ to be the largest i such that $\lambda_i \neq 0$.

Height of linear relations

Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ be a list of elements in V. We put $V_i = \operatorname{span}(v_1, \dots, v_i)$ (with the convention that $V_0 = 0$).

There may or may not be any nontrivial linear relations for $\mathcal{V}$.

If there is a nontrivial relation $\boldsymbol{\lambda}$, so that $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ and $\lambda_k \neq 0$ for some k, then we define the *height* of $\boldsymbol{\lambda}$ to be the largest i such that $\lambda_i \neq 0$.

For example, if $n = 6$ and $5v_1 - 2v_2 - 2v_3 + 3v_4 = 0$ then $[5, -2, -2, 3, 0, 0]^T$ is a nontrivial linear relation of height 4.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Examples

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Examples

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$, so $[1, -2, 1, 0]^T$ is a linear relation

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$, so $[1, -2, 1, 0]^T$ is a linear relation of height 3.

Examples

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$, so $[1, -2, 1, 0]^T$ is a linear relation of height 3.

The equation can be rearranged as $v_3 = -v_1 + 2v_2$

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$, so $[1, -2, 1, 0]^T$ is a linear relation of height 3.

The equation can be rearranged as $v_3 = -v_1 + 2v_2$, showing that $v_3 \in \text{span}(v_1, v_2) = V_2$.

Examples

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$, so $[1, -2, 1, 0]^T$ is a linear relation of height 3.

The equation can be rearranged as $v_3 = -v_1 + 2v_2$, showing that $v_3 \in \text{span}(v_1, v_2) = V_2$. One can check that

$$V_2 = V_3 = \{[x, y, z]^T \mid x + z = 2y\}.$$

The following are equivalent:
(a) $\mathcal{V}$ has a linear relation of height i; (b) $v_i \in V_{i-1}$; (c) $V_i = V_{i-1}$.

Example ??: Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

Then $v_1 - 2v_2 + v_3 = 0$, so $[1, -2, 1, 0]^T$ is a linear relation of height 3.

The equation can be rearranged as $v_3 = -v_1 + 2v_2$, showing that $v_3 \in \operatorname{span}(v_1, v_2) = V_2$. One can check that

$$V_2 = V_3 = \{[x, y, z]^T \mid x + z = 2y\}.$$

Thus, in this example, with $i = 3$, we see that (a), (b) and (c) all hold.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b):

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]^T$ be a nontrivial linear relation of height i

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$.

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$. The fact that the height is i means that $\lambda_i \neq 0$ but $\lambda_{i+1} = \lambda_{i+2} = \cdots = 0$.

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$. The fact that the height is i means that $\lambda_i \neq 0$ but $\lambda_{i+1} = \lambda_{i+2} = \cdots = 0$. We can thus rearrange the linear relation as

$$\lambda_i v_i = -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_n v_n$$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$. The fact that the height is i means that $\lambda_i \neq 0$ but $\lambda_{i+1} = \lambda_{i+2} = \cdots = 0$. We can thus rearrange the linear relation as

$$\begin{aligned} \lambda_i v_i &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_n v_n \\ &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - 0.v_{i+1} - \cdots - 0.v_n \end{aligned}$$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$. The fact that the height is i means that $\lambda_i \neq 0$ but $\lambda_{i+1} = \lambda_{i+2} = \cdots = 0$. We can thus rearrange the linear relation as

$$\begin{aligned}
\lambda_i v_i &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_n v_n \\
&= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - 0.v_{i+1} - \cdots - 0.v_n \\
&= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1}
\end{aligned}$$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. The fact that the height is i means that $\lambda_i \neq 0$ but $\lambda_{i+1} = \lambda_{i+2} = \cdots = 0$. We can thus rearrange the linear relation as

$$\begin{aligned}
\lambda_i v_i &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_n v_n \\
&= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - 0.v_{i+1} - \cdots - 0.v_n \\
&= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} \\
v_i &= -\lambda_1 \lambda_i^{-1} v_1 - \cdots - \lambda_{i-1} \lambda_i^{-1} v_{i-1} \in V_{i-1}.
\end{aligned}$$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (a)⇒(b): Let $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^T$ be a nontrivial linear relation of height i, so $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. The fact that the height is i means that $\lambda_i \neq 0$ but $\lambda_{i+1} = \lambda_{i+2} = \cdots = 0$. We can thus rearrange the linear relation as

$$\begin{aligned} \lambda_i v_i &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_n v_n \\ &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - 0.v_{i+1} - \cdots - 0.v_n \\ &= -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} \\ v_i &= -\lambda_1 \lambda_i^{-1} v_1 - \cdots - \lambda_{i-1} \lambda_i^{-1} v_{i-1} \in V_{i-1}. \end{aligned}$$

so $v_i \in V_{i-1}$.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(a):

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(a): Suppose that $v_i \in V_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(a): Suppose that $v_i \in V_{i-1} = \text{span}(v_1, \ldots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \ldots, \mu_{i-1}$.

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(a): Suppose that $v_i \in V_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We can rewrite this as a nontrivial linear relation

$$\mu_1 v_1 + \dots + \mu_{i-1} v_{i-1} + (-1).v_i + 0.v_{i+1} + \dots + 0.v_n = 0,$$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(a): Suppose that $v_i \in V_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We can rewrite this as a nontrivial linear relation

$$\mu_1 v_1 + \dots + \mu_{i-1} v_{i-1} + (-1).v_i + 0.v_{i+1} + \dots + 0.v_n = 0,$$

which clearly has height i.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Proof that (b)$\Rightarrow$(c):

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)$\Rightarrow$(c): Suppose again that $v_i \in V_{i-1} = \text{span}(v_1, \ldots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \ldots, \mu_{i-1}$.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \operatorname{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$.

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$. Consider an element $w \in V_i$; we must show that $w \in V_{i-1}$.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$. Consider an element $w \in V_i$; we must show that $w \in V_{i-1}$. As $w \in V_i$ we have $w = \lambda_1 v_1 + \dots + \lambda_i v_i$ for some scalars $\lambda_1, \dots, \lambda_i$.

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$. Consider an element $w \in V_i$; we must show that $w \in V_{i-1}$. As $w \in V_i$ we have $w = \lambda_1 v_1 + \dots + \lambda_i v_i$ for some scalars $\lambda_1, \dots, \lambda_i$. This can be rewritten as

$$w = \lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1} + \lambda_i (\mu_1 v_1 + \dots + \mu_{i-1} v_{i-1})$$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$. Consider an element $w \in V_i$; we must show that $w \in V_{i-1}$. As $w \in V_i$ we have $w = \lambda_1 v_1 + \cdots + \lambda_i v_i$ for some scalars $\lambda_1, \dots, \lambda_i$. This can be rewritten as

$$\begin{aligned} w =& \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_i(\mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}) \\ =& (\lambda_1 + \lambda_i \mu_1) v_1 + (\lambda_2 + \lambda_i \mu_2) v_1 + \cdots + (\lambda_{i-1} + \lambda_i \mu_{i-1}) v_{i-1}. \end{aligned}$$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \dots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$. Consider an element $w \in V_i$; we must show that $w \in V_{i-1}$. As $w \in V_i$ we have $w = \lambda_1 v_1 + \cdots + \lambda_i v_i$ for some scalars $\lambda_1, \dots, \lambda_i$. This can be rewritten as

$$\begin{aligned} w =& \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_i(\mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}) \\ =& (\lambda_1 + \lambda_i \mu_1) v_1 + (\lambda_2 + \lambda_i \mu_2) v_1 + \cdots + (\lambda_{i-1} + \lambda_i \mu_{i-1}) v_{i-1}. \end{aligned}$$

This is a linear combination of $v_1, \dots, v_{i-1}$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (b)⇒(c): Suppose again that $v_i \in V_{i-1} = \text{span}(v_1, \ldots, v_{i-1})$, so $v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}$ for some scalars $\mu_1, \ldots, \mu_{i-1}$. We need to show that $V_i = V_{i-1}$, but it is clear that $V_{i-1} \leq V_i$, so it will be enough to show that $V_i \leq V_{i-1}$. Consider an element $w \in V_i$; we must show that $w \in V_{i-1}$. As $w \in V_i$ we have $w = \lambda_1 v_1 + \cdots + \lambda_i v_i$ for some scalars $\lambda_1, \ldots, \lambda_i$. This can be rewritten as

$$\begin{aligned} w =&\lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_i(\mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1}) \\ &= (\lambda_1 + \lambda_i \mu_1) v_1 + (\lambda_2 + \lambda_i \mu_2) v_1 + \cdots + (\lambda_{i-1} + \lambda_i \mu_{i-1}) v_{i-1}. \end{aligned}$$

This is a linear combination of $v_1, \ldots, v_{i-1}$, showing that $w \in V_{i-1}$, as claimed.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Proof that (c)⇒(b):

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\operatorname{span}(v_1, \ldots, v_i) = V_i$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\mathrm{span}(v_1, \ldots, v_i) = V_i$, but $V_i = V_{i-1}$

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\mathrm{span}(v_1, \dots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\operatorname{span}(v_1, \ldots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.
This completes the proof of the Proposition. □

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\operatorname{span}(v_1, \ldots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.
This completes the proof of the Proposition. □

Corollary ??: If for all i we have $v_i \notin V_{i-1}$

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\operatorname{span}(v_1, \ldots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.
This completes the proof of the Proposition. □

Corollary ??: If for all i we have $v_i \notin V_{i-1}$, then there cannot be a linear relation of any height

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\mathrm{span}(v_1, \dots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.
This completes the proof of the Proposition. □

Corollary ??: If for all i we have $v_i \notin V_{i-1}$, then there cannot be a linear relation of any height, so $\mathcal{V}$ must be linearly independent. □

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\text{span}(v_1, \ldots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.
This completes the proof of the Proposition. □

Corollary ??: If for all i we have $v_i \not\in V_{i-1}$, then there cannot be a linear relation of any height, so $\mathcal{V}$ must be linearly independent. □

Corollary ??: The following are equivalent:

(a) The list $\mathcal{V}$ has no nontrivial linear relation of height i

(b) $v_i \not\in V_{i-1}$ (c) $V_i \neq V_{i-1}$.

Jumps

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Proof that (c)⇒(b): Suppose that $V_i = V_{i-1}$. It is clear that the element v_i lies in $\text{span}(v_1, \ldots, v_i) = V_i$, but $V_i = V_{i-1}$, so $v_i \in V_{i-1}$.
This completes the proof of the Proposition. □

Corollary ??: If for all i we have $v_i \notin V_{i-1}$, then there cannot be a linear relation of any height, so $\mathcal{V}$ must be linearly independent. □

Corollary ??: The following are equivalent:

(a) The list $\mathcal{V}$ has no nontrivial linear relation of height i

(b) $v_i \notin V_{i-1}$ (c) $V_i \neq V_{i-1}$.

If these three things are true, we say that i is a *jump*.

Every spanning set contains a basis

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \ldots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \ldots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof:

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \ldots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$. We first claim that $\mathcal{V}'$ is linearly independent.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.
We first claim that $\mathcal{V}'$ is linearly independent.
If not, then there is a nontrivial relation.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.
We first claim that $\mathcal{V}'$ is linearly independent.
If not, then there is a nontrivial relation.
If we write only the nontrivial terms, then the relation takes the form

$$\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_r} v_{i_r} = 0$$

with $i_k \in I'$ for all k, and $\lambda_{i_k} \neq 0$ for all k, and $i_1 < \cdots < i_r$.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.
We first claim that $\mathcal{V}'$ is linearly independent.
If not, then there is a nontrivial relation.
If we write only the nontrivial terms, then the relation takes the form
$$\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_r} v_{i_r} = 0$$
with $i_k \in I'$ for all k, and $\lambda_{i_k} \neq 0$ for all k, and $i_1 < \cdots < i_r$.
This can be regarded as a nontrivial linear relation for $\mathcal{V}$, of height i_r.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.
We first claim that $\mathcal{V}'$ is linearly independent.
If not, then there is a nontrivial relation.
If we write only the nontrivial terms, then the relation takes the form

$$\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_r} v_{i_r} = 0$$

with $i_k \in I'$ for all k, and $\lambda_{i_k} \neq 0$ for all k, and $i_1 < \cdots < i_r$.
This can be regarded as a nontrivial linear relation for $\mathcal{V}$, of height i_r.
Proposition **??** therefore tells us that $v_{i_r} \in V_{i_r - 1}$

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.
We first claim that $\mathcal{V}'$ is linearly independent.
If not, then there is a nontrivial relation.
If we write only the nontrivial terms, then the relation takes the form

$$\lambda_{i_1} v_{i_1} + \dots + \lambda_{i_r} v_{i_r} = 0$$

with $i_k \in I'$ for all k, and $\lambda_{i_k} \neq 0$ for all k, and $i_1 < \dots < i_r$.
This can be regarded as a nontrivial linear relation for $\mathcal{V}$, of height i_r.
Proposition **??** therefore tells us that $v_{i_r} \in V_{i_r - 1}$, which is impossible, as i_r is a jump.

Every spanning set contains a basis

Lemma ??: Let $\mathcal{V} = v_1, \dots, v_n$ be a list that spans a vector space V. Then some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proof: Put $I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$, and $\mathcal{V}' = \{v_i \mid i \in I'\}$.
We first claim that $\mathcal{V}'$ is linearly independent.
If not, then there is a nontrivial relation.
If we write only the nontrivial terms, then the relation takes the form

$$\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_r} v_{i_r} = 0$$

with $i_k \in I'$ for all k, and $\lambda_{i_k} \neq 0$ for all k, and $i_1 < \cdots < i_r$.
This can be regarded as a nontrivial linear relation for $\mathcal{V}$, of height i_r.
Proposition **??** therefore tells us that $v_{i_r} \in V_{i_r - 1}$, which is impossible, as i_r is a jump.
This contradiction shows that $\mathcal{V}'$ must be linearly independent, after all.

Every spanning set contains a basis

$V = \text{span}(\mathcal{V}) = \text{span}(v_1, \ldots, v_n)$; $\qquad V_i = \text{span}(v_1, \ldots, v_i)$;

$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \ldots, v_n);$ $\qquad V_i = \mathrm{span}(v_1, \ldots, v_i);$

$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\};$ $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}.$

Now put $V' = \mathrm{span}(\mathcal{V}')$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\qquad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\mathrm{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$.

Every spanning set contains a basis

$$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n); \qquad V_i = \mathrm{span}(v_1, \dots, v_i);$$
$$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}; \qquad \mathcal{V}' = \{v_i \mid i \in I'\}.$$

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$.

Every spanning set contains a basis

$$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n); \qquad V_i = \mathrm{span}(v_1, \dots, v_i);$$
$$I' = \{\mathrm{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}; \qquad \mathcal{V}' = \{v_i \mid i \in I'\}.$$

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \dots, v_i)$;

$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n);$ $\quad V_i = \mathrm{span}(v_1, \dots, v_i);$
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\};$ $\quad \mathcal{V}' = \{v_i \mid i \in I'\}.$

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\qquad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$.

Every spanning set contains a basis

$V = \text{span}(\mathcal{V}) = \text{span}(v_1, \dots, v_n)$; $\qquad V_i = \text{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \text{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \not\in V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\qquad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \ldots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \ldots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$.

Every spanning set contains a basis

$V = \text{span}(\mathcal{V}) = \text{span}(v_1, \dots, v_n)$; $\qquad V_i = \text{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \not\in V_{i-1}\}$; $\qquad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \text{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$. In particular, we have $V_n \leq V'$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$. In particular, we have $V_n \leq V'$. However, V_n is just $\mathrm{span}(\mathcal{V})$

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$. In particular, we have $V_n \leq V'$. However, V_n is just $\mathrm{span}(\mathcal{V})$, and we assumed that $\mathcal{V}$ spans V

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$. In particular, we have $V_n \leq V'$. However, V_n is just $\mathrm{span}(\mathcal{V})$, and we assumed that $\mathcal{V}$ spans V, so $V_n = V$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \le n \mid v_i \notin V_{i-1}\}$; $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \le V'$ for all $i \le n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \le V'$. Suppose that $V_{i-1} \le V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \le V'$ and $\mathbb{R}v_i \le V'$, we conclude that $V_i \le V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \le V'$, so $V_i \le V'$.

Either way we have $V_i \le V'$, which proves the induction step. We therefore have $V_i \le V'$ for all $i \le n$. In particular, we have $V_n \le V'$. However, V_n is just $\mathrm{span}(\mathcal{V})$, and we assumed that $\mathcal{V}$ spans V, so $V_n = V$. This proves that $V \le V'$, and it is clear that $V' \le V$, so $V = V'$.

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \ldots, v_n)$; $\quad V_i = \mathrm{span}(v_1, \ldots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\quad \mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$. In particular, we have $V_n \leq V'$. However, V_n is just $\mathrm{span}(\mathcal{V})$, and we assumed that $\mathcal{V}$ spans V, so $V_n = V$. This proves that $V \leq V'$, and it is clear that $V' \leq V$, so $V = V'$. This means that $\mathcal{V}'$ is a spanning list as well as being linearly independent

Every spanning set contains a basis

$V = \mathrm{span}(\mathcal{V}) = \mathrm{span}(v_1, \dots, v_n)$; $V_i = \mathrm{span}(v_1, \dots, v_i)$;
$I' = \{\text{jumps}\} = \{i \leq n \mid v_i \notin V_{i-1}\}$; $\mathcal{V}' = \{v_i \mid i \in I'\}$.

Now put $V' = \mathrm{span}(\mathcal{V}')$. We will show by induction that $V_i \leq V'$ for all $i \leq n$. For the initial step, we note that $V_0 = 0$ so certainly $V_0 \leq V'$. Suppose that $V_{i-1} \leq V'$. There are two cases to consider:

(a) Suppose that i is a jump, so $i \in I'$. Then (by the definition of $\mathcal{V}'$) we have $v_i \in \mathcal{V}'$ and so $v_i \in V'$. As $V_i = V_{i-1} + \mathbb{R}v_i$ and $V_{i-1} \leq V'$ and $\mathbb{R}v_i \leq V'$, we conclude that $V_i \leq V'$.

(b) Suppose that i is not a jump, so $v_i \in V_{i-1}$ and so $V_i = V_{i-1}$. By the induction hypothesis we have $V_{i-1} \leq V'$, so $V_i \leq V'$.

Either way we have $V_i \leq V'$, which proves the induction step. We therefore have $V_i \leq V'$ for all $i \leq n$. In particular, we have $V_n \leq V'$. However, V_n is just $\mathrm{span}(\mathcal{V})$, and we assumed that $\mathcal{V}$ spans V, so $V_n = V$. This proves that $V \leq V'$, and it is clear that $V' \leq V$, so $V = V'$. This means that $\mathcal{V}'$ is a spanning list as well as being linearly independent, so it is a basis for V. $\square$

Existence of bases

Existence of bases

Corollary ??: Every finite-dimensional vector space has a basis.

Existence of bases

Corollary ??: Every finite-dimensional vector space has a basis.
Proof: By Definition **??**, we can find a finite list $\mathcal{V}$ that spans V.

Existence of bases

Corollary ??: Every finite-dimensional vector space has a basis.
Proof: By Definition **??**, we can find a finite list $\mathcal{V}$ that spans V. By Lemma **??**, some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis.

Steinitz's lemma

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent.

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \mathrm{span}(v_1, \dots, v_i)$, so $V_n = \mathrm{span}(\mathcal{V}) = V$.

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \dots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i.

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \mathrm{span}(v_1, \dots, v_i)$, so $V_n = \mathrm{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \dots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$.

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \mathrm{span}(v_1, \dots, v_i)$, so $V_n = \mathrm{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$. This means that the only linearly independent list in V_0 is the empy list, which has length 0, as required.

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \ldots, v_n$ and $\mathcal{W} = w_1, \ldots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \ldots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$. This means that the only linearly independent list in V_0 is the empy list, which has length 0, as required.

Now suppose (for the induction step) that every linearly independent list in V_{i-1} has length at most $i - 1$.

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \dots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$. This means that the only linearly independent list in V_0 is the empy list, which has length 0, as required.

Now suppose (for the induction step) that every linearly independent list in V_{i-1} has length at most $i - 1$. Suppose we have a linearly independent list $x_1, \dots, x_p$ in V_i

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \dots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$. This means that the only linearly independent list in V_0 is the empy list, which has length 0, as required.

Now suppose (for the induction step) that every linearly independent list in V_{i-1} has length at most $i - 1$. Suppose we have a linearly independent list $x_1, \dots, x_p$ in V_i; we must show that $p \leq i$.

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \dots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$. This means that the only linearly independent list in V_0 is the empy list, which has length 0, as required.

Now suppose (for the induction step) that every linearly independent list in V_{i-1} has length at most $i - 1$. Suppose we have a linearly independent list $x_1, \dots, x_p$ in V_i; we must show that $p \leq i$. The elements x_j lie in $V_i = \text{span}(v_1, \dots, v_i)$.

Steinitz's lemma

Lemma ??: Let V be a vector space, and let $\mathcal{V} = v_1, \dots, v_n$ and $\mathcal{W} = w_1, \dots, w_m$ be finite lists of elements of $\mathcal{V}$ such that $\mathcal{V}$ spans V and $\mathcal{W}$ is linearly independent. Then $n \geq m$.
(Any spanning list is at least as long as any linearly independent list.)

Proof: As before, we put $V_i = \text{span}(v_1, \dots, v_i)$, so $V_n = \text{span}(\mathcal{V}) = V$. We will show by induction that any linearly independent list in V_i has length at most i. In particular, this will show that any linearly independent list in $V = V_n$ has length at most n, as claimed.

For the initial step, note that $V_0 = 0$. This means that the only linearly independent list in V_0 is the empy list, which has length 0, as required.

Now suppose (for the induction step) that every linearly independent list in V_{i-1} has length at most $i - 1$. Suppose we have a linearly independent list $x_1, \dots, x_p$ in V_i; we must show that $p \leq i$. The elements x_j lie in $V_i = \text{span}(v_1, \dots, v_i)$. We can thus find scalars a_{jk} such that

$$x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{j,i-1}v_{i-1} + a_{ji}v_i.$$

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Case (a): Suppose that for each j the last coefficient a_{ji} is zero.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Case (a): Suppose that for each j the last coefficient a_{ji} is zero. This means that

$$x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{j,i-1}v_{i-1}$$

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Case (a): Suppose that for each j the last coefficient a_{ji} is zero. This means that

$$x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{j,i-1}v_{i-1},$$

so $x_j \in \operatorname{span}(v_1, \ldots, v_{i-1}) = V_{i-1}$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \dots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Case (a): Suppose that for each j the last coefficient a_{ji} is zero. This means that

$$x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{j,i-1}v_{i-1},$$

so $x_j \in \mathrm{span}(v_1, \dots, v_{i-1}) = V_{i-1}$. This means that $x_1, \dots, x_p$ is a linearly independent list in V_{i-1}

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i.$

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Case (a): Suppose that for each j the last coefficient a_{ji} is zero. This means that

$$x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{j,i-1}v_{i-1},$$

so $x_j \in \operatorname{span}(v_1, \ldots, v_{i-1}) = V_{i-1}$. This means that $x_1, \ldots, x_p$ is a linearly independent list in V_{i-1}, so the induction hypothesis tells us that $p \leq i-1$

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

We need to consider two cases:

(a) For each j the last coefficient a_{ji} is zero.

(b) For some j the last coefficient a_{ji} is nonzero.

Case (a): Suppose that for each j the last coefficient a_{ji} is zero. This means that

$$x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{j,i-1}v_{i-1},$$

so $x_j \in \mathrm{span}(v_1, \ldots, v_{i-1}) = V_{i-1}$. This means that $x_1, \ldots, x_p$ is a linearly independent list in V_{i-1}, so the induction hypothesis tells us that $p \leq i-1$, so certainly $p \leq i$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$

$x_1, \dots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i.$

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \dots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i.$

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \dots, x_p$ independent in V_i; $\quad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \dots, y_{p-1}$ is a linearly independent list in V_{i-1}.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p-1 \leq i-1$, so $p \leq i$ as required.

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p - 1 \leq i - 1$, so $p \leq i$ as required. First, we have

$$y_k = x_k - a_{ki}a_{pi}^{-1}x_p$$

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p-1 \leq i-1$, so $p \leq i$ as required. First, we have

$$y_k = x_k - a_{ki}a_{pi}^{-1}x_p = a_{k1}v_1 + \cdots + a_{ki}v_i - a_{ki}a_{pi}^{-1}(a_{p1}v_1 + \cdots + a_{pi}v_i)$$

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p-1 \leq i-1$, so $p \leq i$ as required. First, we have

$$\begin{aligned} y_k =& x_k - a_{ki}a_{pi}^{-1}x_p = a_{k1}v_1 + \cdots + a_{ki}v_i - a_{ki}a_{pi}^{-1}(a_{p1}v_1 + \cdots + a_{pi}v_i) \\ &= (a_{k1} - a_{ki}a_{pi}^{-1}a_{p1})v_1 + (a_{k2} - a_{ki}a_{pi}^{-1}a_{p2})v_2 + \cdots + (a_{ki} - a_{ki}a_{pi}{}^{-1}a_{pi})v_i. \end{aligned}$$

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p-1 \leq i-1$, so $p \leq i$ as required. First, we have

$$
\begin{aligned}
y_k =& x_k - a_{ki}a_{pi}^{-1}x_p = a_{k1}v_1 + \cdots + a_{ki}v_i - a_{ki}a_{pi}^{-1}(a_{p1}v_1 + \cdots + a_{pi}v_i) \\
&= (a_{k1} - a_{ki}a_{pi}^{-1}a_{p1})v_1 + (a_{k2} - a_{ki}a_{pi}^{-1}a_{p2})v_2 + \cdots + (a_{ki} - a_{ki}a_{pi}{}^{-1}a_{pi})v_i.
\end{aligned}
$$

In the last term, the coefficient $a_{ki} - a_{ki}a_{pi}^{-1}a_{pi}$ is zero

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p-1 \leq i-1$, so $p \leq i$ as required. First, we have

$$\begin{aligned} y_k =& x_k - a_{ki}a_{pi}^{-1}x_p = a_{k1}v_1 + \cdots + a_{ki}v_i - a_{ki}a_{pi}^{-1}(a_{p1}v_1 + \cdots + a_{pi}v_i) \\ &= (a_{k1} - a_{ki}a_{pi}^{-1}a_{p1})v_1 + (a_{k2} - a_{ki}a_{pi}^{-1}a_{p2})v_2 + \cdots + (a_{ki} - a_{ki}a_{pi}{}^{-1}a_{pi})v_i. \end{aligned}$$

In the last term, the coefficient $a_{ki} - a_{ki}a_{pi}^{-1}a_{pi}$ is zero, so y_k is actually a linear combination of $v_1, \ldots, v_{i-1}$

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$.

Case (b): Suppose that for some x_j we have $a_{ji} \neq 0$.

It is harmless to reorder the x's, so for notational convenience we move this x_j to the end of the list, which means that $a_{pi} \neq 0$.

Now put $\alpha_k = a_{ki}a_{pi}^{-1}$ and $y_k = x_k - \alpha_k x_p$.

We will show that $y_1, \ldots, y_{p-1}$ is a linearly independent list in V_{i-1}. Assuming this, the induction hypothesis gives $p-1 \leq i-1$, so $p \leq i$ as required. First, we have

$$\begin{aligned} y_k =& x_k - a_{ki}a_{pi}^{-1}x_p = a_{k1}v_1 + \cdots + a_{ki}v_i - a_{ki}a_{pi}^{-1}(a_{p1}v_1 + \cdots + a_{pi}v_i) \\ &= (a_{k1} - a_{ki}a_{pi}^{-1}a_{p1})v_1 + (a_{k2} - a_{ki}a_{pi}^{-1}a_{p2})v_2 + \cdots + ({\color{red}a_{ki}} - {\color{red}a_{ki}}{\color{blue}a_{pi}}^{-1}{\color{blue}a_{pi}})v_i. \end{aligned}$$

In the last term, the coefficient $a_{ki} - a_{ki}a_{pi}^{-1}a_{pi}$ is zero, so y_k is actually a linear combination of $v_1, \ldots, v_{i-1}$, so $y_k \in V_{i-1}$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\quad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \dots, x_p$ independent in V_i; $x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \dots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$.

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \ldots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\quad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \ldots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial. We conclude that the list $y_1, \ldots, y_{p-1}$ is an independent list in V_{i-1}.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \dots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \dots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial. We conclude that the list $y_1, \dots, y_{p-1}$ is an independent list in V_{i-1}. As explained before, the induction hypothesis now tells us that $p - 1 \leq i - 1$, so $p \leq i$.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \ldots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial. We conclude that the list $y_1, \ldots, y_{p-1}$ is an independent list in V_{i-1}. As explained before, the induction hypothesis now tells us that $p - 1 \leq i - 1$, so $p \leq i$.

This completes the induction step.

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \ldots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial. We conclude that the list $y_1, \ldots, y_{p-1}$ is an independent list in V_{i-1}. As explained before, the induction hypothesis now tells us that $p-1 \leq i-1$, so $p \leq i$.

This completes the induction step. So any independent list in V_i has length at most i.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \dots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \dots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial. We conclude that the list $y_1, \dots, y_{p-1}$ is an independent list in V_{i-1}. As explained before, the induction hypothesis now tells us that $p-1 \leq i-1$, so $p \leq i$.

This completes the induction step. So any independent list in V_i has length at most i. In particular, any independent list in $V = V_n$ has length at most n.

Steinitz's lemma

Every independent list in V_{i-1} has length at most $i-1$
$x_1, \ldots, x_p$ independent in V_i; $\qquad x_j = a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{ji}v_i$;
$y_k = x_k - \alpha_k x_p = x_k - a_{ki}a_{pi}^{-1}x_p \in V_{i-1}$

Next, suppose we have a linear relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$. Put

$$\lambda_p = -\lambda_1\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_{p-1}\alpha_{p-1}.$$

By putting $y_k = x_k - \alpha_k x_p$ in the relation $\lambda_1 y_1 + \cdots + \lambda_{p-1}y_{p-1} = 0$ and expanding it out, we get $\lambda_1 x_1 + \ldots + \lambda_{p-1}x_{p-1} + \lambda_p x_p = 0$. As $x_1, \ldots, x_p$ is independent, this means that we must have $\lambda_1 = \cdots = \lambda_{p-1} = \lambda_p = 0$. It follows that our original relation among the y's was trivial. We conclude that the list $y_1, \ldots, y_{p-1}$ is an independent list in V_{i-1}. As explained before, the induction hypothesis now tells us that $p - 1 \leq i - 1$, so $p \leq i$.

This completes the induction step. So any independent list in V_i has length at most i. In particular, any independent list in $V = V_n$ has length at most n. This completes the proof of Steiniz's lemma.

Invariance of dimension

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n.

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V.

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \ldots, v_n$.

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \ldots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V.

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \dots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements.

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \ldots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \dots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \ldots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements. Now let $\mathcal{V}'$ be another basis for V.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \dots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements. Now let $\mathcal{V}'$ be another basis for V. Then $\mathcal{V}'$ has at least n elements (because it spans)

Invariance of dimension

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \ldots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements. Now let $\mathcal{V}'$ be another basis for V. Then $\mathcal{V}'$ has at least n elements (because it spans) and at most n elements (because it is independent)

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \dots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements. Now let $\mathcal{V}'$ be another basis for V. Then $\mathcal{V}'$ has at least n elements (because it spans) and at most n elements (because it is independent) so it must have exactly n elements.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \ldots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements. Now let $\mathcal{V}'$ be another basis for V. Then $\mathcal{V}'$ has at least n elements (because it spans) and at most n elements (because it is independent) so it must have exactly n elements.

Corollary ??: If V is a finite-dimensional vector space over $\mathbb{R}$ with dimension n, then V is isomorphic to $\mathbb{R}^n$.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Proof: We already saw in Corollary **??** that V has a basis, say $\mathcal{V} = v_1, \dots, v_n$. Let $\mathcal{X}$ be a linearly independent list in V. As $\mathcal{V}$ is a spanning list and $\mathcal{X}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{V}$ is at least as long as $\mathcal{X}$, so $\mathcal{X}$ has at most n elements. Now let $\mathcal{Y}$ be a spanning list for V. As $\mathcal{Y}$ spans and $\mathcal{V}$ is linearly independent, Steinitz's Lemma tells us that $\mathcal{Y}$ is at least as long as $\mathcal{V}$, so $\mathcal{Y}$ has at least n elements. Now let $\mathcal{V}'$ be another basis for V. Then $\mathcal{V}'$ has at least n elements (because it spans) and at most n elements (because it is independent) so it must have exactly n elements.

Corollary ??: If V is a finite-dimensional vector space over $\mathbb{R}$ with dimension n, then V is isomorphic to $\mathbb{R}^n$.
Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be any basis; then $\mu_{\mathcal{V}} \colon \mathbb{R}^n \to V$ is an isomorphism.

Recollections

Recollections

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$

(c) $V_i = V_{i-1}$.

Recollections

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Corollary ??: If for all i we have $v_i \notin V_{i-1}$, then there cannot be a linear relation of any height, so $\mathcal{V}$ must be linearly independent.

Recollections

Proposition ??: The following are equivalent (so if any one of them is true, then so are the other two):

(a) The list $\mathcal{V}$ has a nontrivial linear relation of height i

(b) $v_i \in V_{i-1}$ (c) $V_i = V_{i-1}$.

Corollary ??: If for all i we have $v_i \notin V_{i-1}$, then there cannot be a linear relation of any height, so $\mathcal{V}$ must be linearly independent.

Corollary ??: Let V be a finite-dimensional vector space. Then V has a finite basis, and any two bases have the same number of elements, say n. This number is called the *dimension* of V. Moreover, any spanning list for V has at least n elements, and any linearly independent list has at most n elements.

Subspaces are finite-dimensional

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\operatorname{span}(w_1)$.
If $\operatorname{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.
If $\mathrm{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\mathrm{span}(w_1, w_2)$.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.
If $\mathrm{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\mathrm{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \not\in \mathrm{span}(w_1, \dots, w_{i-1})$

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\text{span}(w_1)$.
If $\text{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\text{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \notin \text{span}(w_1, \dots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**).

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.
If $\mathrm{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\mathrm{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \not\in \mathrm{span}(w_1, \dots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**). However, V has a spanning set of length n

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\text{span}(w_1)$.
If $\text{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\text{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \notin \text{span}(w_1, \ldots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**). However, V has a spanning set of length n, so Steinitz's Lemma tells us that we cannot have a linearly independent list of length greater than n

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.
If $\mathrm{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\mathrm{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \notin \mathrm{span}(w_1, \ldots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**). However, V has a spanning set of length n, so Steinitz's Lemma tells us that we cannot have a linearly independent list of length greater than n, so our list of w's must stop before we get to w_{n+1}.

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\text{span}(w_1)$.
If $\text{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\text{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \not\in \text{span}(w_1, \dots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**). However, V has a spanning set of length n, so Steinitz's Lemma tells us that we cannot have a linearly independent list of length greater than n, so our list of w's must stop before we get to w_{n+1}. This means that for some $p \leq n$ we have $W = \text{span}(w_1, \dots, w_p)$

Subspaces are finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \dots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.
If $\mathrm{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\mathrm{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \notin \mathrm{span}(w_1, \dots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**). However, V has a spanning set of length n, so Steinitz's Lemma tells us that we cannot have a linearly independent list of length greater than n, so our list of w's must stop before we get to w_{n+1}. This means that for some $p \leq n$ we have $W = \mathrm{span}(w_1, \dots, w_p)$, so W is finite-dimensional

Proposition ??: Let V be a finite-dimensional vector space, and let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$.

Proof: Put $n = \dim(V)$. We define a list $\mathcal{W} = w_1, w_2, \ldots$ as follows.
If $W = 0$ then we take $\mathcal{W}$ to be the empty list.
Otherwise, we let w_1 be any nonzero vector in W.
If w_1 spans W we take $\mathcal{W} = w_1$.
Otherwise, we can choose an element $w_2 \in W$ that is not in $\mathrm{span}(w_1)$.
If $\mathrm{span}(w_1, w_2) = W$ then we stop and take $\mathcal{W} = w_1, w_2$.
Otherwise, we can choose an element $w_3 \in W$ that is not in $\mathrm{span}(w_1, w_2)$.

We continue in this way, so we always have $w_i \notin \mathrm{span}(w_1, \ldots, w_{i-1})$, so the w's are linearly independent (by Corollary **??**). However, V has a spanning set of length n, so Steinitz's Lemma tells us that we cannot have a linearly independent list of length greater than n, so our list of w's must stop before we get to w_{n+1}. This means that for some $p \leq n$ we have $W = \mathrm{span}(w_1, \ldots, w_p)$, so W is finite-dimensional, with $\dim(W) = p \leq n$.

Extending to a basis

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \ldots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \ldots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof:

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \ldots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \ldots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \ldots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \ldots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\mathrm{span}(v_1, \ldots, v_p) = V$ then we take $\mathcal{V}' = v_1, \ldots, v_p$.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \dots, v_p)$.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \ldots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \ldots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \ldots, v_p) = V$ then we take $\mathcal{V}' = v_1, \ldots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \ldots, v_p)$.
If $\operatorname{span}(v_1, \ldots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \ldots, v_{p+1})$.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \dots, v_p)$.
If $\operatorname{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \operatorname{span}(v_1, \dots, v_{p+1})$ and continue in the same way.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\mathrm{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \mathrm{span}(v_1, \dots, v_p)$.
If $\mathrm{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \mathrm{span}(v_1, \dots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \mathrm{span}(v_1, \dots, v_{i-1})$

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\mathrm{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \mathrm{span}(v_1, \dots, v_p)$.
If $\mathrm{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \mathrm{span}(v_1, \dots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \mathrm{span}(v_1, \dots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**).

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \dots, v_p)$.
If $\operatorname{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \operatorname{span}(v_1, \dots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \operatorname{span}(v_1, \dots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**). Any linearly independent list has length at most n (by Corollary **??**)

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \dots, v_p)$.
If $\operatorname{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \operatorname{span}(v_1, \dots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \operatorname{span}(v_1, \dots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**). Any linearly independent list has length at most n (by Corollary **??**) so our process must stop before we get to v_{n+1}.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \dots, v_p)$.
If $\operatorname{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \operatorname{span}(v_1, \dots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \operatorname{span}(v_1, \dots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**). Any linearly independent list has length at most n (by Corollary **??**) so our process must stop before we get to v_{n+1}. This means that $\mathcal{V}' = v_1, \dots, v_m$ with $m \leq n$

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \ldots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \ldots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\operatorname{span}(v_1, \ldots, v_p) = V$ then we take $\mathcal{V}' = v_1, \ldots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \operatorname{span}(v_1, \ldots, v_p)$.
If $\operatorname{span}(v_1, \ldots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \ldots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \operatorname{span}(v_1, \ldots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \operatorname{span}(v_1, \ldots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**). Any linearly independent list has length at most n (by Corollary **??**) so our process must stop before we get to v_{n+1}. This means that $\mathcal{V}' = v_1, \ldots, v_m$ with $m \leq n$, and as the process has stopped, we must have $\operatorname{span}(\mathcal{V}') = V$.

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \dots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \dots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\mathrm{span}(v_1, \dots, v_p) = V$ then we take $\mathcal{V}' = v_1, \dots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \mathrm{span}(v_1, \dots, v_p)$.
If $\mathrm{span}(v_1, \dots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \dots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \mathrm{span}(v_1, \dots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \mathrm{span}(v_1, \dots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**). Any linearly independent list has length at most n (by Corollary **??**) so our process must stop before we get to v_{n+1}. This means that $\mathcal{V}' = v_1, \dots, v_m$ with $m \leq n$, and as the process has stopped, we must have $\mathrm{span}(\mathcal{V}') = V$. As $\mathcal{V}'$ is also linearly independent, we see that it is a basis

Extending to a basis

Proposition ??: Let V be an n-dimensional vector space, and let $\mathcal{V} = v_1, \ldots, v_p$ be a linearly independent list of elements of V. Then $p \leq n$, and $\mathcal{V}$ can be extended to a list $\mathcal{V}' = v_1, \ldots, v_n$ such that $\mathcal{V}'$ is a basis of V.

Proof: Corollary **??** tells us that $p \leq n$. If $\text{span}(v_1, \ldots, v_p) = V$ then we take $\mathcal{V}' = v_1, \ldots, v_p$.
Otherwise, we choose some $v_{p+1} \notin \text{span}(v_1, \ldots, v_p)$.
If $\text{span}(v_1, \ldots, v_{p+1}) = V$ then we stop and take $\mathcal{V}' = (v_1, \ldots, v_{p+1})$.
Otherwise, we choose some $v_{p+2} \notin \text{span}(v_1, \ldots, v_{p+1})$ and continue in the same way.
We always have $v_i \notin \text{span}(v_1, \ldots, v_{i-1})$, so the v's are linearly independent (by Corollary **??**). Any linearly independent list has length at most n (by Corollary **??**) so our process must stop before we get to v_{n+1}. This means that $\mathcal{V}' = v_1, \ldots, v_m$ with $m \leq n$, and as the process has stopped, we must have $\text{span}(\mathcal{V}') = V$. As $\mathcal{V}'$ is also linearly independent, we see that it is a basis, and so $m = n$ (by Corollary **??** again).

Counting

Proposition ??: Let V be an n-dimensional vector space.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

Counting

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \dots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$.

Counting

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$. Thus, $\mathcal{V}$ itself must be a basis.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \dots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$. Thus, $\mathcal{V}$ itself must be a basis.

(b) Let $\mathcal{W} = (w_1, \dots, w_n)$ be a linearly independent list.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$. Thus, $\mathcal{V}$ itself must be a basis.

(b) Let $\mathcal{W} = (w_1, \ldots, w_n)$ be a linearly independent list. Proposition **??** tells us that $\mathcal{W}$ can be extended to a list $\mathcal{W}' \supseteq \mathcal{W}$ such that $\mathcal{W}'$ is a basis.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$. Thus, $\mathcal{V}$ itself must be a basis.

(b) Let $\mathcal{W} = (w_1, \ldots, w_n)$ be a linearly independent list. Proposition **??** tells us that $\mathcal{W}$ can be extended to a list $\mathcal{W}' \supseteq \mathcal{W}$ such that $\mathcal{W}'$ is a basis. In particular, $\mathcal{W}'$ must have length n

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \dots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$. Thus, $\mathcal{V}$ itself must be a basis.

(b) Let $\mathcal{W} = (w_1, \dots, w_n)$ be a linearly independent list. Proposition **??** tells us that $\mathcal{W}$ can be extended to a list $\mathcal{W}' \supseteq \mathcal{W}$ such that $\mathcal{W}'$ is a basis. In particular, $\mathcal{W}'$ must have length n, so it must just be the same as $\mathcal{W}$

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Proof:

(a) Let $\mathcal{V} = (v_1, \ldots, v_n)$ be a spanning list. Lemma **??** tells us that some sublist $\mathcal{V}' \subseteq \mathcal{V}$ is a basis for V. As $\dim(V) = n$, we see that $\mathcal{V}'$ has length n, but $\mathcal{V}$ also has length n, so $\mathcal{V}'$ must be all of $\mathcal{V}$. Thus, $\mathcal{V}$ itself must be a basis.

(b) Let $\mathcal{W} = (w_1, \ldots, w_n)$ be a linearly independent list. Proposition **??** tells us that $\mathcal{W}$ can be extended to a list $\mathcal{W}' \supseteq \mathcal{W}$ such that $\mathcal{W}'$ is a basis. In particular, $\mathcal{W}'$ must have length n, so it must just be the same as $\mathcal{W}$, so $\mathcal{W}$ itself is a basis.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Counting

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Corollary ??: Let V be an finite-dimensional vector space, and let W be a subspace with $\dim(W) = \dim(V)$; then $W = V$.

Counting

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Corollary ??: Let V be an finite-dimensional vector space, and let W be a subspace with $\dim(W) = \dim(V)$; then $W = V$.

Proof: Put $n = \dim(V) = \dim(W)$

Counting

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Corollary ??: Let V be an finite-dimensional vector space, and let W be a subspace with $\dim(W) = \dim(V)$; then $W = V$.

Proof: Put $n = \dim(V) = \dim(W)$, and let $\mathcal{W} = w_1, \ldots, w_n$ be a basis for W.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Corollary ??: Let V be an finite-dimensional vector space, and let W be a subspace with $\dim(W) = \dim(V)$; then $W = V$.

Proof: Put $n = \dim(V) = \dim(W)$, and let $\mathcal{W} = w_1, \dots, w_n$ be a basis for W. Then $\mathcal{W}$ is a linearly independent list in V with n elements

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Corollary ??: Let V be an finite-dimensional vector space, and let W be a subspace with $\dim(W) = \dim(V)$; then $W = V$.

Proof: Put $n = \dim(V) = \dim(W)$, and let $\mathcal{W} = w_1, \ldots, w_n$ be a basis for W. Then $\mathcal{W}$ is a linearly independent list in V with n elements, so part (b) of the Proposition tells us that $\mathcal{W}$ spans V.

Proposition ??: Let V be an n-dimensional vector space.

(a) Any spanning list for V with exactly n elements is linearly independent, and so is a basis.

(b) Any linearly independent list in V with exactly n elements is a spanning list, and so is a basis.

Corollary ??: Let V be an finite-dimensional vector space, and let W be a subspace with $\dim(W) = \dim(V)$; then $W = V$.

Proof: Put $n = \dim(V) = \dim(W)$, and let $\mathcal{W} = w_1, \ldots, w_n$ be a basis for W. Then $\mathcal{W}$ is a linearly independent list in V with n elements, so part (b) of the Proposition tells us that $\mathcal{W}$ spans V. Thus $V = \mathrm{span}(\mathcal{W}) = W$.

Two subspaces

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U.

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \ldots, u_p)$, $(v_1, \ldots, v_q)$ and $(w_1, \ldots, w_r)$ (for some $p, q, r \geq 0$) such that

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \ldots, u_p)$, $(v_1, \ldots, v_q)$ and $(w_1, \ldots, w_r)$ (for some $p, q, r \geq 0$) such that

- $(u_1, \ldots, u_p)$ is a basis for $V \cap W$

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \ldots, u_p)$, $(v_1, \ldots, v_q)$ and $(w_1, \ldots, w_r)$ (for some $p, q, r \geq 0$) such that

- $(u_1, \ldots, u_p)$ is a basis for $V \cap W$
- $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \ldots, u_p)$, $(v_1, \ldots, v_q)$ and $(w_1, \ldots, w_r)$ (for some $p, q, r \geq 0$) such that

- $(u_1, \ldots, u_p)$ is a basis for $V \cap W$
- $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V
- $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \dots, u_p)$, $(v_1, \dots, v_q)$ and $(w_1, \dots, w_r)$ (for some $p, q, r \geq 0$) such that

- $(u_1, \dots, u_p)$ is a basis for $V \cap W$
- $(u_1, \dots, u_p, v_1, \dots, v_q)$ is a basis for V
- $(u_1, \dots, u_p, w_1, \dots, w_r)$ is a basis for W
- $(u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$ is a basis for $V + W$.

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \dots, u_p)$, $(v_1, \dots, v_q)$ and $(w_1, \dots, w_r)$ (for some $p, q, r \geq 0$) such that

- $(u_1, \dots, u_p)$ is a basis for $V \cap W$
- $(u_1, \dots, u_p, v_1, \dots, v_q)$ is a basis for V
- $(u_1, \dots, u_p, w_1, \dots, w_r)$ is a basis for W
- $(u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$ is a basis for $V + W$.

In particular, we have

$$\dim(V \cap W) = p \quad \dim(V) = p+q \quad \dim(W) = p+r \quad \dim(V+W) = p+q+r,$$

Two subspaces

Proposition ??: Let U be a finite-dimensional vector space, and let V and W be subspaces of U. Then one can find lists $(u_1, \dots, u_p)$, $(v_1, \dots, v_q)$ and $(w_1, \dots, w_r)$ (for some $p, q, r \geq 0$) such that

- $(u_1, \dots, u_p)$ is a basis for $V \cap W$
- $(u_1, \dots, u_p, v_1, \dots, v_q)$ is a basis for V
- $(u_1, \dots, u_p, w_1, \dots, w_r)$ is a basis for W
- $(u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$ is a basis for $V + W$.

In particular, we have

$$\dim(V \cap W) = p \quad \dim(V) = p+q \quad \dim(W) = p+r \quad \dim(V+W) = p+q+r,$$

so $\dim(V) + \dim(W) = 2p + q + r = \dim(V \cap W) + \dim(V + W)$.

Proof:

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \ldots, u_p, v_1, \ldots, v_q)$.

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \ldots, u_p, v_1, \ldots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \ldots, u_p, w_1, \ldots, w_r)$.

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \ldots, u_p, v_1, \ldots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \ldots, u_p, w_1, \ldots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \ldots, u_p, v_1, \ldots, v_q, w_1, \ldots, w_r)$$

is a basis for $V + W$.

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \ldots, u_p, v_1, \ldots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \ldots, u_p, w_1, \ldots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \ldots, u_p, v_1, \ldots, v_q, w_1, \ldots, w_r)$$

is a basis for $V + W$. Consider an element

$$x = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \text{span}(\mathcal{X}).$$

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \ldots, u_p, v_1, \ldots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \ldots, u_p, w_1, \ldots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \ldots, u_p, v_1, \ldots, v_q, w_1, \ldots, w_r)$$

is a basis for $V + W$. Consider an element

$$x = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \text{span}(\mathcal{X}).$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $x = y + z$.

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \dots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \dots, u_p, v_1, \dots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \dots, u_p, w_1, \dots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$$

is a basis for $V + W$. Consider an element

$$x = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \operatorname{span}(\mathcal{X}).$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $x = y + z$. We have $u_i, v_j \in V$ and $w_k \in W$

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \dots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \dots, u_p, v_1, \dots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \dots, u_p, w_1, \dots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$$

is a basis for $V + W$. Consider an element

$$x = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \mathrm{span}(\mathcal{X}).$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $x = y + z$. We have $u_i, v_j \in V$ and $w_k \in W$ so $y \in V$ and $z \in W$

Two subspaces

Proof: Choose a basis $\mathcal{U} = (u_1, \ldots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \ldots, u_p, v_1, \ldots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \ldots, u_p, w_1, \ldots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \ldots, u_p, v_1, \ldots, v_q, w_1, \ldots, w_r)$$

is a basis for $V + W$. Consider an element

$$x = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \operatorname{span}(\mathcal{X}).$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $x = y + z$. We have $u_i, v_j \in V$ and $w_k \in W$ so $y \in V$ and $z \in W$ so $x = y + z \in V + W$.

Proof: Choose a basis $\mathcal{U} = (u_1, \dots, u_p)$ for $V \cap W$.
Then $\mathcal{U}$ is a linearly independent list in V, so it can be extended to a basis for V, say $(u_1, \dots, u_p, v_1, \dots, v_q)$.
Similarly $\mathcal{U}$ is a linearly independent list in W, so it can be extended to a basis for W, say $(u_1, \dots, u_p, w_1, \dots, w_r)$.

All that is left is to prove that the list

$$\mathcal{X} = (u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$$

is a basis for $V + W$. Consider an element

$$x = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \operatorname{span}(\mathcal{X}).$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $x = y + z$. We have $u_i, v_j \in V$ and $w_k \in W$ so $y \in V$ and $z \in W$ so $x = y + z \in V + W$. Thus $\operatorname{span}(\mathcal{X}) \leq V + W$.

Two subspaces

Two subspaces

Now suppose we start with an element $x \in V + W$.

Two subspaces

Now suppose we start with an element $x \in V + W$.
We can then find $y \in V$ and $z \in W$ such that $x = y + z$.

Two subspaces

Now suppose we start with an element $x \in V + W$.
We can then find $y \in V$ and $z \in W$ such that $x = y + z$.
As $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, we have

$$y = \lambda_1 u_1 + \cdots + \lambda_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q$$

for some scalars λ_i, β_j.

Two subspaces

Now suppose we start with an element $x \in V + W$.
We can then find $y \in V$ and $z \in W$ such that $x = y + z$.
As $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, we have

$$y = \lambda_1 u_1 + \cdots + \lambda_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q$$

for some scalars λ_i, β_j.
Similarly, we have

$$z = \mu_1 u_1 + \cdots + \mu_p u_p + \gamma_1 w_1 + \cdots + \gamma_r w_r$$

for some scalars μ_i, γ_k.

Two subspaces

Now suppose we start with an element $x \in V + W$.
We can then find $y \in V$ and $z \in W$ such that $x = y + z$.
As $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, we have

$$y = \lambda_1 u_1 + \cdots + \lambda_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q$$

for some scalars λ_i, β_j.
Similarly, we have

$$z = \mu_1 u_1 + \cdots + \mu_p u_p + \gamma_1 w_1 + \cdots + \gamma_r w_r$$

for some scalars μ_i, γ_k.
If we put $\alpha_i = \lambda_i + \mu_i$ we get

$$x = y + z = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \operatorname{span}(\mathcal{X}).$$

Now suppose we start with an element $x \in V + W$.
We can then find $y \in V$ and $z \in W$ such that $x = y + z$.
As $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, we have

$$y = \lambda_1 u_1 + \cdots + \lambda_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q$$

for some scalars λ_i, β_j.
Similarly, we have

$$z = \mu_1 u_1 + \cdots + \mu_p u_p + \gamma_1 w_1 + \cdots + \gamma_r w_r$$

for some scalars μ_i, γ_k.
If we put $\alpha_i = \lambda_i + \mu_i$ we get

$$x = y + z = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r \in \operatorname{span}(\mathcal{X}).$$

It follows that $\operatorname{span}(\mathcal{X}) = V + W$.

Two subspaces

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$.

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$.

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$.

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$.

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \dots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \dots, u_p, w_1, \dots, w_r)$ is a basis for W

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$.

Two subspaces

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$. Feeding this back into our original relation, we get $\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q = 0$.

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$. Feeding this back into our original relation, we get $\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q = 0$.

The list $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \dots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \dots, u_p, w_1, \dots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$. Feeding this back into our original relation, we get $\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q = 0$.

The list $(u_1, \dots, u_p, v_1, \dots, v_q)$ is a basis for V, so the above gives $\alpha_1 = \cdots = \alpha_p = \beta_1 = \cdots = \beta_q = 0$.

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$. Feeding this back into our original relation, we get $\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q = 0$.

The list $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, so the above gives $\alpha_1 = \cdots = \alpha_p = \beta_1 = \cdots = \beta_q = 0$. As all α's, β's and γ's are zero, we see that our original linear relation was trivial.

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$. Feeding this back into our original relation, we get $\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q = 0$.

The list $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, so the above gives $\alpha_1 = \cdots = \alpha_p = \beta_1 = \cdots = \beta_q = 0$. As all α's, β's and γ's are zero, we see that our original linear relation was trivial. This shows that the list $\mathcal{X}$ is linearly independent

Finally, suppose we have a linear relation

$$\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0.$$

Put $y = \sum_i \alpha_i u_i + \sum_j \beta_j v_j$ and $z = \sum_k \gamma_k w_k$, so $y + z = 0$, so $z = -y$. Now $y \in V$, so z also lies in V, because $z = -y$. Also $z \in W$, so $z \in V \cap W$. We know that $\mathcal{U}$ is a basis for $V \cap W$, so $z = \lambda_1 u_1 + \cdots + \lambda_p u_p$ for some $\lambda_1, \ldots, \lambda_p$. This means that

$$\lambda_1 u_1 + \cdots + \lambda_p u_p - \gamma_1 w_1 - \cdots - \gamma_r w_r = 0.$$

We also know that $(u_1, \ldots, u_p, w_1, \ldots, w_r)$ is a basis for W, so the above gives $\lambda_1 = \cdots = \lambda_p = \gamma_1 = \cdots = \gamma_r = 0$. Feeding this back into our original relation, we get $\alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 v_1 + \cdots + \beta_q v_q = 0$.

The list $(u_1, \ldots, u_p, v_1, \ldots, v_q)$ is a basis for V, so the above gives $\alpha_1 = \cdots = \alpha_p = \beta_1 = \cdots = \beta_q = 0$. As all α's, β's and γ's are zero, we see that our original linear relation was trivial. This shows that the list $\mathcal{X}$ is linearly independent, so it gives a basis for $V + W$ as claimed.

An example

An example

Put $U = M_3\mathbb{R}$ and

$$V = \{A \in U \mid \text{ all rows sum to } 0 \} = \{A \in U \mid A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$W = \{A \in U \mid \text{ all columns sum to } 0 \} = \{A \in U \mid [1,1,1]A = [0,0,0]\}$$

An example

Put $U = M_3\mathbb{R}$ and

$$V = \{A \in U \mid \text{ all rows sum to } 0 \ \} = \{A \in U \mid A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$W = \{A \in U \mid \text{ all columns sum to } 0 \ \} = \{A \in U \mid [1,1,1]A = [0,0,0]\}$$

Then $V \cap W$ is the set of all matrices of the form

$$A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ -a-c & -b-d & a+b+c+d \end{bmatrix} = a \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} + d \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

An example

Put $U = M_3\mathbb{R}$ and

$$V = \{A \in U \mid \text{ all rows sum to } 0 \ \} = \{A \in U \mid A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$W = \{A \in U \mid \text{ all columns sum to } 0 \ \} = \{A \in U \mid [1,1,1]A = [0,0,0]\}$$

Then $V \cap W$ is the set of all matrices of the form

$$A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ -a-c & -b-d & a+b+c+d \end{bmatrix} = a \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} + d \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

It follows that the list

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

is a basis for $V \cap W$.

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix},\ u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix},\ u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix},\ u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix},\ v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix},\ w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},\ w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$.

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$. A typical element of V has the form

$$A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ e & f & -e-f \end{bmatrix}$$

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$. A typical element of V has the form

$$A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ e & f & -e-f \end{bmatrix} = au_1 + bu_2 + cu_3 + du_4 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e-a-c & f-b-d & a+b+c+d-e-f \end{bmatrix}$$

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$. A typical element of V has the form

$$A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ e & f & -e-f \end{bmatrix} = au_1 + bu_2 + cu_3 + du_4 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e-a-c & f-b-d & a+b+c+d-e-f \end{bmatrix}$$

$$= au_1 + bu_2 + cu_3 + du_4 + (e-a-c)v_1 + (f-b-d)v_2.$$

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$. A typical element of V has the form

$$\begin{aligned} A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ e & f & -e-f \end{bmatrix} &= au_1 + bu_2 + cu_3 + du_4 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e-a-c & f-b-d & a+b+c+d-e-f \end{bmatrix} \\ &= au_1 + bu_2 + cu_3 + du_4 + (e-a-c)v_1 + (f-b-d)v_2 . \end{aligned}$$

Using this, we see that $u_1, \ldots, u_4, v_1, v_2$ is a basis for V.

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$. A typical element of V has the form

$$\begin{aligned} A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ e & f & -e-f \end{bmatrix} &= au_1 + bu_2 + cu_3 + du_4 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e-a-c & f-b-d & a+b+c+d-e-f \end{bmatrix} \\ &= au_1 + bu_2 + cu_3 + du_4 + (e-a-c)v_1 + (f-b-d)v_2 . \end{aligned}$$

Using this, we see that $u_1, \ldots, u_4, v_1, v_2$ is a basis for V. Similarly, $u_1, \ldots, u_4, w_1, w_2$ is a basis for W.

An example

$$u_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, u_3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, u_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Now put

$$v_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, w_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

so $v_i \in V$ and $w_i \in W$. A typical element of V has the form

$$A = \begin{bmatrix} a & b & -a-b \\ c & d & -c-d \\ e & f & -e-f \end{bmatrix} = au_1 + bu_2 + cu_3 + du_4 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e-a-c & f-b-d & a+b+c+d-e-f \end{bmatrix}$$
$$= au_1 + bu_2 + cu_3 + du_4 + (e - a - c)v_1 + (f - b - d)v_2.$$

Using this, we see that $u_1, \dots, u_4, v_1, v_2$ is a basis for V. Similarly, $u_1, \dots, u_4, w_1, w_2$ is a basis for W. It follows that

$$u_1, u_2, u_3, u_4, v_1, v_2, w_1, w_2$$

is a basis for $V + W$.

Another example

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$

so $V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3 - x) + b(x^2 - 1)$.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1 \end{aligned}$$

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3 - x) + b(x^2 - 1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$ then $a(x^3-x) + b(x^2-1) + c(x-1) = 0$ for all x

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3 - x) + b(x^2 - 1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$ then $a(x^3 - x) + b(x^2 - 1) + c(x-1) = 0$ for all x, so $ax^3 + bx^2 + (c-a)x - c = 0$ for all x

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$ then $a(x^3-x) + b(x^2-1) + c(x-1) = 0$ for all x, so $ax^3 + bx^2 + (c-a)x - c = 0$ for all x, which implies that $a = b = c = 0$.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$ then $a(x^3 - x) + b(x^2 - 1) + c(x-1) = 0$ for all x, so $ax^3 + bx^2 + (c-a)x - c = 0$ for all x, which implies that $a = b = c = 0$. Our list is thus independent as well as spanning V, so it is a basis.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$V = \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$W = \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\}$$
$$\text{so } V \cap W = \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$ then $a(x^3-x) + b(x^2-1) + c(x-1) = 0$ for all x, so $ax^3 + bx^2 + (c-a)x - c = 0$ for all x, which implies that $a = b = c = 0$. Our list is thus independent as well as spanning V, so it is a basis. Similarly u_1, u_2, w_1 is a basis for W.

Another example

Put $U = \mathbb{R}[x]_{\leq 3}$ and

$$\begin{aligned} V &= \{f \in U \mid f(1) = 0\} = \{(x-1)g(x) \mid g(x) \in \mathbb{R}[x]_{\leq 2}\} \\ W &= \{f \in U \mid f(-1) = 0\} = \{(x+1)h(x) \mid h(x) \in \mathbb{R}[x]_{\leq 2}\} \\ \text{so } V \cap W &= \{f \in U \mid f \text{ is divisible by } (x+1)(x-1) = x^2 - 1\} \end{aligned}$$

Any $f(x) \in V \cap W$ has the form $(ax+b)(x^2-1) = a(x^3-x) + b(x^2-1)$. It follows that the list $u_1 = x^3 - x, u_2 = x^2 - 1$ is a basis for $V \cap W$. Now put $v_1 = x - 1 \in V$ and $w_1 = x + 1 \in W$. We claim that u_1, u_2, v_1 is a basis for V. Indeed, any element of V has the form

$$\begin{aligned} f(x) &= (ax^2 + bx + c).(x-1) = ax^3 + (b-a)x^2 + (c-b)x - c \\ &= au_1 + (b-a)u_2 + (a-b+c)v_1, \end{aligned}$$

so the list spans V. If we have a linear relation $au_1 + bu_2 + cv_1 = 0$ then $a(x^3 - x) + b(x^2 - 1) + c(x-1) = 0$ for all x, so $ax^3 + bx^2 + (c-a)x - c = 0$ for all x, which implies that $a = b = c = 0$. Our list is thus independent as well as spanning V, so it is a basis. Similarly u_1, u_2, w_1 is a basis for W. It follows that u_1, u_2, v_1, w_1 is a basis for $V + W$.

Adapted bases

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces.

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces. Then one can choose a basis $\mathcal{U} = u_1, \dots, u_m$ for U, and a basis $\mathcal{V} = v_1, \dots, v_n$ for V, and an integer $r \leq \min(m, n)$ such that

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces. Then one can choose a basis $\mathcal{U} = u_1, \dots, u_m$ for U, and a basis $\mathcal{V} = v_1, \dots, v_n$ for V, and an integer $r \leq \min(m, n)$ such that

(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces. Then one can choose a basis $\mathcal{U} = u_1, \dots, u_m$ for U, and a basis $\mathcal{V} = v_1, \dots, v_n$ for V, and an integer $r \leq \min(m, n)$ such that

(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$

(b) $\alpha(u_i) = 0$ for $r < i \leq m$

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces. Then one can choose a basis $\mathcal{U} = u_1, \dots, u_m$ for U, and a basis $\mathcal{V} = v_1, \dots, v_n$ for V, and an integer $r \leq \min(m, n)$ such that

(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$

(b) $\alpha(u_i) = 0$ for $r < i \leq m$

(c) $u_{r+1}, \dots, u_m$ is a basis for $\ker(\alpha) \leq U$

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces. Then one can choose a basis $\mathcal{U} = u_1, \dots, u_m$ for U, and a basis $\mathcal{V} = v_1, \dots, v_n$ for V, and an integer $r \leq \min(m, n)$ such that

(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$

(b) $\alpha(u_i) = 0$ for $r < i \leq m$

(c) $u_{r+1}, \dots, u_m$ is a basis for $\ker(\alpha) \leq U$

(d) $v_1, \dots, v_r$ is a basis for $\mathrm{image}(\alpha) \leq V$.

Adapted bases

Theorem ??: Let $\alpha\colon U \to V$ be a linear map between finite-dimensional vector spaces. Then one can choose a basis $\mathcal{U} = u_1, \ldots, u_m$ for U, and a basis $\mathcal{V} = v_1, \ldots, v_n$ for V, and an integer $r \leq \min(m, n)$ such that

(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$

(b) $\alpha(u_i) = 0$ for $r < i \leq m$

(c) $u_{r+1}, \ldots, u_m$ is a basis for $\ker(\alpha) \leq U$

(d) $v_1, \ldots, v_r$ is a basis for $\mathrm{image}(\alpha) \leq V$.

Remark ??: If we use bases as in the theorem, then the matrix of α with respect to those bases has the form

$$A = \left[\begin{array}{c|c} I_r & 0_{r,m-r} \\ \hline 0_{n-r,r} & 0_{n-r,m-r} \end{array}\right]$$

Adapted bases

$\mathcal{U} = u_1, \ldots, u_m$ a basis for U	$\mathcal{V} = v_1, \ldots, v_n$ a basis for V
$\alpha(u_i) = v_i$ for $1 \leq i \leq r$	$\alpha(u_i) = 0$ for $r < i \leq m$
$u_{r+1}, \ldots, u_m$ a basis for $\ker(\alpha) \leq U$	$v_1, \ldots, v_r$ a basis for $\text{image}(\alpha) \leq V$

Adapted bases

$\mathcal{U} = u_1, \ldots, u_m$ a basis for U	$\mathcal{V} = v_1, \ldots, v_n$ a basis for V
$\alpha(u_i) = v_i$ for $1 \leq i \leq r$	$\alpha(u_i) = 0$ for $r < i \leq m$
$u_{r+1}, \ldots, u_m$ a basis for $\ker(\alpha) \leq U$	$v_1, \ldots, v_r$ a basis for $\mathrm{image}(\alpha) \leq V$

Corollary ??: If $\alpha \colon U \to V$ is a linear map then

$$\dim(\ker(\alpha)) + \dim(\mathrm{image}(\alpha)) = \dim(U).$$

Adapted bases

$\mathcal{U} = u_1, \ldots, u_m$ a basis for U
$\alpha(u_i) = v_i$ for $1 \leq i \leq r$
$u_{r+1}, \ldots, u_m$ a basis for $\ker(\alpha) \leq U$

$\mathcal{V} = v_1, \ldots, v_n$ a basis for V
$\alpha(u_i) = 0$ for $r < i \leq m$
$v_1, \ldots, v_r$ a basis for $\operatorname{image}(\alpha) \leq V$

Corollary ??: If $\alpha\colon U \to V$ is a linear map then

$$\dim(\ker(\alpha)) + \dim(\operatorname{image}(\alpha)) = \dim(U).$$

Proof: Choose bases as in the theorem. Then $\dim(U) = m$ and $\dim(\operatorname{image}(\alpha)) = r$ and

$$\dim(\ker(\alpha)) = |\{u_{r+1}, \ldots, u_m\}| = m - r.$$

The claim follows.

Adapted bases

$\mathcal{U} = u_1, \ldots, u_m$ a basis for U	$\mathcal{V} = v_1, \ldots, v_n$ a basis for V
(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$	(b) $\alpha(u_i) = 0$ for $r < i \leq m$
(c) $u_{r+1}, \ldots, u_m$ a basis for $\ker(\alpha)$	(d) $v_1, \ldots, v_r$ a basis for $\text{image}(\alpha)$

Proof of Theorem ??:

Adapted bases

$\mathcal{U} = u_1, \dots, u_m$ a basis for U	$\mathcal{V} = v_1, \dots, v_n$ a basis for V
(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$	(b) $\alpha(u_i) = 0$ for $r < i \leq m$
(c) $u_{r+1}, \dots, u_m$ a basis for $\ker(\alpha)$	(d) $v_1, \dots, v_r$ a basis for $\mathrm{image}(\alpha)$

Proof of Theorem ??:
Let $v_1, \dots, v_r$ be any basis for image(α) (so (d) is satisfied).

Adapted bases

$\mathcal{U} = u_1, \dots, u_m$ a basis for U	$\mathcal{V} = v_1, \dots, v_n$ a basis for V
(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$	(b) $\alpha(u_i) = 0$ for $r < i \leq m$
(c) $u_{r+1}, \dots, u_m$ a basis for $\ker(\alpha)$	(d) $v_1, \dots, v_r$ a basis for $\operatorname{image}(\alpha)$

Proof of Theorem ??:
Let $v_1, \dots, v_r$ be any basis for image(α) (so (d) is satisfied).

By Proposition **??**, this can be extended to a list $\mathcal{V} = v_1, \dots, v_n$ which is a basis for all of V.

Adapted bases

$\mathcal{U} = u_1, \dots, u_m$ a basis for U	$\mathcal{V} = v_1, \dots, v_n$ a basis for V
(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$	(b) $\alpha(u_i) = 0$ for $r < i \leq m$
(c) $u_{r+1}, \dots, u_m$ a basis for $\ker(\alpha)$	(d) $v_1, \dots, v_r$ a basis for $\text{image}(\alpha)$

Proof of Theorem ??:
Let $v_1, \dots, v_r$ be any basis for $\text{image}(\alpha)$ (so (d) is satisfied).

By Proposition **??**, this can be extended to a list $\mathcal{V} = v_1, \dots, v_n$ which is a basis for all of V.

Next, for $j \leq r$ we have $v_j \in \text{image}(\alpha)$, so we can choose $u_j \in U$ with $\alpha(u_j) = v_j$ (so (a) is satisfied).

$\mathcal{U} = u_1, \dots, u_m$ a basis for U	$\mathcal{V} = v_1, \dots, v_n$ a basis for V
(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$	(b) $\alpha(u_i) = 0$ for $r < i \leq m$
(c) $u_{r+1}, \dots, u_m$ a basis for $\ker(\alpha)$	(d) $v_1, \dots, v_r$ a basis for $\operatorname{image}(\alpha)$

Proof of Theorem ??:
Let $v_1, \dots, v_r$ be any basis for image(α) (so (d) is satisfied).

By Proposition **??**, this can be extended to a list $\mathcal{V} = v_1, \dots, v_n$ which is a basis for all of V.

Next, for $j \leq r$ we have $v_j \in \operatorname{image}(\alpha)$, so we can choose $u_j \in U$ with $\alpha(u_j) = v_j$ (so (a) is satisfied).

This gives us a list $u_1, \dots, u_r$ of elements of U; to these, we add vectors $u_{r+1}, \dots, u_m$ forming a basis for ker(α) (so that (b) and (c) are satisfied).

$\mathcal{U} = u_1, \dots, u_m$ a basis for U	$\mathcal{V} = v_1, \dots, v_n$ a basis for V
(a) $\alpha(u_i) = v_i$ for $1 \leq i \leq r$	(b) $\alpha(u_i) = 0$ for $r < i \leq m$
(c) $u_{r+1}, \dots, u_m$ a basis for $\ker(\alpha)$	(d) $v_1, \dots, v_r$ a basis for $\operatorname{image}(\alpha)$

Proof of Theorem ??:
Let $v_1, \dots, v_r$ be any basis for image(α) (so (d) is satisfied).

By Proposition **??**, this can be extended to a list $\mathcal{V} = v_1, \dots, v_n$ which is a basis for all of V.

Next, for $j \leq r$ we have $v_j \in \operatorname{image}(\alpha)$, so we can choose $u_j \in U$ with $\alpha(u_j) = v_j$ (so (a) is satisfied).

This gives us a list $u_1, \dots, u_r$ of elements of U; to these, we add vectors $u_{r+1}, \dots, u_m$ forming a basis for ker(α) (so that (b) and (c) are satisfied).

Now everything is as claimed except that we have not shown that the list $\mathcal{U} = u_1, \dots, u_m$ is a basis for U.

Adapted bases

Adapted bases

Consider an element $x \in U$.

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \mathrm{image}(\alpha)$

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \ldots, v_r$ is a basis for $\text{image}(\alpha)$

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \dots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \dots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \dots + \lambda_r v_r$.

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \dots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \dots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \dots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \dots + \lambda_r u_r$, and $x'' = x - x'$.

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \dots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \dots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x)$$

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \dots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \dots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \mathrm{image}(\alpha)$, and $v_1, \ldots, v_r$ is a basis for $\mathrm{image}(\alpha)$, so there exist numbers $\lambda_1, \ldots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$, so $x'' \in \ker(\alpha)$.

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \ldots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \ldots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$, so $x'' \in \ker(\alpha)$. We also know that $u_{r+1}, \ldots, u_m$ is a basis for $\ker(\alpha)$

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \dots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \dots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \dots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \dots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$, so $x'' \in \ker(\alpha)$. We also know that $u_{r+1}, \dots, u_m$ is a basis for $\ker(\alpha)$, so there exist numbers $\lambda_{r+1}, \dots, \lambda_m$ with $x'' = \lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m$.

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \ldots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \ldots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$, so $x'' \in \ker(\alpha)$. We also know that $u_{r+1}, \ldots, u_m$ is a basis for $\ker(\alpha)$, so there exist numbers $\lambda_{r+1}, \ldots, \lambda_m$ with $x'' = \lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m$. Putting this together, we get

$$x = x' + x'' = (\lambda_1 u_1 + \cdots + \lambda_r u_r) + (\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m)$$

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \ldots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \ldots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$, so $x'' \in \ker(\alpha)$. We also know that $u_{r+1}, \ldots, u_m$ is a basis for $\ker(\alpha)$, so there exist numbers $\lambda_{r+1}, \ldots, \lambda_m$ with $x'' = \lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m$. Putting this together, we get

$$x = x' + x'' = (\lambda_1 u_1 + \cdots + \lambda_r u_r) + (\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m),$$

which is a linear combination of $u_1, \ldots, u_m$.

Adapted bases

Consider an element $x \in U$. We then have $\alpha(x) \in \text{image}(\alpha)$, and $v_1, \dots, v_r$ is a basis for $\text{image}(\alpha)$, so there exist numbers $\lambda_1, \dots, \lambda_r$ such that $\alpha(x) = \lambda_1 v_1 + \ldots + \lambda_r v_r$. Now put $x' = \lambda_1 u_1 + \ldots + \lambda_r u_r$, and $x'' = x - x'$. We have

$$\alpha(x') = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) = \lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha(x),$$

so $\alpha(x'') = \alpha(x) - \alpha(x') = 0$, so $x'' \in \ker(\alpha)$. We also know that $u_{r+1}, \dots, u_m$ is a basis for $\ker(\alpha)$, so there exist numbers $\lambda_{r+1}, \dots, \lambda_m$ with $x'' = \lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m$. Putting this together, we get

$$x = x' + x'' = (\lambda_1 u_1 + \cdots + \lambda_r u_r) + (\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m),$$

which is a linear combination of $u_1, \dots, u_m$. It follows that the list $\mathcal{U}$ spans U.

Adapted bases

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$.

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$0 = \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1} \alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m)$$

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1\alpha(u_1) + \cdots + \lambda_r\alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m\alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \end{aligned}$$

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1\alpha(u_1) + \cdots + \lambda_r\alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m\alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1} \alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α)

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1} \alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned}0 &= \lambda_1\alpha(u_1) + \cdots + \lambda_r\alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m\alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r.\end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$.

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1} \alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \ldots, u_m$ is a basis for ker(α), these vectors are linearly independent

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1} \alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \ldots, u_m$ is a basis for ker(α), these vectors are linearly independent, so the above relation must be trivial

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1\alpha(u_1) + \cdots + \lambda_r\alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m\alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \dots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \dots, u_m$ is a basis for ker(α), these vectors are linearly independent, so the above relation must be trivial, so $\lambda_{r+1} = \cdots = \lambda_m = 0$.

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \ldots, u_m$ is a basis for ker(α), these vectors are linearly independent, so the above relation must be trivial, so $\lambda_{r+1} = \cdots = \lambda_m = 0$. This shows that all the λ's are zero

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned}
0 &= \lambda_1\alpha(u_1) + \cdots + \lambda_r\alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m\alpha(u_m) \\
&= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\
&= \lambda_1 v_1 + \cdots + \lambda_r v_r.
\end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \ldots, u_m$ is a basis for ker(α), these vectors are linearly independent, so the above relation must be trivial, so $\lambda_{r+1} = \cdots = \lambda_m = 0$. This shows that all the λ's are zero, so the original relation was trivial.

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1 \alpha(u_1) + \cdots + \lambda_r \alpha(u_r) + \lambda_{r+1} \alpha(u_{r+1}) + \cdots + \lambda_m \alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \ldots, u_m$ is a basis for ker(α), these vectors are linearly independent, so the above relation must be trivial, so $\lambda_{r+1} = \cdots = \lambda_m = 0$. This shows that all the λ's are zero, so the original relation was trivial. Thus, the vectors $u_1, \ldots, u_m$ are linearly independent.

Adapted bases

Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$. We apply α to both sides of this equation to get

$$\begin{aligned} 0 &= \lambda_1\alpha(u_1) + \cdots + \lambda_r\alpha(u_r) + \lambda_{r+1}\alpha(u_{r+1}) + \cdots + \lambda_m\alpha(u_m) \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1}.0 + \cdots + \lambda_m.0 \\ &= \lambda_1 v_1 + \cdots + \lambda_r v_r. \end{aligned}$$

This is a linear relation between the vectors $v_1, \ldots, v_r$, but these form a basis for image(α), so this must be the trivial relation, so $\lambda_1 = \cdots = \lambda_r = 0$. This means that our original relation has the form

$$\lambda_{r+1} u_{r+1} + \cdots + \lambda_m u_m = 0$$

As $u_{r+1}, \ldots, u_m$ is a basis for ker(α), these vectors are linearly independent, so the above relation must be trivial, so $\lambda_{r+1} = \cdots = \lambda_m = 0$. This shows that all the λ's are zero, so the original relation was trivial. Thus, the vectors $u_1, \ldots, u_m$ are linearly independent. We have already seen that they span U, so they give a basis for U.

An example

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, or equivalently

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right] = \left[\begin{smallmatrix} b+d & a+c \\ b+d & a+c \end{smallmatrix}\right]$$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, or equivalently

$$\begin{aligned} \phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] &= \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right] = \left[\begin{smallmatrix} b+d & a+c \\ b+d & a+c \end{smallmatrix}\right] \\ &= (b+d)\left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]. \end{aligned}$$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, or equivalently

$$\begin{aligned} \phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] &= \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right] = \left[\begin{smallmatrix} b+d & a+c \\ b+d & a+c \end{smallmatrix}\right] \\ &= (b+d)\left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]. \end{aligned}$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]$$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, or equivalently

$$\begin{aligned} \phi \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] &= \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right] = \left[\begin{smallmatrix} b+d & a+c \\ b+d & a+c \end{smallmatrix}\right] \\ &= (b+d) \left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] + (a+c) \left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]. \end{aligned}$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for image(ϕ).

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right] A \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$, or equivalently

$$\begin{aligned}\phi\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right] &= \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}b&a\\d&c\end{smallmatrix}\right] = \left[\begin{smallmatrix}b+d&a+c\\b+d&a+c\end{smallmatrix}\right]\\ &= (b+d)\left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right].\end{aligned}$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for $\text{image}(\phi)$. It can be extended to a basis for all of $M_2\mathbb{R}$ by adding $v_3 = \left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right]$ and $v_4 = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$.

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right] A \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$, or equivalently

$$\phi\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}b&a\\d&c\end{smallmatrix}\right] = \left[\begin{smallmatrix}b+d&a+c\\b+d&a+c\end{smallmatrix}\right]$$
$$= (b+d)\left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right].$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for image(ϕ). It can be extended to a basis for all of $M_2\mathbb{R}$ by adding $v_3 = \left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right]$ and $v_4 = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$. Moreover, we have $\phi(A) = 0$ iff $a + c = b + d = 0$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right] A \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$, or equivalently

$$\begin{aligned}\phi\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right] &= \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}b&a\\d&c\end{smallmatrix}\right] = \left[\begin{smallmatrix}b+d&a+c\\b+d&a+c\end{smallmatrix}\right] \\ &= (b+d)\left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right].\end{aligned}$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for image(ϕ). It can be extended to a basis for all of $M_2\mathbb{R}$ by adding $v_3 = \left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right]$ and $v_4 = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$. Moreover, we have $\phi(A) = 0$ iff $a + c = b + d = 0$ iff $c = -a$ and $d = -b$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right] A \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$, or equivalently

$$\phi\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right]\left[\begin{smallmatrix}b&a\\d&c\end{smallmatrix}\right] = \left[\begin{smallmatrix}b+d&a+c\\b+d&a+c\end{smallmatrix}\right]$$
$$= (b+d)\left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right].$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for image(ϕ). It can be extended to a basis for all of $M_2\mathbb{R}$ by adding $v_3 = \left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right]$ and $v_4 = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$. Moreover, we have $\phi(A) = 0$ iff $a+c = b+d = 0$ iff $c = -a$ and $d = -b$, in which case

$$A = \left[\begin{smallmatrix}a&b\\-a&-b\end{smallmatrix}\right] = a\left[\begin{smallmatrix}1&0\\-1&0\end{smallmatrix}\right] + b\left[\begin{smallmatrix}0&1\\0&-1\end{smallmatrix}\right].$$

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, or equivalently

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right] = \left[\begin{smallmatrix} b+d & a+c \\ b+d & a+c \end{smallmatrix}\right]$$
$$= (b+d)\left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right].$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for image(ϕ). It can be extended to a basis for all of $M_2\mathbb{R}$ by adding $v_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ and $v_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. Moreover, we have $\phi(A) = 0$ iff $a + c = b + d = 0$ iff $c = -a$ and $d = -b$, in which case

$$A = \left[\begin{smallmatrix} a & b \\ -a & -b \end{smallmatrix}\right] = a\left[\begin{smallmatrix} 1 & 0 \\ -1 & 0 \end{smallmatrix}\right] + b\left[\begin{smallmatrix} 0 & 1 \\ 0 & -1 \end{smallmatrix}\right].$$

This means that the matrices $u_3 = \left[\begin{smallmatrix} 1 & 0 \\ -1 & 0 \end{smallmatrix}\right]$ and $u_4 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & -1 \end{smallmatrix}\right]$ form a basis for $\ker(\phi)$.

An example

Consider the map $\phi\colon M_2\mathbb{R} \to M_2\mathbb{R}$ given by $\phi(A) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, or equivalently

$$\phi\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right] = \left[\begin{smallmatrix} b+d & a+c \\ b+d & a+c \end{smallmatrix}\right]$$
$$= (b+d)\left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] + (a+c)\left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right].$$

It follows that if we put

$$u_1 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \quad u_2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] \quad v_1 = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] \quad v_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right]$$

then $\phi(u_1) = v_1$, $\phi(u_2) = v_2$, and v_1, v_2 is a basis for image(ϕ). It can be extended to a basis for all of $M_2\mathbb{R}$ by adding $v_3 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ and $v_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. Moreover, we have $\phi(A) = 0$ iff $a + c = b + d = 0$ iff $c = -a$ and $d = -b$, in which case

$$A = \left[\begin{smallmatrix} a & b \\ -a & -b \end{smallmatrix}\right] = a\left[\begin{smallmatrix} 1 & 0 \\ -1 & 0 \end{smallmatrix}\right] + b\left[\begin{smallmatrix} 0 & 1 \\ 0 & -1 \end{smallmatrix}\right].$$

This means that the matrices $u_3 = \left[\begin{smallmatrix} 1 & 0 \\ -1 & 0 \end{smallmatrix}\right]$ and $u_4 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & -1 \end{smallmatrix}\right]$ form a basis for $\ker(\phi)$. Putting this together, we see that $u_1, \dots, u_4$ and $v_1, \dots, v_4$ are bases for $M_2\mathbb{R}$ such that $\phi(u_i) = v_i$ for $i \leq 2$, and $\phi(u_i) = 0$ for $i > 2$.

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors

Definition ??:

Eigenvalues and eigenvectors

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha\colon V \to V$ be a $\mathbb{C}$-linear map.

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha\colon V \to V$ be a $\mathbb{C}$-linear map. Let λ be a complex number.

Eigenvalues and eigenvectors

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha\colon V \to V$ be a $\mathbb{C}$-linear map. Let λ be a complex number. An *eigenvector* for α, with *eigenvalue* λ is a nonzero element $v \in V$ such that $\alpha(v) = \lambda v$.

Eigenvalues and eigenvectors

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha\colon V \to V$ be a $\mathbb{C}$-linear map. Let λ be a complex number. An *eigenvector* for α, with *eigenvalue* λ is a nonzero element $v \in V$ such that $\alpha(v) = \lambda v$. If such a v exists, we say that λ is an *eigenvalue of* α.

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha\colon V \to V$ be a $\mathbb{C}$-linear map. Let λ be a complex number. An *eigenvector* for α, with *eigenvalue* λ is a nonzero element $v \in V$ such that $\alpha(v) = \lambda v$. If such a v exists, we say that λ is an *eigenvalue of* α.

Remark ??: Suppose we choose a basis $\mathcal{V}$ for V, and let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$.

Eigenvalues and eigenvectors

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha \colon V \to V$ be a $\mathbb{C}$-linear map. Let λ be a complex number. An *eigenvector* for α, with *eigenvalue* λ is a nonzero element $v \in V$ such that $\alpha(v) = \lambda v$. If such a v exists, we say that λ is an *eigenvalue of* α.

Remark ??: Suppose we choose a basis $\mathcal{V}$ for V, and let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$. Then the eigenvalues of α are the same as the eigenvalues of the matrix A

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $\alpha\colon V \to V$ be a $\mathbb{C}$-linear map. Let λ be a complex number. An *eigenvector* for α, with *eigenvalue* λ is a nonzero element $v \in V$ such that $\alpha(v) = \lambda v$. If such a v exists, we say that λ is an *eigenvalue of* α.

Remark ??: Suppose we choose a basis $\mathcal{V}$ for V, and let A be the matrix of α with respect to $\mathcal{V}$ and $\mathcal{V}$. Then the eigenvalues of α are the same as the eigenvalues of the matrix A, which are the roots of the characteristic polynomial $\det(tI - A)$.

Eigenvalue examples

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi \colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \dots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \dots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P)$$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \dots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = \det \begin{bmatrix} t-1 & -1 & -1 & -1 & -1 \\ 0 & t-1 & -2 & -3 & -4 \\ 0 & 0 & t-1 & -3 & -6 \\ 0 & 0 & 0 & t-1 & -4 \\ 0 & 0 & 0 & 0 & t-1 \end{bmatrix}$$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi \colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \ldots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = \det \begin{bmatrix} t-1 & -1 & -1 & -1 & -1 \\ 0 & t-1 & -2 & -3 & -4 \\ 0 & 0 & t-1 & -3 & -6 \\ 0 & 0 & 0 & t-1 & -4 \\ 0 & 0 & 0 & 0 & t-1 \end{bmatrix} = (t-1)^5$$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \ldots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = \det \begin{bmatrix} t-1 & -1 & -1 & -1 & -1 \\ 0 & t-1 & -2 & -3 & -4 \\ 0 & 0 & t-1 & -3 & -6 \\ 0 & 0 & 0 & t-1 & -4 \\ 0 & 0 & 0 & 0 & t-1 \end{bmatrix} = (t-1)^5$$

so 1 is the only root of the characteristic polynomial.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \ldots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = \det \begin{bmatrix} t-1 & -1 & -1 & -1 & -1 \\ 0 & t-1 & -2 & -3 & -4 \\ 0 & 0 & t-1 & -3 & -6 \\ 0 & 0 & 0 & t-1 & -4 \\ 0 & 0 & 0 & 0 & t-1 \end{bmatrix} = (t-1)^5$$

so 1 is the only root of the characteristic polynomial. The eigenvectors are just the polynomials f with $\phi(f) = 1.f$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \ldots, x^4$) is

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = \det \begin{bmatrix} t-1 & -1 & -1 & -1 & -1 \\ 0 & t-1 & -2 & -3 & -4 \\ 0 & 0 & t-1 & -3 & -6 \\ 0 & 0 & 0 & t-1 & -4 \\ 0 & 0 & 0 & 0 & t-1 \end{bmatrix} = (t-1)^5$$

so 1 is the only root of the characteristic polynomial. The eigenvectors are just the polynomials f with $\phi(f) = 1.f$ or equivalently $f(x+1) = f(x)$ for all x.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(x+1)$ so $\phi(x^k) = (x+1)^k$ for all $k \leq 4$. We claim that 1 is the only eigenvalue. Indeed, the corresponding matrix P (with respect to the basis $1, x, \dots, x^4$) is

$$P = \begin{bmatrix} 1&1&1&1&1 \\ 0&1&2&3&4 \\ 0&0&1&3&6 \\ 0&0&0&1&4 \\ 0&0&0&0&1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = \det \begin{bmatrix} t-1 & -1 & -1 & -1 & -1 \\ 0 & t-1 & -2 & -3 & -4 \\ 0 & 0 & t-1 & -3 & -6 \\ 0 & 0 & 0 & t-1 & -4 \\ 0 & 0 & 0 & 0 & t-1 \end{bmatrix} = (t-1)^5$$

so 1 is the only root of the characteristic polynomial. The eigenvectors are just the polynomials f with $\phi(f) = 1.f$ or equivalently $f(x+1) = f(x)$ for all x. These are just the constant polynomials.

Eigenvalue examples

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi \colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = (t-1)(t-i)(t+1)(t+i)(t-1)$$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi \colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = (t-1)(t-i)(t+1)(t+i)(t-1) = (t-1)(t^2+1)(t^2-1) = t^5 - t^4 - t + 1$$

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI-P) = (t-1)(t-i)(t+1)(t+i)(t-1) = (t-1)(t^2+1)(t^2-1) = t^5-t^4-t+1$$

so the eigenvalues are $1, i, -1$ and $-i$.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = (t-1)(t-i)(t+1)(t+i)(t-1) = (t-1)(t^2+1)(t^2-1) = t^5 - t^4 - t + 1$$

so the eigenvalues are $1, i, -1$ and $-i$.

- The eigenvectors of eigenvalue 1 are functions $f \in V$ with $f(ix) = f(x)$. These are the functions of the form $f(x) = a + ex^4$.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = (t-1)(t-i)(t+1)(t+i)(t-1) = (t-1)(t^2+1)(t^2-1) = t^5 - t^4 - t + 1$$

so the eigenvalues are $1, i, -1$ and $-i$.

- The eigenvectors of eigenvalue 1 are functions $f \in V$ with $f(ix) = f(x)$. These are the functions of the form $f(x) = a + ex^4$.
- The eigenvectors of eigenvalue i are functions $f \in V$ with $f(ix) = if(x)$. These are the functions of the form $f(x) = bx$.

Eigenvalue examples

Example ??: Put $V = \mathbb{C}[x]_{\leq 4}$, and define $\phi\colon V \to V$ by $\phi(f)(x) = f(ix)$, so $\phi(x^k) = i^k x^k$. The corresponding matrix P (with respect to $1, x, x^2, x^3, x^4$) is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - P) = (t-1)(t-i)(t+1)(t+i)(t-1) = (t-1)(t^2+1)(t^2-1) = t^5 - t^4 - t + 1$$

so the eigenvalues are $1, i, -1$ and $-i$.

- The eigenvectors of eigenvalue 1 are functions $f \in V$ with $f(ix) = f(x)$. These are the functions of the form $f(x) = a + ex^4$.
- The eigenvectors of eigenvalue i are functions $f \in V$ with $f(ix) = if(x)$. These are the functions of the form $f(x) = bx$.
- The eigenvectors of eigenvalue -1 are functions $f \in V$ with $f(ix) = -f(x)$. These are the functions of the form $f(x) = cx^2$.

Eigenvalue examples

Eigenvalue examples

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$.

Eigenvalue examples

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Eigenvalue examples

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix}$$

Eigenvalue examples

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix} = t^3 + t$$

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix} = t^3 + t = t(t+i)(t-i).$$

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix} = t^3 + t = t(t+i)(t-i).$$

The eigenvalues are thus 0, i and $-i$.

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix} = t^3 + t = t(t+i)(t-i).$$

The eigenvalues are thus 0, i and $-i$.

- The eigenvectors of eigenvalue 0 are the multiples of $\mathbf{a}$.

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix} = t^3 + t = t(t+i)(t-i).$$

The eigenvalues are thus 0, i and $-i$.

- The eigenvectors of eigenvalue 0 are the multiples of $\mathbf{a}$.
- The eigenvectors of eigenvalue i are the multiples of $\mathbf{b} - i\mathbf{c}$.

Eigenvalue examples

Example ??: Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^3$ and define $\alpha\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\alpha(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$. Choose a unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. We saw previously that the matrix of α with respect to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is thus

$$\det(tI - A) = \det \begin{bmatrix} t & 0 & 0 \\ 0 & t & 1 \\ 0 & -1 & t \end{bmatrix} = t^3 + t = t(t+i)(t-i).$$

The eigenvalues are thus 0, i and $-i$.

- The eigenvectors of eigenvalue 0 are the multiples of $\mathbf{a}$.
- The eigenvectors of eigenvalue i are the multiples of $\mathbf{b} - i\mathbf{c}$.
- The eigenvectors of eigenvalue $-i$ are the multiples of $\mathbf{b} + i\mathbf{c}$.

Eigenvalue examples

Eigenvalue examples

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{v}$.

Eigenvalue examples

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{v}$. We claim that the characteristic polynomial of ϕ is $t^2(t - \mathbf{u}.\mathbf{v})$.

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{v}$. We claim that the characteristic polynomial of ϕ is $t^2(t - \mathbf{u}.\mathbf{v})$. Indeed, the matrix P with respect to the standard basis is calculated as follows:

$$\phi(\mathbf{e}_1) = u_1\mathbf{v} = \begin{bmatrix} u_1 v_1 \\ u_1 v_2 \\ u_1 v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_2) = u_2\mathbf{v} = \begin{bmatrix} u_2 v_1 \\ u_2 v_2 \\ u_2 v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_3) = u_3\mathbf{v} = \begin{bmatrix} u_3 v_1 \\ u_3 v_2 \\ u_3 v_3 \end{bmatrix}$$

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{v}$. We claim that the characteristic polynomial of ϕ is $t^2(t - \mathbf{u}.\mathbf{v})$. Indeed, the matrix P with respect to the standard basis is calculated as follows:

$$\phi(\mathbf{e}_1) = u_1\mathbf{v} = \begin{bmatrix} u_1v_1 \\ u_1v_2 \\ u_1v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_2) = u_2\mathbf{v} = \begin{bmatrix} u_2v_1 \\ u_2v_2 \\ u_2v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_3) = u_3\mathbf{v} = \begin{bmatrix} u_3v_1 \\ u_3v_2 \\ u_3v_3 \end{bmatrix}$$

$$P = \begin{bmatrix} u_1v_1 & u_2v_1 & u_3v_1 \\ u_1v_2 & u_2v_2 & u_3v_2 \\ u_1v_3 & u_2v_3 & u_3v_3 \end{bmatrix}$$

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{v}$. We claim that the characteristic polynomial of ϕ is $t^2(t - \mathbf{u}.\mathbf{v})$. Indeed, the matrix P with respect to the standard basis is calculated as follows:

$$\phi(\mathbf{e}_1) = u_1 \mathbf{v} = \begin{bmatrix} u_1 v_1 \\ u_1 v_2 \\ u_1 v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_2) = u_2 \mathbf{v} = \begin{bmatrix} u_2 v_1 \\ u_2 v_2 \\ u_2 v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_3) = u_3 \mathbf{v} = \begin{bmatrix} u_3 v_1 \\ u_3 v_2 \\ u_3 v_3 \end{bmatrix}$$

$$P = \begin{bmatrix} u_1 v_1 & u_2 v_1 & u_3 v_1 \\ u_1 v_2 & u_2 v_2 & u_3 v_2 \\ u_1 v_3 & u_2 v_3 & u_3 v_3 \end{bmatrix}$$

The characteristic polynomial is $\det(tI - P) = -\det(P - tI)$

Eigenvalue examples

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x}\rangle \mathbf{v}$. We claim that the characteristic polynomial of ϕ is $t^2(t - \mathbf{u}.\mathbf{v})$. Indeed, the matrix P with respect to the standard basis is calculated as follows:

$$\phi(\mathbf{e}_1) = u_1\mathbf{v} = \begin{bmatrix} u_1v_1 \\ u_1v_2 \\ u_1v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_2) = u_2\mathbf{v} = \begin{bmatrix} u_2v_1 \\ u_2v_2 \\ u_2v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_3) = u_3\mathbf{v} = \begin{bmatrix} u_3v_1 \\ u_3v_2 \\ u_3v_3 \end{bmatrix}$$

$$P = \begin{bmatrix} u_1v_1 & u_2v_1 & u_3v_1 \\ u_1v_2 & u_2v_2 & u_3v_2 \\ u_1v_3 & u_2v_3 & u_3v_3 \end{bmatrix}$$

The characteristic polynomial is $\det(tI - P) = -\det(P - tI)$, which is found as follows:

$$\begin{aligned} &\det(P - tI) \\ &= \det \begin{bmatrix} u_1v_1 - t & u_2v_1 & u_3v_1 \\ u_1v_2 & u_2v_2 - t & u_3v_2 \\ u_1v_3 & u_2v_3 & u_3v_3 - t \end{bmatrix} \end{aligned}$$

Example ??: Let $\mathbf{u}$ and $\mathbf{v}$ be non-orthogonal vectors in $\mathbb{R}^3$, and define $\phi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\phi(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x} \rangle \mathbf{v}$. We claim that the characteristic polynomial of ϕ is $t^2(t - \mathbf{u}.\mathbf{v})$. Indeed, the matrix P with respect to the standard basis is calculated as follows:

$$\phi(\mathbf{e}_1) = u_1 \mathbf{v} = \begin{bmatrix} u_1 v_1 \\ u_1 v_2 \\ u_1 v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_2) = u_2 \mathbf{v} = \begin{bmatrix} u_2 v_1 \\ u_2 v_2 \\ u_2 v_3 \end{bmatrix} \qquad \phi(\mathbf{e}_3) = u_3 \mathbf{v} = \begin{bmatrix} u_3 v_1 \\ u_3 v_2 \\ u_3 v_3 \end{bmatrix}$$

$$P = \begin{bmatrix} u_1 v_1 & u_2 v_1 & u_3 v_1 \\ u_1 v_2 & u_2 v_2 & u_3 v_2 \\ u_1 v_3 & u_2 v_3 & u_3 v_3 \end{bmatrix}$$

The characteristic polynomial is $\det(tI - P) = -\det(P - tI)$, which is found as follows:

$$\begin{aligned}
&\det(P - tI) \\
&= \det \begin{bmatrix} u_1 v_1 - t & u_2 v_1 & u_3 v_1 \\ u_1 v_2 & u_2 v_2 - t & u_3 v_2 \\ u_1 v_3 & u_2 v_3 & u_3 v_3 - t \end{bmatrix} \\
&= (u_1 v_1 - t) \det \begin{bmatrix} u_2 v_2 - t & u_3 v_2 \\ u_2 v_3 & u_3 v_3 - t \end{bmatrix} - u_2 v_1 \det \begin{bmatrix} u_1 v_2 & u_3 v_2 \\ u_1 v_3 & u_3 v_3 - t \end{bmatrix} + u_3 v_1 \det \begin{bmatrix} u_1 v_2 & u_2 v_2 - t \\ u_1 v_3 & u_2 v_3 \end{bmatrix}
\end{aligned}$$

Eigenvalue examples

Eigenvalue examples

$$\det \begin{bmatrix} u_2 v_2 - t & u_3 v_2 \\ u_2 v_3 & u_3 v_3 - t \end{bmatrix} = (u_2 v_2 - t)(u_3 v_3 - t) - u_2 v_3 u_3 v_2$$

Eigenvalue examples

$$\det \begin{bmatrix} u_2 v_2 - t & u_3 v_2 \\ u_2 v_3 & u_3 v_3 - t \end{bmatrix} = (u_2 v_2 - t)(u_3 v_3 - t) - u_2 v_3 u_3 v_2 = t^2 - (u_2 v_2 + u_3 v_3)t$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

Eigenvalue examples

$$\det \begin{bmatrix} u_2 v_2 - t & u_3 v_2 \\ u_2 v_3 & u_3 v_3 - t \end{bmatrix} = (u_2 v_2 - t)(u_3 v_3 - t) - u_2 v_3 u_3 v_2 = t^2 - (u_2 v_2 + u_3 v_3) t$$

$$\det \begin{bmatrix} u_1 v_2 & u_3 v_2 \\ u_1 v_3 & u_3 v_3 - t \end{bmatrix} = u_1 v_2 (u_3 v_3 - t) - u_1 v_3 u_3 v_2 = -u_1 v_2 t$$

$$\det \begin{bmatrix} u_1 v_2 & u_2 v_2 - t \\ u_1 v_3 & u_2 v_3 \end{bmatrix} = u_1 v_2 u_2 v_3 - u_1 v_3 (u_2 v_2 - t)$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2 - t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\det(P - tI) = (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\begin{aligned} \det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \end{aligned}$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2-t)(u_3v_3-t) - u_2v_3u_3v_2 = t^2-(u_2v_2+u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3-t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2-t) = u_1v_3t$$

$$\begin{aligned} \det(P-tI) &= (u_1v_1-t)(t^2-(u_2v_2+u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1+u_2v_2+u_3v_3)t^2 - t^3 \\ \det(tI-P) &= t^3-(u_1v_1+u_2v_2+u_3v_3)t^2 \end{aligned}$$

Eigenvalue examples

$$
\begin{aligned}
\det \begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} &= (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t \\
\det \begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} &= u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t \\
\det \begin{bmatrix} u_1v_2 & u_2v_2 - t \\ u_1v_3 & u_2v_3 \end{bmatrix} &= u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t \\
\det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\
&= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \\
\det(tI - P) &= t^3 - (u_1v_1 + u_2v_2 + u_3v_3)t^2 = t^2(t - \langle \mathbf{u}, \mathbf{v} \rangle)
\end{aligned}
$$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2 - t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\begin{aligned} \det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \\ \det(tI - P) &= t^3 - (u_1v_1 + u_2v_2 + u_3v_3)t^2 = t^2(t - \langle \mathbf{u}, \mathbf{v} \rangle) \end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u}, \mathbf{v} \rangle$.

Eigenvalue examples

$$\det \begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det \begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det \begin{bmatrix} u_1v_2 & u_2v_2 - t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\begin{aligned} \det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \\ \det(tI - P) &= t^3 - (u_1v_1 + u_2v_2 + u_3v_3)t^2 = t^2(t - \langle \mathbf{u}, \mathbf{v} \rangle) \end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u}, \mathbf{v} \rangle$. The eigenvectors of eigenvalue 0 are the vectors orthogonal to $\mathbf{u}$.

Eigenvalue examples

$$\det \begin{bmatrix} u_2v_2 - t & u_3v_2 \\ u_2v_3 & u_3v_3 - t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det \begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3 - t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det \begin{bmatrix} u_1v_2 & u_2v_2 - t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\begin{aligned} \det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \\ \det(tI - P) &= t^3 - (u_1v_1 + u_2v_2 + u_3v_3)t^2 = t^2(t - \langle \mathbf{u}, \mathbf{v} \rangle) \end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u}, \mathbf{v} \rangle$. The eigenvectors of eigenvalue 0 are the vectors orthogonal to $\mathbf{u}$. The eigenvectors of eigenvalue $\langle \mathbf{u}, \mathbf{v} \rangle$ are the multiples of $\mathbf{v}$.

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\begin{aligned} \det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \\ \det(tI - P) &= t^3 - (u_1v_1 + u_2v_2 + u_3v_3)t^2 = t^2(t - \langle \mathbf{u}, \mathbf{v} \rangle) \end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u}, \mathbf{v} \rangle$. The eigenvectors of eigenvalue 0 are the vectors orthogonal to $\mathbf{u}$. The eigenvectors of eigenvalue $\langle \mathbf{u}, \mathbf{v} \rangle$ are the multiples of $\mathbf{v}$.

If we had noticed this in advance then the whole argument would have been much easier.

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2-t)(u_3v_3-t)-u_2v_3u_3v_2 = t^2-(u_2v_2+u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3-t)-u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3-u_1v_3(u_2v_2-t) = u_1v_3t$$

$$\begin{aligned} \det(P-tI) &= (u_1v_1-t)(t^2-(u_2v_2+u_3v_3)t)-u_2v_1(-u_1v_2t)+u_3v_1u_1v_3t \\ &= (u_1v_1+u_2v_2+u_3v_3)t^2-t^3 \\ \det(tI-P) &= t^3-(u_1v_1+u_2v_2+u_3v_3)t^2 = t^2(t-\langle \mathbf{u},\mathbf{v}\rangle) \end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u},\mathbf{v}\rangle$. The eigenvectors of eigenvalue 0 are the vectors orthogonal to $\mathbf{u}$. The eigenvectors of eigenvalue $\langle \mathbf{u},\mathbf{v}\rangle$ are the multiples of $\mathbf{v}$.

If we had noticed this in advance then the whole argument would have been much easier. We could have chosen a basis of the form $\mathbf{a},\mathbf{b},\mathbf{v}$ with $\mathbf{a}$ and $\mathbf{b}$ orthogonal to $\mathbf{u}$.

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2 - t)(u_3v_3 - t) - u_2v_3u_3v_2 = t^2 - (u_2v_2 + u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3 - t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2 - t) = u_1v_3t$$

$$\begin{aligned}
\det(P - tI) &= (u_1v_1 - t)(t^2 - (u_2v_2 + u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\
&= (u_1v_1 + u_2v_2 + u_3v_3)t^2 - t^3 \\
\det(tI - P) &= t^3 - (u_1v_1 + u_2v_2 + u_3v_3)t^2 = t^2(t - \langle \mathbf{u}, \mathbf{v}\rangle)
\end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u}, \mathbf{v}\rangle$. The eigenvectors of eigenvalue 0 are the vectors orthogonal to $\mathbf{u}$. The eigenvectors of eigenvalue $\langle \mathbf{u}, \mathbf{v}\rangle$ are the multiples of $\mathbf{v}$.

If we had noticed this in advance then the whole argument would have been much easier. We could have chosen a basis of the form $\mathbf{a}, \mathbf{b}, \mathbf{v}$ with $\mathbf{a}$ and $\mathbf{b}$ orthogonal to $\mathbf{u}$. With respect to that basis, ϕ would have matrix $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \langle \mathbf{u}, \mathbf{v}\rangle \end{bmatrix}$

Eigenvalue examples

$$\det\begin{bmatrix} u_2v_2-t & u_3v_2 \\ u_2v_3 & u_3v_3-t \end{bmatrix} = (u_2v_2-t)(u_3v_3-t) - u_2v_3u_3v_2 = t^2 - (u_2v_2+u_3v_3)t$$

$$\det\begin{bmatrix} u_1v_2 & u_3v_2 \\ u_1v_3 & u_3v_3-t \end{bmatrix} = u_1v_2(u_3v_3-t) - u_1v_3u_3v_2 = -u_1v_2t$$

$$\det\begin{bmatrix} u_1v_2 & u_2v_2-t \\ u_1v_3 & u_2v_3 \end{bmatrix} = u_1v_2u_2v_3 - u_1v_3(u_2v_2-t) = u_1v_3t$$

$$\begin{aligned}\det(P-tI) &= (u_1v_1-t)(t^2-(u_2v_2+u_3v_3)t) - u_2v_1(-u_1v_2t) + u_3v_1u_1v_3t \\ &= (u_1v_1+u_2v_2+u_3v_3)t^2 - t^3 \\ \det(tI-P) &= t^3 - (u_1v_1+u_2v_2+u_3v_3)t^2 = t^2(t-\langle \mathbf{u},\mathbf{v}\rangle)\end{aligned}$$

The eigenvalues are thus 0 and $\langle \mathbf{u},\mathbf{v}\rangle$. The eigenvectors of eigenvalue 0 are the vectors orthogonal to $\mathbf{u}$. The eigenvectors of eigenvalue $\langle \mathbf{u},\mathbf{v}\rangle$ are the multiples of $\mathbf{v}$.

If we had noticed this in advance then the whole argument would have been much easier. We could have chosen a basis of the form $\mathbf{a},\mathbf{b},\mathbf{v}$ with $\mathbf{a}$ and $\mathbf{b}$ orthogonal to $\mathbf{u}$. With respect to that basis, ϕ would have matrix $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \langle \mathbf{u},\mathbf{v}\rangle \end{bmatrix}$ which immediately gives the characteristic polynomial.

Inner products

Inner products

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

Inner products

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{R}$.

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{R}$.

(c) $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

Inner products

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{R}$.

(c) $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

(d) $\langle u, u\rangle \geq 0$ for all $u \in V$, and $\langle u, u\rangle = 0$ iff $u = 0$.

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{R}$.

(c) $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

(d) $\langle u, u\rangle \geq 0$ for all $u \in V$, and $\langle u, u\rangle = 0$ iff $u = 0$.

Given an inner product, we write $\|u\| = \sqrt{\langle u, u\rangle}$, and call this the *norm* of u.

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{R}$.

(c) $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

(d) $\langle u, u\rangle \geq 0$ for all $u \in V$, and $\langle u, u\rangle = 0$ iff $u = 0$.

Given an inner product, we write $\|u\| = \sqrt{\langle u, u\rangle}$, and call this the *norm* of u.

We say that u is a *unit vector* if $\|u\| = 1$.

Definition ??: Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{R}$ for each $u, v \in V$, with the following properties:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{R}$.

(c) $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

(d) $\langle u, u\rangle \geq 0$ for all $u \in V$, and $\langle u, u\rangle = 0$ iff $u = 0$.

Given an inner product, we write $\|u\| = \sqrt{\langle u, u\rangle}$, and call this the *norm* of u.

We say that u is a *unit vector* if $\|u\| = 1$.

We say that u and v are *orthogonal* if $\langle u, v\rangle = 0$.

Other fields

Remark ??:

Other fields

Remark ??: Unlike most of the other things we have done, this does not immediately generalise to fields K other than $\mathbb{R}$.

Remark ??: Unlike most of the other things we have done, this does not immediately generalise to fields K other than $\mathbb{R}$. The reason is that axiom (d) involves the condition $\langle u, u \rangle \geq 0$

Other fields

Remark ??: Unlike most of the other things we have done, this does not immediately generalise to fields K other than $\mathbb{R}$. The reason is that axiom (d) involves the condition $\langle u, u \rangle \geq 0$, and in an arbitrary field K (such as $\mathbb{Z}/5$, for example) we do not have a good notion of positivity.

Other fields

Remark ??: Unlike most of the other things we have done, this does not immediately generalise to fields K other than $\mathbb{R}$. The reason is that axiom (d) involves the condition $\langle u, u \rangle \geq 0$, and in an arbitrary field K (such as $\mathbb{Z}/5$, for example) we do not have a good notion of positivity.

Moreover, all our examples will rely heavily on the fact that $x^2 \geq 0$ for all $x \in \mathbb{R}$

Other fields

Remark ??: Unlike most of the other things we have done, this does not immediately generalise to fields K other than $\mathbb{R}$. The reason is that axiom (d) involves the condition $\langle u, u\rangle \geq 0$, and in an arbitrary field K (such as $\mathbb{Z}/5$, for example) we do not have a good notion of positivity.

Moreover, all our examples will rely heavily on the fact that $x^2 \geq 0$ for all $x \in \mathbb{R}$, and of course this ceases to be true if we work over $\mathbb{C}$.

Other fields

Remark ??: Unlike most of the other things we have done, this does not immediately generalise to fields K other than $\mathbb{R}$. The reason is that axiom (d) involves the condition $\langle u, u \rangle \geq 0$, and in an arbitrary field K (such as $\mathbb{Z}/5$, for example) we do not have a good notion of positivity.

Moreover, all our examples will rely heavily on the fact that $x^2 \geq 0$ for all $x \in \mathbb{R}$, and of course this ceases to be true if we work over $\mathbb{C}$. We will see in Section **??** how to fix things up in the complex case.

The standard inner product on $\mathbb{R}^n$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$

(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

The standard inner product on $\mathbb{R}^n$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

The standard inner product on $\mathbb{R}^n$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Properties (a) to (c) are obvious.

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Properties (a) to (c) are obvious. For property (d), note that if $\mathbf{u} = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$ then

$$\langle \mathbf{u}, \mathbf{u}\rangle = u_1^2 + u_2^2 + \cdots + u_n^2.$$

The standard inner product on $\mathbb{R}^n$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Properties (a) to (c) are obvious. For property (d), note that if $\mathbf{u} = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$ then

$$\langle \mathbf{u}, \mathbf{u} \rangle = u_1^2 + u_2^2 + \cdots + u_n^2.$$

All the terms in this sum are at least zero, so the sum must be at least zero.

The standard inner product on $\mathbb{R}^n$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Properties (a) to (c) are obvious. For property (d), note that if $\mathbf{u} = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$ then

$$\langle \mathbf{u}, \mathbf{u}\rangle = u_1^2 + u_2^2 + \cdots + u_n^2.$$

All the terms in this sum are at least zero, so the sum must be at least zero. Moreover, there can be no cancellation, so the only way that $\langle \mathbf{u}, \mathbf{u}\rangle$ can be zero is if all the individual terms are zero

The standard inner product on $\mathbb{R}^n$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Properties (a) to (c) are obvious. For property (d), note that if $\mathbf{u} = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$ then

$$\langle \mathbf{u}, \mathbf{u}\rangle = u_1^2 + u_2^2 + \cdots + u_n^2.$$

All the terms in this sum are at least zero, so the sum must be at least zero. Moreover, there can be no cancellation, so the only way that $\langle \mathbf{u}, \mathbf{u}\rangle$ can be zero is if all the individual terms are zero, which means $u_1 = u_2 = \cdots = u_n = 0$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $\mathbb{R}^n$ by

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Properties (a) to (c) are obvious. For property (d), note that if $\mathbf{u} = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$ then

$$\langle \mathbf{u}, \mathbf{u}\rangle = u_1^2 + u_2^2 + \cdots + u_n^2.$$

All the terms in this sum are at least zero, so the sum must be at least zero. Moreover, there can be no cancellation, so the only way that $\langle \mathbf{u}, \mathbf{u}\rangle$ can be zero is if all the individual terms are zero, which means $u_1 = u_2 = \cdots = u_n = 0$, so $\mathbf{u} = 0$ as a vector.

Inner products and matrix multiplication

Inner products and matrix multiplication

Remark ??: If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then we can regard $\mathbf{x}$ and $\mathbf{y}$ as $n \times 1$ matrices

Remark ??: If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then we can regard $\mathbf{x}$ and $\mathbf{y}$ as $n \times 1$ matrices, so $\mathbf{x}^T$ is a $1 \times n$ matrix

Inner products and matrix multiplication

Remark ??: If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then we can regard $\mathbf{x}$ and $\mathbf{y}$ as $n \times 1$ matrices, so $\mathbf{x}^T$ is a $1 \times n$ matrix, so $\mathbf{x}^T\mathbf{y}$ is a 1×1 matrix

Remark ??: If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then we can regard $\mathbf{x}$ and $\mathbf{y}$ as $n \times 1$ matrices, so $\mathbf{x}^T$ is a $1 \times n$ matrix, so $\mathbf{x}^T\mathbf{y}$ is a 1×1 matrix, or in other words a number.

Inner products and matrix multiplication

Remark ??: If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then we can regard $\mathbf{x}$ and $\mathbf{y}$ as $n \times 1$ matrices, so $\mathbf{x}^T$ is a $1 \times n$ matrix, so $\mathbf{x}^T\mathbf{y}$ is a 1×1 matrix, or in other words a number. This number is just $\langle \mathbf{x}, \mathbf{y} \rangle$.

Remark ??: If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then we can regard $\mathbf{x}$ and $\mathbf{y}$ as $n \times 1$ matrices, so $\mathbf{x}^T$ is a $1 \times n$ matrix, so $\mathbf{x}^T\mathbf{y}$ is a 1×1 matrix, or in other words a number. This number is just $\langle \mathbf{x}, \mathbf{y} \rangle$. This is most easily explained by example: in the case $n = 4$ we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}^T \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Inner products of physical vectors

Example ??: Let U be the set of physical vectors, as in Example **??**.

Inner products of physical vectors

Example ??: Let U be the set of physical vectors, as in Example **??**. Given $\mathbf{u}, \mathbf{v} \in U$ we can define

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle = &(\text{length of } \mathbf{u} \text{ in miles}) \times (\text{length of } \mathbf{v} \text{ in miles}) \times \\ &\cos(\text{ angle between } \mathbf{u} \text{ and } \mathbf{v} \text{ }). \end{aligned}$$

Inner products of physical vectors

Example ??: Let U be the set of physical vectors, as in Example **??**. Given $\mathbf{u}, \mathbf{v} \in U$ we can define

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle =& (\text{length of } \mathbf{u} \text{ in miles}) \times (\text{length of } \mathbf{v} \text{ in miles}) \times \\ & \cos(\text{ angle between } \mathbf{u} \text{ and } \mathbf{v} \text{ }).\end{aligned}$$

This turns out to give an inner product on U.

Example ??: Let U be the set of physical vectors, as in Example **??**. Given $\mathbf{u}, \mathbf{v} \in U$ we can define

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle =& (\text{length of } \mathbf{u} \text{ in miles}) \times (\text{length of } \mathbf{v} \text{ in miles}) \times \\ & \cos(\text{ angle between } \mathbf{u} \text{ and } \mathbf{v} \text{ }).\end{aligned}$$

This turns out to give an inner product on U. Of course we could use a different unit of length instead of miles, and that would just change the inner product by a constant factor.

Inner products of functions

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $C[0,1]$ by

$$\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx.$$

Inner products of functions

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $C[0,1]$ by

$$\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx.$$

Properties (a) to (c) are obvious.

Inner products of functions

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $C[0,1]$ by

$$\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx.$$

Properties (a) to (c) are obvious. For property (d), note that if $u \in C[0,1]$ then

$$\langle u, u\rangle = \int_0^1 u(x)^2\,dx$$

Inner products of functions

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on $C[0,1]$ by

$$\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx.$$

Properties (a) to (c) are obvious. For property (d), note that if $u \in C[0,1]$ then

$$\langle u, u\rangle = \int_0^1 u(x)^2\,dx$$

As $u(x)^2 \geq 0$ for all x, we have $\langle u, u\rangle \geq 0$.

Inner products of functions

(a) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (b) $\langle tu,v\rangle = t\langle u,v\rangle$
(c) $\langle u,v\rangle = \langle v,u\rangle$ (d) $\langle u,u\rangle \geq 0$, equality iff $u=0$

Example ??: We can define an inner product on $C[0,1]$ by

$$\langle u,v\rangle = \int_0^1 u(x)v(x)\,dx.$$

Properties (a) to (c) are obvious. For property (d), note that if $u \in C[0,1]$ then

$$\langle u,u\rangle = \int_0^1 u(x)^2\,dx$$

As $u(x)^2 \geq 0$ for all x, we have $\langle u,u\rangle \geq 0$.

If $\langle u,u\rangle = 0$ then the area between the x-axis and the graph of $u(x)^2$ is zero, so $u(x)^2$ must be zero for all x, so $u=0$ as required.

(a) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (b) $\langle tu,v\rangle = t\langle u,v\rangle$
(c) $\langle u,v\rangle = \langle v,u\rangle$ (d) $\langle u,u\rangle \geq 0$, equality iff $u=0$

Example ??: We can define an inner product on $C[0,1]$ by

$$\langle u,v\rangle = \int_0^1 u(x)v(x)\,dx.$$

Properties (a) to (c) are obvious. For property (d), note that if $u \in C[0,1]$ then

$$\langle u,u\rangle = \int_0^1 u(x)^2\,dx$$

As $u(x)^2 \geq 0$ for all x, we have $\langle u,u\rangle \geq 0$.

If $\langle u,u\rangle = 0$ then the area between the x-axis and the graph of $u(x)^2$ is zero, so $u(x)^2$ must be zero for all x, so $u=0$ as required.

(There is a more careful proof in the notes.)

Inner products of matrices

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A, B\rangle = \text{trace}(AB^T)$.

(a) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (b) $\langle tu,v\rangle = t\langle u,v\rangle$
(c) $\langle u,v\rangle = \langle v,u\rangle$ (d) $\langle u,u\rangle \geq 0$, equality iff $u=0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A,B\rangle = \text{trace}(AB^T)$.

Consider for example the case $n=3$, so $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$

Inner products of matrices

(a) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (b) $\langle tu,v\rangle = t\langle u,v\rangle$
(c) $\langle u,v\rangle = \langle v,u\rangle$ (d) $\langle u,u\rangle \geq 0$, equality iff $u=0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A,B\rangle = \mathrm{trace}(AB^T)$.

Consider for example the case $n=3$, so $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$

so

$$AB^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{12}+a_{13}b_{13} & a_{11}b_{21}+a_{12}b_{22}+a_{13}b_{23} & a_{11}b_{31}+a_{12}b_{32}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{12}+a_{23}b_{13} & a_{21}b_{21}+a_{22}b_{22}+a_{23}b_{23} & a_{21}b_{31}+a_{22}b_{32}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{12}+a_{33}b_{13} & a_{31}b_{21}+a_{32}b_{22}+a_{33}b_{23} & a_{31}b_{31}+a_{32}b_{32}+a_{33}b_{33} \end{bmatrix}$$

Inner products of matrices

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A, B\rangle = \text{trace}(AB^T)$.

Consider for example the case $n = 3$, so $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$

so

$$AB^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{12}+a_{13}b_{13} & a_{11}b_{21}+a_{12}b_{22}+a_{13}b_{23} & a_{11}b_{31}+a_{12}b_{32}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{12}+a_{23}b_{13} & a_{21}b_{21}+a_{22}b_{22}+a_{23}b_{23} & a_{21}b_{31}+a_{22}b_{32}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{12}+a_{33}b_{13} & a_{31}b_{21}+a_{32}b_{22}+a_{33}b_{23} & a_{31}b_{31}+a_{32}b_{32}+a_{33}b_{33} \end{bmatrix}$$

$$\text{so } \langle A, B\rangle = \text{trace}(AB^T) = \begin{array}{c} a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + \\ a_{21}b_{21} + a_{22}b_{22} + a_{23}b_{23} + \\ a_{31}b_{31} + a_{32}b_{32} + a_{33}b_{33} \end{array}$$

Inner products of matrices

(a) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (b) $\langle tu, v\rangle = t\langle u, v\rangle$
(c) $\langle u, v\rangle = \langle v, u\rangle$ (d) $\langle u, u\rangle \geq 0$, equality iff $u = 0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A, B\rangle = \mathrm{trace}(AB^T)$.

Consider for example the case $n = 3$, so $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$

so

$$AB^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{12}+a_{13}b_{13} & a_{11}b_{21}+a_{12}b_{22}+a_{13}b_{23} & a_{11}b_{31}+a_{12}b_{32}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{12}+a_{23}b_{13} & a_{21}b_{21}+a_{22}b_{22}+a_{23}b_{23} & a_{21}b_{31}+a_{22}b_{32}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{12}+a_{33}b_{13} & a_{31}b_{21}+a_{32}b_{22}+a_{33}b_{23} & a_{31}b_{31}+a_{32}b_{32}+a_{33}b_{33} \end{bmatrix}$$

$$\text{so } \langle A, B\rangle = \mathrm{trace}(AB^T) = \begin{array}{c} a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + \\ a_{21}b_{21} + a_{22}b_{22} + a_{23}b_{23} + \\ a_{31}b_{31} + a_{32}b_{32} + a_{33}b_{33} \end{array} = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}b_{ij}.$$

Inner products of matrices

(a) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (b) $\langle tu,v\rangle = t\langle u,v\rangle$
(c) $\langle u,v\rangle = \langle v,u\rangle$ (d) $\langle u,u\rangle \geq 0$, equality iff $u=0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A,B\rangle = \mathrm{trace}(AB^T)$.

Consider for example the case $n=3$, so $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$

so

$$AB^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{12}+a_{13}b_{13} & a_{11}b_{21}+a_{12}b_{22}+a_{13}b_{23} & a_{11}b_{31}+a_{12}b_{32}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{12}+a_{23}b_{13} & a_{21}b_{21}+a_{22}b_{22}+a_{23}b_{23} & a_{21}b_{31}+a_{22}b_{32}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{12}+a_{33}b_{13} & a_{31}b_{21}+a_{32}b_{22}+a_{33}b_{23} & a_{31}b_{31}+a_{32}b_{32}+a_{33}b_{33} \end{bmatrix}$$

$$\text{so } \langle A,B\rangle = \mathrm{trace}(AB^T) = \begin{matrix} a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + \\ a_{21}b_{21} + a_{22}b_{22} + a_{23}b_{23} + \\ a_{31}b_{31} + a_{32}b_{32} + a_{33}b_{33} \end{matrix} = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}b_{ij}.$$

In other words $\langle A,B\rangle$ is the sum of the entries of A multiplied by the corresponding entries in B.

Inner products of matrices

(a) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$ (b) $\langle tu,v\rangle = t\langle u,v\rangle$
(c) $\langle u,v\rangle = \langle v,u\rangle$ (d) $\langle u,u\rangle \geq 0$, equality iff $u=0$

Example ??: We can define an inner product on the space $M_n\mathbb{R}$ by $\langle A,B\rangle = \mathrm{trace}(AB^T)$.

Consider for example the case $n=3$, so $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$

so

$$AB^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{12}+a_{13}b_{13} & a_{11}b_{21}+a_{12}b_{22}+a_{13}b_{23} & a_{11}b_{31}+a_{12}b_{32}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{12}+a_{23}b_{13} & a_{21}b_{21}+a_{22}b_{22}+a_{23}b_{23} & a_{21}b_{31}+a_{22}b_{32}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{12}+a_{33}b_{13} & a_{31}b_{21}+a_{32}b_{22}+a_{33}b_{23} & a_{31}b_{31}+a_{32}b_{32}+a_{33}b_{33} \end{bmatrix}$$

$$\text{so } \langle A,B\rangle = \mathrm{trace}(AB^T) = \begin{matrix} a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + \\ a_{21}b_{21} + a_{22}b_{22} + a_{23}b_{23} + \\ a_{31}b_{31} + a_{32}b_{32} + a_{33}b_{33} \end{matrix} = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}b_{ij}.$$

In other words $\langle A,B\rangle$ is the sum of the entries of A multiplied by the corresponding entries in B. Thus, if we identify $M_n\mathbb{R}$ with $\mathbb{R}^{n^2}$, our inner product on $M_n\mathbb{R}$ corresponds to the standard inner product on $\mathbb{R}^{n^2}$.

Inner products for quadratic polynomials

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\,dx.$$

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\, dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\, dx$$

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\,dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\,dx = \left[\frac{x^{i+j+1}}{i+j+1} \right]_a^b$$

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\, dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\, dx = \left[\frac{x^{i+j+1}}{i+j+1} \right]_a^b = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}$$

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\, dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\, dx = \left[\frac{x^{i+j+1}}{i+j+1} \right]_a^b = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}$$

This gives an infinite family of different inner products on $\mathbb{R}[x]_{\leq 2}$.

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\, dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\, dx = \left[\frac{x^{i+j+1}}{i+j+1} \right]_a^b = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}$$

This gives an infinite family of different inner products on $\mathbb{R}[x]_{\leq 2}$.

$$\langle 1, x^2 \rangle_{[-1,1]} = \frac{1^3 - (-1)^3}{3} = 2/3$$

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\,dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\,dx = \left[\frac{x^{i+j+1}}{i+j+1}\right]_a^b = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}$$

This gives an infinite family of different inner products on $\mathbb{R}[x]_{\leq 2}$.

$$\langle 1, x^2 \rangle_{[-1,1]} = \frac{1^3 - (-1)^3}{3} = 2/3$$

$$\langle x, x^2 \rangle_{[-1,1]} = \frac{1^4 - (-1)^4}{4} = 0$$

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\, dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\, dx = \left[\frac{x^{i+j+1}}{i+j+1} \right]_a^b = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}$$

This gives an infinite family of different inner products on $\mathbb{R}[x]_{\leq 2}$.

$$\langle 1, x^2 \rangle_{[-1,1]} = \frac{1^3 - (-1)^3}{3} = 2/3$$

$$\langle x, x^2 \rangle_{[-1,1]} = \frac{1^4 - (-1)^4}{4} = 0$$

$$\|x^2\|_{[-1,1]} = \sqrt{\frac{1^5 - (-1)^5}{5}} = \sqrt{2/5}$$

Inner products for quadratic polynomials

Example ??: For any $a < b$ we can define an inner product $\langle \cdot, \cdot \rangle_{[a,b]}$ on $\mathbb{R}[x]_{\leq 2}$ by

$$\langle u, v \rangle_{[a,b]} = \int_a^b u(x)v(x)\,dx.$$

In particular, we have

$$\langle x^i, x^j \rangle_{[a,b]} = \int_a^b x^{i+j}\,dx = \left[\frac{x^{i+j+1}}{i+j+1}\right]_a^b = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}$$

This gives an infinite family of different inner products on $\mathbb{R}[x]_{\leq 2}$.

$$\langle 1, x^2 \rangle_{[-1,1]} = \frac{1^3 - (-1)^3}{3} = 2/3$$

$$\langle x, x^2 \rangle_{[-1,1]} = \frac{1^4 - (-1)^4}{4} = 0$$

$$\|x^2\|_{[-1,1]} = \sqrt{\frac{1^5 - (-1)^5}{5}} = \sqrt{2/5}$$

$$\|x^2\|_{[0,5]} = \sqrt{\frac{5^5 - 0^5}{5}} = 25$$

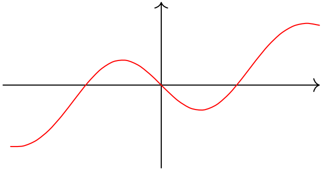
Another example

Another example

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial.

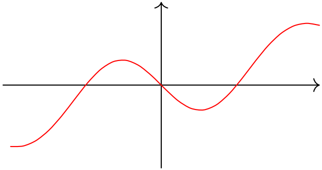
Another example

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. For example, the function $f(x) = (x^3 - x)e^{-x^2/2}$, shown in the graph below, is an element of V:



Another example

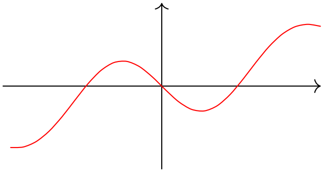
Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. For example, the function $f(x) = (x^3 - x)e^{-x^2/2}$, shown in the graph below, is an element of V:



We can define an inner product on V by $\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$

Another example

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. For example, the function $f(x) = (x^3 - x)e^{-x^2/2}$, shown in the graph below, is an element of V:

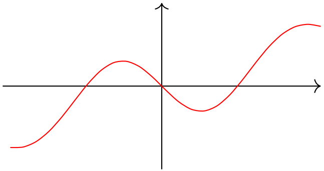


We can define an inner product on V by $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$

Note that this only works because of the special form of the functions in V.

Another example

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. For example, the function $f(x) = (x^3 - x)e^{-x^2/2}$, shown in the graph below, is an element of V:

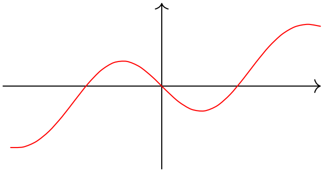


We can define an inner product on V by $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$

Note that this only works because of the special form of the functions in V. For most functions f and g that you might think of, the integral $\int_{-\infty}^{\infty} f(x)g(x)\,dx$ will give an infinite or undefined answer.

Another example

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. For example, the function $f(x) = (x^3 - x)e^{-x^2/2}$, shown in the graph below, is an element of V:



We can define an inner product on V by $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$

Note that this only works because of the special form of the functions in V. For most functions f and g that you might think of, the integral $\int_{-\infty}^{\infty} f(x)g(x)\,dx$ will give an infinite or undefined answer. However, the function e^{-x^2} decays very rapidly to zero as $|x|$ tends to infinity, and one can check that this is enough to make the integral well-defined and finite when f and g are in V.

In fact, we have the formula

$$\langle x^n e^{-x^2/2}, x^m e^{-x^2/2} \rangle = \int_{-\infty}^{\infty} x^{n+m} e^{-x^2}\, dx$$

$$= \begin{cases} \frac{\sqrt{\pi}}{2^{n+m}} \frac{(n+m)!}{((n+m)/2)!} & \text{if } n+m \text{ is even} \\ 0 & \text{if } n+m \text{ is odd} \end{cases}$$

The Cauchy-Schwartz inequality

The Cauchy-Schwartz inequality

If **v** and **w** are vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, you should be familiar with the fact that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \, \|\mathbf{w}\| \, \cos(\theta),$$

where θ is the angle between **v** and **w**.

The Cauchy-Schwartz inequality

If $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, you should be familiar with the fact that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \, \|\mathbf{w}\| \, \cos(\theta),$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

In particular, as the cosine lies between -1 and 1, we see that $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \, \|\mathbf{w}\|$.

The Cauchy-Schwartz inequality

If $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, you should be familiar with the fact that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos(\theta),$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

In particular, as the cosine lies between -1 and 1, we see that $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \, \|\mathbf{w}\|$.

We would like to extend all this to arbitrary inner-product spaces.

The Cauchy-Schwartz inequality

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V.

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V.
Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent.

Proof:

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have
$0 \leq \|sv + tw\|^2$

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have
$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle$

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have
$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle$

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have

$$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle = s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2.$$

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have

$$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle = s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2.$$

Now take $s = \langle w, w\rangle = \|w\|^2$ and $t = -\langle v, w\rangle$.

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have

$$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle =$$
$$s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2.$$

Now take $s = \langle w, w\rangle = \|w\|^2$ and $t = -\langle v, w\rangle$. The above inequality gives

$$0 \leq \|w\|^4\|v\|^2 - 2\|w\|^2\langle v, w\rangle^2 + \langle v, w\rangle^2\|w\|^2$$

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have

$$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle =$$
$$s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2.$$

Now take $s = \langle w, w\rangle = \|w\|^2$ and $t = -\langle v, w\rangle$. The above inequality gives

$$0 \leq \|w\|^4\|v\|^2 - 2\|w\|^2\langle v, w\rangle^2 + \langle v, w\rangle^2\|w\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2).$$

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have

$$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle = s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2.$$

Now take $s = \langle w, w\rangle = \|w\|^2$ and $t = -\langle v, w\rangle$. The above inequality gives

$$0 \leq \|w\|^4\|v\|^2 - 2\|w\|^2\langle v, w\rangle^2 + \langle v, w\rangle^2\|w\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2).$$

We have assumed that $w \neq 0$, so $\|w\|^2 > 0$.

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have
$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle =$
$s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2.$

Now take $s = \langle w, w\rangle = \|w\|^2$ and $t = -\langle v, w\rangle$. The above inequality gives

$$0 \leq \|w\|^4\|v\|^2 - 2\|w\|^2\langle v, w\rangle^2 + \langle v, w\rangle^2\|w\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2).$$

We have assumed that $w \neq 0$, so $\|w\|^2 > 0$. We can thus divide by $\|w\|^2$ and rearrange to see that $\langle v, w\rangle^2 \leq \|v\|^2\|w\|^2$.

The Cauchy-Schwartz inequality

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V. Then $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent.

Proof: We may assume $w \neq 0$ (otherwise everything is trivial).

For any real numbers s and t, we have
$0 \leq \|sv + tw\|^2 = \langle sv + tw, sv + tw\rangle = s^2\langle v, v\rangle + st\langle v, w\rangle + st\langle w, v\rangle + t^2\langle w, w\rangle =$
$s^2\|v\|^2 + 2st\langle v, w\rangle + t^2\|w\|^2$.

Now take $s = \langle w, w\rangle = \|w\|^2$ and $t = -\langle v, w\rangle$. The above inequality gives

$$0 \leq \|w\|^4\|v\|^2 - 2\|w\|^2\langle v, w\rangle^2 + \langle v, w\rangle^2\|w\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2).$$

We have assumed that $w \neq 0$, so $\|w\|^2 > 0$. We can thus divide by $\|w\|^2$ and rearrange to see that $\langle v, w\rangle^2 \leq \|v\|^2\|w\|^2$. It follows that $|\langle v, w\rangle| \leq \|v\|\|w\|$ as claimed.

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w \rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w \rangle^2)$$

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$.

The Cauchy-Schwartz inequality

$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent.

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent. As $w \neq 0$, this means that $v = \lambda w$ for some $\lambda \in \mathbb{R}$.

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent. As $w \neq 0$, this means that $v = \lambda w$ for some $\lambda \in \mathbb{R}$. It follows that
$\langle v, w\rangle = \lambda\|w\|^2$

The Cauchy-Schwartz inequality

$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent. As $w \neq 0$, this means that $v = \lambda w$ for some $\lambda \in \mathbb{R}$. It follows that $\langle v, w\rangle = \lambda\|w\|^2$, so $|\langle v, w\rangle\| = |\lambda|\|w\|^2$.

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent. As $w \neq 0$, this means that $v = \lambda w$ for some $\lambda \in \mathbb{R}$. It follows that $\langle v, w\rangle = \lambda\|w\|^2$, so $|\langle v, w\rangle\| = |\lambda|\|w\|^2$. On the other hand, we have $\|v\| = |\lambda|\|w\|$

The Cauchy-Schwartz inequality

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent. As $w \neq 0$, this means that $v = \lambda w$ for some $\lambda \in \mathbb{R}$. It follows that $\langle v, w\rangle = \lambda\|w\|^2$, so $|\langle v, w\rangle\| = |\lambda|\|w\|^2$. On the other hand, we have $\|v\| = |\lambda|\|w\|$, so $\|v\|\,\|w\| = |\lambda|\|w\|^2$

$$s = \|w\|^2 \qquad t = -\langle v, w\rangle \qquad \|sv + tw\|^2 = \|w\|^2(\|w\|^2\|v\|^2 - \langle v, w\rangle^2)$$

If we have equality (i.e. $|\langle v, w\rangle| = \|v\|\|w\|$) then our calculation shows that $\|sv + tw\|^2 = 0$, so $sv + tw = 0$. Here $s = \|w\|^2 > 0$, so we have a *nontrivial* linear relation between v and w, so they are linearly dependent.

Conversely, suppose we start by assuming that v and w are linearly dependent. As $w \neq 0$, this means that $v = \lambda w$ for some $\lambda \in \mathbb{R}$. It follows that $\langle v, w\rangle = \lambda\|w\|^2$, so $|\langle v, w\rangle\| = |\lambda|\|w\|^2$. On the other hand, we have $\|v\| = |\lambda|\|w\|$, so $\|v\|\,\|w\| = |\lambda|\|w\|^2$, which is the same.

Theorem ??:
Let V be an inner product space over $\mathbb{R}$, and let v and w be elements of V.
Then $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent.

An example of Cauchy-Schwartz

An example of Cauchy-Schwartz

Example ??:

An example of Cauchy-Schwartz

Example ??: We claim that for any vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$|x_1 + \cdots + x_n| \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}.$$

An example of Cauchy-Schwartz

Example ??: We claim that for any vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$|x_1 + \cdots + x_n| \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}.$$

To see this, use the standard inner product on $\mathbb{R}^n$, and consider the vector $\mathbf{e} = [1, 1, \ldots, 1]^T$.

An example of Cauchy-Schwartz

Example ??: We claim that for any vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$|x_1 + \cdots + x_n| \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}.$$

To see this, use the standard inner product on $\mathbb{R}^n$, and consider the vector $\mathbf{e} = [1, 1, \ldots, 1]^T$. We have

$$\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

An example of Cauchy-Schwartz

Example ??: We claim that for any vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$|x_1 + \cdots + x_n| \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}.$$

To see this, use the standard inner product on $\mathbb{R}^n$, and consider the vector $\mathbf{e} = [1, 1, \ldots, 1]^T$. We have

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{x_1^2 + \cdots + x_n^2} \\ \|\mathbf{e}\| &= \sqrt{n} \end{aligned}$$

An example of Cauchy-Schwartz

Example ??: We claim that for any vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$|x_1 + \cdots + x_n| \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}.$$

To see this, use the standard inner product on $\mathbb{R}^n$, and consider the vector $\mathbf{e} = [1, 1, \ldots, 1]^T$. We have

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{x_1^2 + \cdots + x_n^2} \\ \|\mathbf{e}\| &= \sqrt{n} \\ \langle \mathbf{x}, \mathbf{e} \rangle &= x_1 + \cdots + x_n. \end{aligned}$$

An example of Cauchy-Schwartz

Example ??: We claim that for any vector $\mathbf{x} \in \mathbb{R}^n$, we have

$$|x_1 + \cdots + x_n| \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}.$$

To see this, use the standard inner product on $\mathbb{R}^n$, and consider the vector $\mathbf{e} = [1, 1, \ldots, 1]^T$. We have

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{x_1^2 + \cdots + x_n^2} \\ \|\mathbf{e}\| &= \sqrt{n} \\ \langle \mathbf{x}, \mathbf{e} \rangle &= x_1 + \cdots + x_n. \end{aligned}$$

The Cauchy-Schwartz inequality therefore tells us that

$$\begin{aligned} |x_1 + \cdots + x_n| &= |\langle \mathbf{x}, \mathbf{e} \rangle| \\ &\leq \|\mathbf{x}\|\|\mathbf{e}\| = \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2}, \end{aligned}$$

as claimed.

An example with functions

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\,dx \right| \le \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\,dx}.$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\, dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v \rangle = \int_0^1 u(x)v(x)\, dx$.

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\, dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v \rangle = \int_0^1 u(x) v(x)\, dx$.
We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\, dx}$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\, dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v \rangle = \int_0^1 u(x) v(x)\, dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\, dx}$ and

$$\|1-x^2\|^2 = \langle 1-x^2, 1-x^2 \rangle$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\, dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v \rangle = \int_0^1 u(x)v(x)\, dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\, dx}$ and

$$\|1-x^2\|^2 = \langle 1-x^2, 1-x^2 \rangle = \int_0^1 1 - 2x^2 + x^4\, dx$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\,dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\,dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\,dx}$ and

$$\begin{aligned} \|1-x^2\|^2 = \langle 1-x^2, 1-x^2\rangle &= \int_0^1 1 - 2x^2 + x^4\,dx \\ &= \left[x - \tfrac{2}{3}x^3 + \tfrac{1}{5}x^5\right]_0^1 \end{aligned}$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left|\int_0^1 (1-x^2)f(x)\,dx\right| \leq \sqrt{\frac{8}{15}}\sqrt{\int_0^1 f(x)^2\,dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\,dx}$ and

$$\begin{aligned}\|1-x^2\|^2 = \langle 1-x^2, 1-x^2\rangle &= \int_0^1 1-2x^2+x^4\,dx \\ &= \left[x - \tfrac{2}{3}x^3 + \tfrac{1}{5}x^5\right]_0^1 = 1 - \tfrac{2}{3} + \tfrac{1}{5}\end{aligned}$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left|\int_0^1 (1-x^2)f(x)\,dx\right| \leq \sqrt{\frac{8}{15}}\sqrt{\int_0^1 f(x)^2\,dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\,dx}$ and

$$\begin{aligned}\|1-x^2\|^2 = \langle 1-x^2, 1-x^2\rangle &= \int_0^1 1-2x^2+x^4\,dx \\ &= \left[x - \tfrac{2}{3}x^3 + \tfrac{1}{5}x^5\right]_0^1 = 1 - \tfrac{2}{3} + \tfrac{1}{5} = 8/15\end{aligned}$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\, dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v \rangle = \int_0^1 u(x)v(x)\, dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\, dx}$ and

$$\begin{aligned}
\|1-x^2\|^2 &= \langle 1-x^2, 1-x^2 \rangle = \int_0^1 1 - 2x^2 + x^4\, dx \\
&= \left[x - \tfrac{2}{3}x^3 + \tfrac{1}{5}x^5 \right]_0^1 = 1 - \tfrac{2}{3} + \tfrac{1}{5} = 8/15 \\
\|1-x^2\| &= \sqrt{8/15}
\end{aligned}$$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\,dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\,dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v\rangle = \int_0^1 u(x)v(x)\,dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\,dx}$ and

$$\begin{aligned} \|1-x^2\|^2 &= \langle 1-x^2, 1-x^2\rangle = \int_0^1 1 - 2x^2 + x^4\,dx \\ &= \left[x - \tfrac{2}{3}x^3 + \tfrac{1}{5}x^5\right]_0^1 = 1 - \tfrac{2}{3} + \tfrac{1}{5} = 8/15 \\ \|1-x^2\| &= \sqrt{8/15} \end{aligned}$$

The Cauchy-Schwartz inequality tells us that $|\langle u, f\rangle| \leq \|u\|\,\|f\|$

An example with functions

We claim that for any continuous function $f\colon [0,1] \to \mathbb{R}$ we have

$$\left| \int_0^1 (1-x^2) f(x)\, dx \right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}.$$

Indeed, we can define an inner product on $C[0,1]$ by $\langle u, v\rangle = \int_0^1 u(x)v(x)\, dx$. We then have $\|f\| = \sqrt{\int_0^1 f(x)^2\, dx}$ and

$$\begin{aligned} \|1-x^2\|^2 &= \langle 1-x^2, 1-x^2\rangle = \int_0^1 1 - 2x^2 + x^4\, dx \\ &= \left[x - \tfrac{2}{3}x^3 + \tfrac{1}{5}x^5\right]_0^1 = 1 - \tfrac{2}{3} + \tfrac{1}{5} = 8/15 \\ \|1-x^2\| &= \sqrt{8/15} \end{aligned}$$

The Cauchy-Schwartz inequality tells us that $|\langle u, f\rangle| \leq \|u\|\,\|f\|$, so $\left|\int_0^1 (1-x^2) f(x)\, dx\right| \leq \sqrt{\frac{8}{15}} \sqrt{\int_0^1 f(x)^2\, dx}$ as claimed.

Examples with matrices

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$.

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\,\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\operatorname{trace}(A^2)| \leq \operatorname{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \operatorname{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\text{trace}(A)^2 \leq n\,\text{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\text{trace}(A^2)| \leq \text{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \text{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I\rangle| \leq \|A\|\|I\|$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\operatorname{trace}(A^2)| \leq \operatorname{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \operatorname{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I\rangle| \leq \|A\|\|I\|$, or equivalently

$$\langle A, I\rangle^2 \leq \|A\|^2\|I\|^2 = \operatorname{trace}(AA^T)\operatorname{trace}(II^T)$$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I\rangle| \leq \|A\|\|I\|$, or equivalently

$$\langle A, I\rangle^2 \leq \|A\|^2\|I\|^2 = \mathrm{trace}(AA^T)\,\mathrm{trace}(II^T)$$

Here $\langle A, I\rangle = \mathrm{trace}(A)$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\operatorname{trace}(A^2)| \leq \operatorname{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \operatorname{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I\rangle| \leq \|A\|\|I\|$, or equivalently

$$\langle A, I\rangle^2 \leq \|A\|^2\|I\|^2 = \operatorname{trace}(AA^T)\operatorname{trace}(II^T)$$

Here $\langle A, I\rangle = \operatorname{trace}(A)$ and $\operatorname{trace}(II^T) = \operatorname{trace}(I) = n$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\operatorname{trace}(A^2)| \leq \operatorname{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \operatorname{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I\rangle| \leq \|A\|\|I\|$, or equivalently

$$\langle A, I\rangle^2 \leq \|A\|^2\|I\|^2 = \operatorname{trace}(AA^T)\operatorname{trace}(II^T)$$

Here $\langle A, I\rangle = \operatorname{trace}(A)$ and $\operatorname{trace}(II^T) = \operatorname{trace}(I) = n$, so we get $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$ as claimed.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\operatorname{trace}(A^2)| \leq \operatorname{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B \rangle = \operatorname{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I \rangle| \leq \|A\| \|I\|$, or equivalently

$$\langle A, I \rangle^2 \leq \|A\|^2 \|I\|^2 = \operatorname{trace}(AA^T) \operatorname{trace}(II^T)$$

Here $\langle A, I \rangle = \operatorname{trace}(A)$ and $\operatorname{trace}(II^T) = \operatorname{trace}(I) = n$, so we get $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$ as claimed. This is an equality iff A and I are linearly dependent

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\operatorname{trace}(A^2)| \leq \operatorname{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B \rangle = \operatorname{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(a) Apply the inequality to A and I, giving $|\langle A, I \rangle| \leq \|A\| \|I\|$, or equivalently

$$\langle A, I \rangle^2 \leq \|A\|^2 \|I\|^2 = \operatorname{trace}(AA^T) \operatorname{trace}(II^T)$$

Here $\langle A, I \rangle = \operatorname{trace}(A)$ and $\operatorname{trace}(II^T) = \operatorname{trace}(I) = n$, so we get $\operatorname{trace}(A)^2 \leq n \operatorname{trace}(AA^T)$ as claimed. This is an equality iff A and I are linearly dependent, which means that A is a multiple of I.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\,\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\,\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\,\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$. The conclusion is that $|\mathrm{trace}(A^2)| \leq \sqrt{\mathrm{trace}(AA^T)}\sqrt{\mathrm{trace}(AA^T)}$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\text{trace}(A)^2 \leq n\,\text{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\,\text{trace}(A^2)| \leq \text{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \text{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\text{trace}(AA^T)}$ and $\langle A, A^T\rangle = \text{trace}(AA^{TT}) = \text{trace}(A^2)$. The conclusion is that $|\,\text{trace}(A^2)| \leq \sqrt{\text{trace}(AA^T)}\sqrt{\text{trace}(AA^T)}$, which gives $|\,\text{trace}(A^2)| \leq \text{trace}(AA^T)$.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$. The conclusion is that $|\mathrm{trace}(A^2)| \leq \sqrt{\mathrm{trace}(AA^T)}\sqrt{\mathrm{trace}(AA^T)}$, which gives $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$. This is an equality iff A^T is a multiple of A

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\,\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$. The conclusion is that $|\,\mathrm{trace}(A^2)| \leq \sqrt{\mathrm{trace}(AA^T)}\sqrt{\mathrm{trace}(AA^T)}$, which gives $|\,\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$. This is an equality iff A^T is a multiple of A, say $A^T = \lambda A$ for some λ.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$. The conclusion is that $|\mathrm{trace}(A^2)| \leq \sqrt{\mathrm{trace}(AA^T)}\sqrt{\mathrm{trace}(AA^T)}$, which gives $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$. This is an equality iff A^T is a multiple of A, say $A^T = \lambda A$ for some λ. This means that $A = A^{TT} = \lambda A^T = \lambda^2 A$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\text{trace}(A)^2 \leq n\,\text{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\text{trace}(A^2)| \leq \text{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \text{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\text{trace}(AA^T)}$ and $\langle A, A^T\rangle = \text{trace}(AA^{TT}) = \text{trace}(A^2)$. The conclusion is that $|\text{trace}(A^2)| \leq \sqrt{\text{trace}(AA^T)}\sqrt{\text{trace}(AA^T)}$, which gives $|\text{trace}(A^2)| \leq \text{trace}(AA^T)$. This is an equality iff A^T is a multiple of A, say $A^T = \lambda A$ for some λ. This means that $A = A^{TT} = \lambda A^T = \lambda^2 A$, and $A \neq 0$, so this means that $\lambda^2 = 1$

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$. The conclusion is that $|\mathrm{trace}(A^2)| \leq \sqrt{\mathrm{trace}(AA^T)}\sqrt{\mathrm{trace}(AA^T)}$, which gives $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$. This is an equality iff A^T is a multiple of A, say $A^T = \lambda A$ for some λ. This means that $A = A^{TT} = \lambda A^T = \lambda^2 A$, and $A \neq 0$, so this means that $\lambda^2 = 1$, or equivalently $\lambda = \pm 1$.

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\mathrm{trace}(A)^2 \leq n\,\mathrm{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \mathrm{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\mathrm{trace}(AA^T)}$ and $\langle A, A^T\rangle = \mathrm{trace}(AA^{TT}) = \mathrm{trace}(A^2)$. The conclusion is that $|\mathrm{trace}(A^2)| \leq \sqrt{\mathrm{trace}(AA^T)}\sqrt{\mathrm{trace}(AA^T)}$, which gives $|\mathrm{trace}(A^2)| \leq \mathrm{trace}(AA^T)$. This is an equality iff A^T is a multiple of A, say $A^T = \lambda A$ for some λ. This means that $A = A^{TT} = \lambda A^T = \lambda^2 A$, and $A \neq 0$, so this means that $\lambda^2 = 1$, or equivalently $\lambda = \pm 1$. If $\lambda = 1$ then $A^T = A$ and A is symmetric

Examples with matrices

Let A be a nonzero $n \times n$ matrix over $\mathbb{R}$. We claim that

(a) $\text{trace}(A)^2 \leq n\,\text{trace}(AA^T)$, with equality iff A is a multiple of the identity.

(b) $|\text{trace}(A^2)| \leq \text{trace}(AA^T)$, with equality iff A is either symmetric or antisymmetric.

In both cases we use the inner product $\langle A, B\rangle = \text{trace}(AB^T)$ on $M_n\mathbb{R}$ and the Cauchy-Schwartz inequality.

(b) Now instead apply the inequality to A and A^T, noting that $\|A\| = \|A^T\| = \sqrt{\text{trace}(AA^T)}$ and $\langle A, A^T\rangle = \text{trace}(AA^{TT}) = \text{trace}(A^2)$. The conclusion is that $|\text{trace}(A^2)| \leq \sqrt{\text{trace}(AA^T)}\sqrt{\text{trace}(AA^T)}$, which gives $|\text{trace}(A^2)| \leq \text{trace}(AA^T)$. This is an equality iff A^T is a multiple of A, say $A^T = \lambda A$ for some λ. This means that $A = A^{TT} = \lambda A^T = \lambda^2 A$, and $A \neq 0$, so this means that $\lambda^2 = 1$, or equivalently $\lambda = \pm 1$. If $\lambda = 1$ then $A^T = A$ and A is symmetric; if $\lambda = -1$ then $A^T = -A$ and A is antisymmetric.

Angles

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\| \|w\| > 0$.

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$.

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$.

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\| \, \|w\| > 0$. Put $c = \langle v, w \rangle / (\|v\| \, \|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\| \, \|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\| \, \|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$.

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0,1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Example ??: Take $V = M_3\mathbb{R}$ $\quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ $\quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Example ??: Take $V = M_3\mathbb{R}$ $\quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$$\|A\| = \sqrt{0^2 + 1^2 + 0^2 + 1^2 + 2^2 + 1^2 + 0^2 + 1^2 + 0^2} = \sqrt{8} = 2\sqrt{2}$$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\| \, \|w\| > 0$. Put $c = \langle v, w \rangle / (\|v\| \, \|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0,1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w \rangle = 1/2$, so $\langle v, w \rangle / (\|v\| \, \|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Example ??: Take $V = M_3\mathbb{R}$ $\quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$$\|A\| = \sqrt{0^2 + 1^2 + 0^2 + 1^2 + 2^2 + 1^2 + 0^2 + 1^2 + 0^2} = \sqrt{8} = 2\sqrt{2}$$
$$\|B\| = \sqrt{1^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2} = \sqrt{4} = 2$$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Example ??: Take $V = M_3\mathbb{R}$ $\quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$$\begin{aligned} \|A\| &= \sqrt{0^2 + 1^2 + 0^2 + 1^2 + 2^2 + 1^2 + 0^2 + 1^2 + 0^2} = \sqrt{8} = 2\sqrt{2} \\ \|B\| &= \sqrt{1^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2} = \sqrt{4} = 2 \\ \langle A, B\rangle &= 0.1 + 1.0 + 0.0 + 1.1 + 2.1 + 1.1 + 0.0 + 1.0 + 0.0 = 4 \end{aligned}$$

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0, 1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Example ??: Take $V = M_3\mathbb{R}$ $\quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$$\begin{aligned} \|A\| &= \sqrt{0^2 + 1^2 + 0^2 + 1^2 + 2^2 + 1^2 + 0^2 + 1^2 + 0^2} = \sqrt{8} = 2\sqrt{2} \\ \|B\| &= \sqrt{1^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2} = \sqrt{4} = 2 \\ \langle A, B\rangle &= 0.1 + 1.0 + 0.0 + 1.1 + 2.1 + 1.1 + 0.0 + 1.0 + 0.0 = 4 \end{aligned}$$

so $\langle A, B\rangle/(\|A\|\|B\|) = 4/(4\sqrt{2}) = 1/\sqrt{2} = \cos(\pi/4)$.

Angles

Definition ??: Let V be an inner product space over $\mathbb{R}$, and let v and w be nonzero elements of V, so $\|v\|\,\|w\| > 0$. Put $c = \langle v, w\rangle/(\|v\|\,\|w\|)$. The Cauchy-Schwartz inequality tells us that $-1 \leq c \leq 1$, so there is a unique angle $\theta \in [0, \pi]$ such that $\cos(\theta) = c$. We call this *the angle between v and w*.

Example ??: Take $V = C[0,1]$ (with the usual inner product), and $v(t) = 1$, and $w(t) = t$. Then $\|v\| = 1$ and $\|w\| = 1/\sqrt{3}$ and $\langle v, w\rangle = 1/2$, so $\langle v, w\rangle/(\|v\|\,\|w\|) = \sqrt{3}/2 = \cos(\pi/6)$, so the angle between v and w is $\pi/6$.

Example ??: Take $V = M_3\mathbb{R}$ $\quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$$\begin{aligned}
\|A\| &= \sqrt{0^2 + 1^2 + 0^2 + 1^2 + 2^2 + 1^2 + 0^2 + 1^2 + 0^2} = \sqrt{8} = 2\sqrt{2} \\
\|B\| &= \sqrt{1^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2} = \sqrt{4} = 2 \\
\langle A, B\rangle &= 0.1 + 1.0 + 0.0 + 1.1 + 2.1 + 1.1 + 0.0 + 1.0 + 0.0 = 4
\end{aligned}$$

so $\langle A, B\rangle/(\|A\|\|B\|) = 4/(4\sqrt{2}) = 1/\sqrt{2} = \cos(\pi/4)$. The angle between A and B is thus $\pi/4$.

Orthogonal complements

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}.$$

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Lemma ??: We always have $W \cap W^{\perp} = 0$.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Lemma ??: We always have $W \cap W^{\perp} = 0$. (Thus, if W is complemented, we have $V = W \oplus W^{\perp}$.)

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Lemma ??: We always have $W \cap W^{\perp} = 0$. (Thus, if W is complemented, we have $V = W \oplus W^{\perp}$.)

Proof: Suppose that $v \in W \cap W^{\perp}$.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Lemma ??: We always have $W \cap W^{\perp} = 0$. (Thus, if W is complemented, we have $V = W \oplus W^{\perp}$.)

Proof: Suppose that $v \in W \cap W^{\perp}$. As $v \in W^{\perp}$, we have $\langle v, w \rangle = 0$ for all $w \in W$.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Lemma ??: We always have $W \cap W^{\perp} = 0$. (Thus, if W is complemented, we have $V = W \oplus W^{\perp}$.)

Proof: Suppose that $v \in W \cap W^{\perp}$. As $v \in W^{\perp}$, we have $\langle v, w \rangle = 0$ for all $w \in W$. As $v \in W$, we can take $w = v$, which gives $\|v\|^2 = \langle v, v \rangle = 0$.

Orthogonal complements

Definition ??: Let V be a vector space with inner product, and let W be a subspace. We then put

$$W^{\perp} = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}.$$

This is called the *orthogonal complement* (or *annihilator*) of W. We say that W is *complemented* if $W + W^{\perp} = V$.

Lemma ??: We always have $W \cap W^{\perp} = 0$. (Thus, if W is complemented, we have $V = W \oplus W^{\perp}$.)

Proof: Suppose that $v \in W \cap W^{\perp}$. As $v \in W^{\perp}$, we have $\langle v, w\rangle = 0$ for all $w \in W$. As $v \in W$, we can take $w = v$, which gives $\|v\|^2 = \langle v, v\rangle = 0$. This implies that $v = 0$, as required.

Orthogonal sequences

Orthogonal sequences

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \dots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$.

Orthogonal sequences

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \dots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. We say that the sequence is *strictly orthogonal* if it is orthogonal, and all the elements v_i are nonzero.

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \dots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. We say that the sequence is *strictly orthogonal* if it is orthogonal, and all the elements v_i are nonzero. We say that the sequence is *orthonormal* if it is orthogonal, and also $\langle v_i, v_i \rangle = 1$ for all i.

Orthogonal sequences

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \ldots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. We say that the sequence is *strictly orthogonal* if it is orthogonal, and all the elements v_i are nonzero. We say that the sequence is *orthonormal* if it is orthogonal, and also $\langle v_i, v_i \rangle = 1$ for all i.

Remark ??: If $\mathcal{V}$ is a strictly orthogonal sequence then we can define an orthonormal sequence $\hat{v}_1, \ldots, \hat{v}_n$ by $\hat{v}_i = v_i / \|v_i\|$.

Orthonormal examples

Orthonormal examples

Example ??: The standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for $\mathbb{R}^n$ is an orthonormal sequence.

Orthonormal examples

Example ??: The standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ for $\mathbb{R}^n$ is an orthonormal sequence.

Example ??: Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be the vectors joining the centre of the earth to the North Pole, the mouth of the river Amazon, and the city of Mogadishu.

Orthonormal examples

Example ??: The standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ for $\mathbb{R}^n$ is an orthonormal sequence.

Example ??: Let **a**, **b** and **c** be the vectors joining the centre of the earth to the North Pole, the mouth of the river Amazon, and the city of Mogadishu. These are elements of the inner product space U discussed in Examples **??** and **??**.

Orthonormal examples

Example ??: The standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ for $\mathbb{R}^n$ is an orthonormal sequence.

Example ??: Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be the vectors joining the centre of the earth to the North Pole, the mouth of the river Amazon, and the city of Mogadishu. These are elements of the inner product space U discussed in Examples **??** and **??**. Then $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is an orthogonal sequence, and $\mathbf{a}/4000, \mathbf{b}/4000, \mathbf{c}/4000$ is an orthonormal sequence.

Orthonormal examples

Example ??: The standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for $\mathbb{R}^n$ is an orthonormal sequence.

Example ??: Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be the vectors joining the centre of the earth to the North Pole, the mouth of the river Amazon, and the city of Mogadishu. These are elements of the inner product space U discussed in Examples **??** and **??**. Then $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is an orthogonal sequence, and $\mathbf{a}/4000, \mathbf{b}/4000, \mathbf{c}/4000$ is an orthonormal sequence.

(Of course, these statements are only approximations. You can take it as an exercise to work out the size of the errors involved.)

Pythagoras

Pythagoras

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \cdots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}.$$

Pythagoras

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \dots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \dots + \|v_n\|^2}.$$

Proof: We have

$$\|v\|^2 = \langle \sum_i v_i, \sum_j v_j \rangle = \sum_{i,j} \langle v_i, v_j \rangle.$$

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \cdots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}.$$

Proof: We have

$$\|v\|^2 = \langle \sum_i v_i, \sum_j v_j \rangle = \sum_{i,j} \langle v_i, v_j \rangle.$$

Because the sequence is orthogonal, all terms in the sum are zero except those for which $i = j$.

Pythagoras

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \cdots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}.$$

Proof: We have

$$\|v\|^2 = \langle \sum_i v_i, \sum_j v_j \rangle = \sum_{i,j} \langle v_i, v_j \rangle.$$

Because the sequence is orthogonal, all terms in the sum are zero except those for which $i = j$. We thus have

$$\|v\|^2 = \sum_i \langle v_i, v_i \rangle = \sum_i \|v_i\|^2.$$

Pythagoras

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \cdots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}.$$

Proof: We have

$$\|v\|^2 = \langle \sum_i v_i, \sum_j v_j \rangle = \sum_{i,j} \langle v_i, v_j \rangle.$$

Because the sequence is orthogonal, all terms in the sum are zero except those for which $i = j$. We thus have

$$\|v\|^2 = \sum_i \langle v_i, v_i \rangle = \sum_i \|v_i\|^2.$$

We can now take square roots to get the equation in the lemma.

Linear independence

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof:

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$.

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \ldots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \ldots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$, so the only nonzero term on the left hand side is $\lambda_i \langle v_i, v_i \rangle$

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$, so the only nonzero term on the left hand side is $\lambda_i \langle v_i, v_i \rangle$, so we conclude that $\lambda_i \langle v_i, v_i \rangle = 0$.

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$, so the only nonzero term on the left hand side is $\lambda_i \langle v_i, v_i \rangle$, so we conclude that $\lambda_i \langle v_i, v_i \rangle = 0$. Moreover, the sequence is *strictly* orthogonal, so $v_i \neq 0$, so $\langle v_i, v_i \rangle > 0$.

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$, so the only nonzero term on the left hand side is $\lambda_i \langle v_i, v_i \rangle$, so we conclude that $\lambda_i \langle v_i, v_i \rangle = 0$. Moreover, the sequence is *strictly* orthogonal, so $v_i \neq 0$, so $\langle v_i, v_i \rangle > 0$. It follows that we must have $\lambda_i = 0$

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \dots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \dots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$, so the only nonzero term on the left hand side is $\lambda_i \langle v_i, v_i \rangle$, so we conclude that $\lambda_i \langle v_i, v_i \rangle = 0$. Moreover, the sequence is *strictly* orthogonal, so $v_i \neq 0$, so $\langle v_i, v_i \rangle > 0$. It follows that we must have $\lambda_i = 0$, so our original linear relation was the trivial one.

Linear independence

Lemma ??: Any strictly orthogonal sequence is linearly independent.

Proof: Let $\mathcal{V} = v_1, \dots, v_n$ be a strictly orthogonal sequence, and suppose we have a linear relation $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. For each i it follows that

$$\langle v_i, \lambda_1 v_1 + \cdots + \lambda_n v_n \rangle = \langle v_i, 0 \rangle = 0.$$

The left hand side here is just

$$\lambda_1 \langle v_i, v_1 \rangle + \lambda_2 \langle v_i, v_2 \rangle + \cdots + \lambda_n \langle v_i, v_n \rangle.$$

Moreover, the sequence $\mathcal{V}$ is orthogonal, so the inner products $\langle v_i, v_j \rangle$ are zero unless $j = i$, so the only nonzero term on the left hand side is $\lambda_i \langle v_i, v_i \rangle$, so we conclude that $\lambda_i \langle v_i, v_i \rangle = 0$. Moreover, the sequence is *strictly* orthogonal, so $v_i \neq 0$, so $\langle v_i, v_i \rangle > 0$. It follows that we must have $\lambda_i = 0$, so our original linear relation was the trivial one. We conclude that $\mathcal{V}$ is linearly independent, as claimed.

Orthogonal projections

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace.

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \ldots, w_p$ that spans W

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$).

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^{\perp}$

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^\perp$, so $v = \pi(v) + (v - \pi(v)) \in W + W^\perp$.

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^\perp$, so $v = \pi(v) + (v - \pi(v)) \in W + W^\perp$. In particular, we have $W + W^\perp = V$, so W is complemented.

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^\perp$, so $v = \pi(v) + (v - \pi(v)) \in W + W^\perp$. In particular, we have $W + W^\perp = V$, so W is complemented.

Remark ??: If the sequence $\mathcal{W}$ is orthonormal, then of course we have $\langle w_k, w_k \rangle = 1$ and the formula reduces to

$$\pi(v) = \langle v, w_1 \rangle w_1 + \ldots + \langle v, w_p \rangle w_p.$$

Orthogonal projections

Proof:

Orthogonal projections

Proof: First note that the coefficients $\lambda_i = \langle v, w_i \rangle / \langle w_i, w_i \rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W.

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1 \langle w_i, w_1\rangle + \ldots + \lambda_i \langle w_i, w_i\rangle + \ldots + \lambda_p \langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$

Orthogonal projections

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero.

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero. This means that

$$\langle w_i, \pi(v)\rangle = \lambda_i\langle w_i, w_i\rangle$$

Orthogonal projections

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero. This means that

$$\langle w_i, \pi(v)\rangle = \lambda_i\langle w_i, w_i\rangle = \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_i\rangle$$

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle/\langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero. This means that

$$\langle w_i, \pi(v)\rangle = \lambda_i\langle w_i, w_i\rangle = \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_i\rangle = \langle v, w_i\rangle = \langle w_i, v\rangle$$

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero. This means that

$$\langle w_i, \pi(v)\rangle = \lambda_i\langle w_i, w_i\rangle = \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_i\rangle = \langle v, w_i\rangle = \langle w_i, v\rangle,$$

so $\langle w_i, v - \pi(v)\rangle = \langle w_i, v\rangle - \langle w_i, \pi(v)\rangle = 0$.

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle/\langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero. This means that

$$\langle w_i, \pi(v)\rangle = \lambda_i\langle w_i, w_i\rangle = \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_i\rangle = \langle v, w_i\rangle = \langle w_i, v\rangle,$$

so $\langle w_i, v - \pi(v)\rangle = \langle w_i, v\rangle - \langle w_i, \pi(v)\rangle = 0$. As this holds for all i, and the elements w_i span W, we see that $\langle w, v - \pi(v)\rangle = 0$ for all $w \in W$

Proof: First note that the coefficients $\lambda_i = \langle v, w_i\rangle / \langle w_i, w_i\rangle$ are just numbers, so the element $\pi(v) = \lambda_1 w_1 + \ldots + \lambda_p w_p$ lies in the span of $w_1, \ldots, w_p$, which is W. Next, we have

$$\langle w_i, \pi(v)\rangle = \lambda_1\langle w_i, w_1\rangle + \ldots + \lambda_i\langle w_i, w_i\rangle + \ldots + \lambda_p\langle w_i, w_p\rangle.$$

As the sequence $\mathcal{W}$ is orthogonal, we have $\langle w_i, w_j\rangle = 0$ for $j \neq i$, so only the i'th term in the above sum is nonzero. This means that

$$\langle w_i, \pi(v)\rangle = \lambda_i\langle w_i, w_i\rangle = \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle}\langle w_i, w_i\rangle = \langle v, w_i\rangle = \langle w_i, v\rangle,$$

so $\langle w_i, v - \pi(v)\rangle = \langle w_i, v\rangle - \langle w_i, \pi(v)\rangle = 0$. As this holds for all i, and the elements w_i span W, we see that $\langle w, v - \pi(v)\rangle = 0$ for all $w \in W$, or in other words, that $v - \pi(v) \in W^{\perp}$, as claimed.

Orthogonal sequences

Orthogonal sequences

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \dots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$.

Orthogonal sequences

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \dots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. We say that the sequence is *strictly orthogonal* if it is orthogonal, and all the elements v_i are nonzero.

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \cdots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}.$$

Orthogonal sequences

Definition ??: Let V be a vector space with inner product. We say that a sequence $\mathcal{V} = v_1, \dots, v_n$ of elements of V is *orthogonal* if we have $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. We say that the sequence is *strictly orthogonal* if it is orthogonal, and all the elements v_i are nonzero. We say that the sequence is *orthonormal* if it is orthogonal, and also $\langle v_i, v_i \rangle = 1$ for all i.

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence, and put $v = v_1 + \cdots + v_n$. Then

$$\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}.$$

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace.

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \ldots, w_p$ that spans W

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$).

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^\perp$

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \dots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^\perp$, so $v = \pi(v) + (v - \pi(v)) \in W + W^\perp$.

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^{\perp}$, so $v = \pi(v) + (v - \pi(v)) \in W + W^{\perp}$. In particular, we have $W + W^{\perp} = V$, so W is complemented.

Orthogonal projections

Proposition ??: Let V be a vector space with inner product, and let W be a subspace. Suppose that we have a strictly orthogonal sequence $\mathcal{W} = w_1, \dots, w_p$ that spans W, and we define

$$\pi(v) = \frac{\langle v, w_1\rangle}{\langle w_1, w_1\rangle} w_1 + \cdots + \frac{\langle v, w_p\rangle}{\langle w_p, w_p\rangle} w_p$$

(for all $v \in V$). Then $\pi(v) \in W$ and $v - \pi(v) \in W^\perp$, so $v = \pi(v) + (v - \pi(v)) \in W + W^\perp$. In particular, we have $W + W^\perp = V$, so W is complemented.

Remark ??: If the sequence $\mathcal{W}$ is orthonormal, then of course we have $\langle w_k, w_k\rangle = 1$ and the formula reduces to

$$\pi(v) = \langle v, w_1\rangle w_1 + \ldots + \langle v, w_p\rangle w_p.$$

Parseval's inequality

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \ldots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \text{span}(\mathcal{W})$.

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \mathrm{span}(\mathcal{W})$.

Proof:

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \text{span}(\mathcal{W})$.

Proof: Put $W = \text{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**.

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \ldots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \mathrm{span}(\mathcal{W})$.

Proof: Put $W = \mathrm{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^{\perp}$.

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \operatorname{span}(\mathcal{W})$.

Proof: Put $W = \operatorname{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^\perp$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \text{span}(\mathcal{W})$.

Proof: Put $W = \text{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^{\perp}$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$.

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^p \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \operatorname{span}(\mathcal{W})$.

Proof: Put $W = \operatorname{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^p \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^\perp$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$. Lemma **??** therefore tells us that

$$\|v\|^2 = \|\langle v, w_1 \rangle w_1\|^2 + \dots + \|\langle v, w_p \rangle w_p\|^2 + \|\epsilon(v)\|^2 = \|\epsilon(v)\|^2 + \sum_i \langle v, w_i \rangle^2.$$

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \mathrm{span}(\mathcal{W})$.

Proof: Put $W = \mathrm{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^\perp$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$. Lemma **??** therefore tells us that

$$\|v\|^2 = \|\langle v, w_1 \rangle w_1\|^2 + \dots + \|\langle v, w_p \rangle w_p\|^2 + \|\epsilon(v)\|^2 = \|\epsilon(v)\|^2 + \sum_i \langle v, w_i \rangle^2.$$

All terms are ≥ 0, so $\|v\|^2 \geq \sum_i \langle v, w_i \rangle^2$

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.
Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \operatorname{span}(\mathcal{W})$.

Proof: Put $W = \operatorname{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^{\perp}$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$. Lemma **??** therefore tells us that

$$\|v\|^2 = \|\langle v, w_1 \rangle w_1\|^2 + \dots + \|\langle v, w_p \rangle w_p\|^2 + \|\epsilon(v)\|^2 = \|\epsilon(v)\|^2 + \sum_i \langle v, w_i \rangle^2.$$

All terms are ≥ 0, so $\|v\|^2 \geq \sum_i \langle v, w_i \rangle^2$, with equality iff $\|\epsilon(v)\|^2 = 0$.

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V. Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \operatorname{span}(\mathcal{W})$.

Proof: Put $W = \operatorname{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^{\perp}$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$. Lemma **??** therefore tells us that

$$\|v\|^2 = \|\langle v, w_1 \rangle w_1\|^2 + \dots + \|\langle v, w_p \rangle w_p\|^2 + \|\epsilon(v)\|^2 = \|\epsilon(v)\|^2 + \sum_i \langle v, w_i \rangle^2.$$

All terms are ≥ 0, so $\|v\|^2 \geq \sum_i \langle v, w_i \rangle^2$, with equality iff $\|\epsilon(v)\|^2 = 0$. Moreover, we have $\|\epsilon(v)\|^2 = 0$ iff $\epsilon(v) = 0$

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \ldots, w_p$ be an orthonormal sequence in V. Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \text{span}(\mathcal{W})$.

Proof: Put $W = \text{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^\perp$. The sequence

$$\langle v, w_1 \rangle w_1, \ldots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$. Lemma **??** therefore tells us that

$$\|v\|^2 = \|\langle v, w_1 \rangle w_1\|^2 + \cdots + \|\langle v, w_p \rangle w_p\|^2 + \|\epsilon(v)\|^2 = \|\epsilon(v)\|^2 + \sum_i \langle v, w_i \rangle^2.$$

All terms are ≥ 0, so $\|v\|^2 \geq \sum_i \langle v, w_i \rangle^2$, with equality iff $\|\epsilon(v)\|^2 = 0$. Moreover, we have $\|\epsilon(v)\|^2 = 0$ iff $\epsilon(v) = 0$ iff $v = \pi(v)$

Parseval's inequality

Corollary ??: Let V be a vector space with inner product, and let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V. Then for any $v \in V$ we have $\|v\|^2 \geq \sum_{i=1}^{p} \langle v, w_i \rangle^2$.

Moreover, this is actually an equality iff $v \in \mathrm{span}(\mathcal{W})$.

Proof: Put $W = \mathrm{span}(\mathcal{W})$, and put $\pi(v) = \sum_{i=1}^{p} \langle v, w_i \rangle w_i$ as in Proposition **??**. Put $\epsilon(v) = v - \pi(v)$, which lies in $W^{\perp}$. The sequence

$$\langle v, w_1 \rangle w_1, \dots, \langle v, w_p \rangle w_p, \epsilon(v)$$

is orthogonal, and the sum of the sequence is $\pi(v) + \epsilon(v) = v$. Lemma **??** therefore tells us that

$$\|v\|^2 = \|\langle v, w_1 \rangle w_1\|^2 + \cdots + \|\langle v, w_p \rangle w_p\|^2 + \|\epsilon(v)\|^2 = \|\epsilon(v)\|^2 + \sum_i \langle v, w_i \rangle^2.$$

All terms are ≥ 0, so $\|v\|^2 \geq \sum_i \langle v, w_i \rangle^2$, with equality iff $\|\epsilon(v)\|^2 = 0$. Moreover, we have $\|\epsilon(v)\|^2 = 0$ iff $\epsilon(v) = 0$ iff $v = \pi(v)$ iff $v \in W$.

The closest point

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof:

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^\perp$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^\perp$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$)

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^\perp$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^\perp$

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$.

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$. It follows that

$$\|v - w\|^2 = \|x + y\|^2$$

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$. It follows that

$$\|v - w\|^2 = \|x + y\|^2 = \langle x + y, x + y\rangle$$

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^\perp$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^\perp$, so $\langle x, y \rangle = 0 = \langle y, x \rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$. It follows that

$$\begin{aligned} \|v - w\|^2 = \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \end{aligned}$$

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$. It follows that

$$\begin{aligned}
\|v - w\|^2 = \|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \|x\|^2 + 0 + 0 + \|y\|^2
\end{aligned}$$

The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$. It follows that

$$\begin{aligned}\|v - w\|^2 &= \|x + y\|^2 = \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 0 + 0 + \|y\|^2 > \|x\|^2.\end{aligned}$$

This shows that $\|v - w\| > \|x\| = \|v - \pi(v)\|$

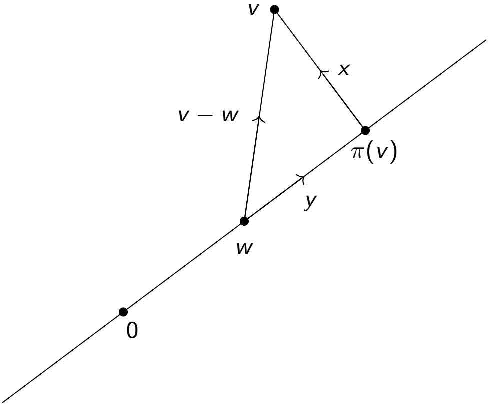
The closest point

Proposition ??: Let $\mathcal{W}$ and π be as in Proposition **??**. Then $\pi(v)$ is the point in W that is closest to v.

Proof: Put $x = v - \pi(v)$, so $x \in W^{\perp}$. The distance from v to $\pi(v)$ is just $\|v - \pi(v)\| = \|x\|$. Now consider another point $w \in W$, with $w \neq \pi(v)$. The distance from v to w is just $\|v - w\|$; we must show that this is larger than $\|x\|$. Put $y = \pi(v) - w$, and note that $v - w = \pi(v) + x - w = x + y$. Note also that $y \in W$ (because $\pi(v) \in W$ and $w \in W$) and $x \in W^{\perp}$, so $\langle x, y\rangle = 0 = \langle y, x\rangle$. Finally, note that $y \neq 0$ and so $\|y\| > 0$. It follows that

$$\begin{aligned} \|v - w\|^2 &= \|x + y\|^2 = \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 0 + 0 + \|y\|^2 > \|x\|^2. \end{aligned}$$

This shows that $\|v - w\| > \|x\| = \|v - \pi(v)\|$, so w is further from v than $\pi(v)$ is.



v
x
$v - w$
$\pi(v)$
y
w
0

The Gram-Schmidt procedure

The Gram-Schmidt procedure

Theorem ??:

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \ldots, u_n$ be a linearly independent list of elements of V.

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \ldots, u_n$ be a linearly independent list of elements of V. Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \ldots, v_n$ such that $\text{span}(v_1, \ldots, v_i) = \text{span}(u_1, \ldots, u_i)$ for all i.

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \ldots, u_n$ be a linearly independent list of elements of V.
Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \ldots, v_n$ such that $\text{span}(v_1, \ldots, v_i) = \text{span}(u_1, \ldots, u_i)$ for all i.

Proof:

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \ldots, u_n$ be a linearly independent list of elements of V.
Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \ldots, v_n$ such that $\mathrm{span}(v_1, \ldots, v_i) = \mathrm{span}(u_1, \ldots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V.
Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\text{span}(v_1, \dots, v_i) = \text{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \text{span}(u_1, \dots, u_i)$.

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V. Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\text{span}(v_1, \dots, v_i) = \text{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \text{span}(u_1, \dots, u_i)$. We will construct the elements v_i by induction.

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V. Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\text{span}(v_1, \dots, v_i) = \text{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \text{span}(u_1, \dots, u_i)$. We will construct the elements v_i by induction.

For the initial step, we take $v_1 = u_1$, so (v_1) is an orthogonal basis for U_1.

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V. Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\mathrm{span}(v_1, \dots, v_i) = \mathrm{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \mathrm{span}(u_1, \dots, u_i)$. We will construct the elements v_i by induction.

For the initial step, we take $v_1 = u_1$, so (v_1) is an orthogonal basis for U_1.

Suppose we have constructed an orthogonal basis $v_1, \dots, v_{i-1}$ for U_{i-1}.

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V. Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\mathrm{span}(v_1, \dots, v_i) = \mathrm{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \mathrm{span}(u_1, \dots, u_i)$. We will construct the elements v_i by induction.

For the initial step, we take $v_1 = u_1$, so (v_1) is an orthogonal basis for U_1.

Suppose we have constructed an orthogonal basis $v_1, \dots, v_{i-1}$ for U_{i-1}. Proposition **??** then tells us that U_{i-1} is complemented

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V.
Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\mathrm{span}(v_1, \dots, v_i) = \mathrm{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \mathrm{span}(u_1, \dots, u_i)$. We will construct the elements v_i by induction.

For the initial step, we take $v_1 = u_1$, so (v_1) is an orthogonal basis for U_1.

Suppose we have constructed an orthogonal basis $v_1, \dots, v_{i-1}$ for U_{i-1}.
Proposition **??** then tells us that U_{i-1} is complemented, so $V = U_{i-1}^{\perp} + U_{i-1}$.

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \ldots, u_n$ be a linearly independent list of elements of V.
Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \ldots, v_n$ such that $\mathrm{span}(v_1, \ldots, v_i) = \mathrm{span}(u_1, \ldots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \mathrm{span}(u_1, \ldots, u_i)$. We will construct the elements v_i by induction.

For the initial step, we take $v_1 = u_1$, so (v_1) is an orthogonal basis for U_1.

Suppose we have constructed an orthogonal basis $v_1, \ldots, v_{i-1}$ for U_{i-1}.
Proposition **??** then tells us that U_{i-1} is complemented, so $V = U_{i-1}^{\perp} + U_{i-1}$.
In particular, we can write $u_i = v_i + w_i$ with $v_i \in U_{i-1}^{\perp}$ and $w_i \in U_{i-1}$.

The Gram-Schmidt procedure

Theorem ??: Let V be a vector space with inner product, and let $\mathcal{U} = u_1, \dots, u_n$ be a linearly independent list of elements of V. Then there is a strictly orthogonal sequence $\mathcal{V} = v_1, \dots, v_n$ such that $\mathrm{span}(v_1, \dots, v_i) = \mathrm{span}(u_1, \dots, u_i)$ for all i.

Proof: The sequence $\mathcal{W}$ is generated by the *Gram-Schmidt procedure*, which we now describe.

Put $U_i = \mathrm{span}(u_1, \dots, u_i)$. We will construct the elements v_i by induction.

For the initial step, we take $v_1 = u_1$, so (v_1) is an orthogonal basis for U_1.

Suppose we have constructed an orthogonal basis $v_1, \dots, v_{i-1}$ for U_{i-1}. Proposition **??** then tells us that U_{i-1} is complemented, so $V = U_{i-1}^{\perp} + U_{i-1}$. In particular, we can write $u_i = v_i + w_i$ with $v_i \in U_{i-1}^{\perp}$ and $w_i \in U_{i-1}$. Explicitly, the formulae are

$$w_i = \sum_{j=1}^{i-1} \frac{\langle u_i, v_j \rangle}{\langle v_j, v_j \rangle} v_j \qquad v_i = u_i - w_i.$$

The Gram-Schmidt procedure

$U_k = \text{span}(u_1, \ldots, u_k)$ $v_1, \ldots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $v_i \in U_{i-1}^{\perp}$ $w_i \in U_{i-1}$

The Gram-Schmidt procedure

$U_k = \text{span}(u_1, \ldots, u_k)$ $\qquad$ $v_1, \ldots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $\qquad$ $v_i \in U_{i-1}^{\perp}$ $\qquad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^{\perp}$ and $v_1, \ldots, v_{i-1} \in U_{i-1}$

The Gram-Schmidt procedure

$U_k = \mathrm{span}(u_1, \ldots, u_k)$ $\qquad$ $v_1, \ldots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $\qquad$ $v_i \in U_{i-1}^{\perp}$ $\qquad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^{\perp}$ and $v_1, \ldots, v_{i-1} \in U_{i-1}$, we have $\langle v_i, v_j \rangle = 0$ for $j < i$

The Gram-Schmidt procedure

$U_k = \mathrm{span}(u_1, \dots, u_k)$ $\qquad$ $v_1, \dots, v_{i-1}$ an orthogonal basis for U_{i-1}
$u_i = v_i + w_i$ $\qquad$ $v_i \in U_{i-1}^{\perp}$ $\qquad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^{\perp}$ and $v_1, \dots, v_{i-1} \in U_{i-1}$, we have $\langle v_i, v_j \rangle = 0$ for $j < i$, so $(v_1, \dots, v_i)$ is an orthogonal sequence.

Next, note that $U_i = U_{i-1} + \mathbb{R} u_i$.

The Gram-Schmidt procedure

$U_k = \text{span}(u_1, \ldots, u_k)$ $\qquad$ $v_1, \ldots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $\qquad$ $v_i \in U_{i-1}^{\perp}$ $\qquad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^{\perp}$ and $v_1, \ldots, v_{i-1} \in U_{i-1}$, we have $\langle v_i, v_j \rangle = 0$ for $j < i$, so $(v_1, \ldots, v_i)$ is an orthogonal sequence.

Next, note that $U_i = U_{i-1} + \mathbb{R}u_i$. As $u_i = v_i + w_i$ with $w_i \in U_{i-1}$, we see that this is the same as $U_{i-1} + \mathbb{R}v_i$.

The Gram-Schmidt procedure

$U_k = \mathrm{span}(u_1, \dots, u_k)$ $\quad$ $v_1, \dots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $\quad$ $v_i \in U_{i-1}^{\perp}$ $\quad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^{\perp}$ and $v_1, \dots, v_{i-1} \in U_{i-1}$, we have $\langle v_i, v_j \rangle = 0$ for $j < i$, so $(v_1, \dots, v_i)$ is an orthogonal sequence.

Next, note that $U_i = U_{i-1} + \mathbb{R}u_i$. As $u_i = v_i + w_i$ with $w_i \in U_{i-1}$, we see that this is the same as $U_{i-1} + \mathbb{R}v_i$. By our induction hypothesis, we have $U_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$

The Gram-Schmidt procedure

$U_k = \text{span}(u_1, \dots, u_k)$ $\qquad$ $v_1, \dots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $\qquad$ $v_i \in U_{i-1}^\perp$ $\qquad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^\perp$ and $v_1, \dots, v_{i-1} \in U_{i-1}$, we have $\langle v_i, v_j \rangle = 0$ for $j < i$, so $(v_1, \dots, v_i)$ is an orthogonal sequence.

Next, note that $U_i = U_{i-1} + \mathbb{R}u_i$. As $u_i = v_i + w_i$ with $w_i \in U_{i-1}$, we see that this is the same as $U_{i-1} + \mathbb{R}v_i$. By our induction hypothesis, we have $U_{i-1} = \text{span}(v_1, \dots, v_{i-1})$, and it follows that $U_i = U_{i-1} + \mathbb{R}v_i = \text{span}(v_1, \dots, v_i)$.

The Gram-Schmidt procedure

$U_k = \mathrm{span}(u_1, \dots, u_k)$ $\qquad$ $v_1, \dots, v_{i-1}$ an orthogonal basis for U_{i-1}

$u_i = v_i + w_i$ $\qquad$ $v_i \in U_{i-1}^{\perp}$ $\qquad$ $w_i \in U_{i-1}$

As $v_i \in U_{i-1}^{\perp}$ and $v_1, \dots, v_{i-1} \in U_{i-1}$, we have $\langle v_i, v_j \rangle = 0$ for $j < i$, so $(v_1, \dots, v_i)$ is an orthogonal sequence.

Next, note that $U_i = U_{i-1} + \mathbb{R}u_i$. As $u_i = v_i + w_i$ with $w_i \in U_{i-1}$, we see that this is the same as $U_{i-1} + \mathbb{R}v_i$. By our induction hypothesis, we have $U_{i-1} = \mathrm{span}(v_1, \dots, v_{i-1})$, and it follows that $U_i = U_{i-1} + \mathbb{R}v_i = \mathrm{span}(v_1, \dots, v_i)$.

This means that $v_1, \dots, v_i$ is a spanning set of the i-dimensional space U_i, so it must be a basis. $\square$

The Gram-Schmidt procedure

Corollary ??: If V and $\mathcal{U}$ are as above, then there is an *orthonormal* sequence $\hat{v}_1, \ldots, \hat{v}_n$ with $\text{span}(\hat{v}_1, \ldots, \hat{v}_i) = \text{span}(u_1, \ldots, u_i)$ for all i.

Corollary ??: If V and $\mathcal{U}$ are as above, then there is an *orthonormal* sequence $\hat{v}_1, \ldots, \hat{v}_n$ with $\text{span}(\hat{v}_1, \ldots, \hat{v}_i) = \text{span}(u_1, \ldots, u_i)$ for all i.

Proof:

Corollary ??: If V and $\mathcal{U}$ are as above, then there is an *orthonormal* sequence $\hat{v}_1, \ldots, \hat{v}_n$ with $\mathrm{span}(\hat{v}_1, \ldots, \hat{v}_i) = \mathrm{span}(u_1, \ldots, u_i)$ for all i.

Proof: Just find a strictly orthogonal sequence $v_1, \ldots, v_n$ as in the Proposition

Corollary ??: If V and $\mathcal{U}$ are as above, then there is an *orthonormal* sequence $\hat{v}_1, \dots, \hat{v}_n$ with $\text{span}(\hat{v}_1, \dots, \hat{v}_i) = \text{span}(u_1, \dots, u_i)$ for all i.

Proof: Just find a strictly orthogonal sequence $v_1, \dots, v_n$ as in the Proposition, and put $\hat{v}_i = v_i / \|v_i\|$ as in Remark **??**.

An example

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \mathrm{span}(u_1, u_2, u_3, u_4)$.

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$.

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \mathrm{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

It follows that $\langle v_2, v_2 \rangle = 3/2$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \mathrm{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

It follows that $\langle v_2, v_2 \rangle = 3/2$ and $\langle u_3, v_2 \rangle = 1$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \mathrm{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

It follows that $\langle v_2, v_2 \rangle = 3/2$ and $\langle u_3, v_2 \rangle = 1$, whereas $\langle u_3, v_1 \rangle = 0$.

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

It follows that $\langle v_2, v_2 \rangle = 3/2$ and $\langle u_3, v_2 \rangle = 1$, whereas $\langle u_3, v_1 \rangle = 0$. It follows that

$$v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2$$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \mathrm{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

It follows that $\langle v_2, v_2 \rangle = 3/2$ and $\langle u_3, v_2 \rangle = 1$, whereas $\langle u_3, v_1 \rangle = 0$. It follows that

$$v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3/2} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

An example

Example ??: Consider the following elements of $\mathbb{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We apply the Gram-Schmidt procedure to get an orthogonal basis for the space $U = \text{span}(u_1, u_2, u_3, u_4)$. We have $v_1 = u_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$, so $\langle v_1, v_1 \rangle = 2$ and $\langle u_2, v_1 \rangle = 1$. Next, we have

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

It follows that $\langle v_2, v_2 \rangle = 3/2$ and $\langle u_3, v_2 \rangle = 1$, whereas $\langle u_3, v_1 \rangle = 0$. It follows that

$$v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3/2} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}.$$

An example

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}$$

An example

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}$$

It now follows that $\langle v_3, v_3 \rangle = 4/3$ and $\langle u_4, v_3 \rangle = 1$

An example

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}$$

It now follows that $\langle v_3, v_3 \rangle = 4/3$ and $\langle u_4, v_3 \rangle = 1$, whereas $\langle u_4, v_1 \rangle = \langle u_4, v_2 \rangle = 0$.

An example

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}$$

It now follows that $\langle v_3, v_3 \rangle = 4/3$ and $\langle u_4, v_3 \rangle = 1$, whereas $\langle u_4, v_1 \rangle = \langle u_4, v_2 \rangle = 0$. It follows that

$$v_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_4, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \frac{\langle u_4, v_3 \rangle}{\langle v_3, v_3 \rangle} v_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{4/3} \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1/4 \\ -1/4 \\ 1/4 \\ 1 \end{bmatrix}.$$

An example

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}$$

It now follows that $\langle v_3, v_3 \rangle = 4/3$ and $\langle u_4, v_3 \rangle = 1$, whereas $\langle u_4, v_1 \rangle = \langle u_4, v_2 \rangle = 0$. It follows that

$$v_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_4, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \frac{\langle u_4, v_3 \rangle}{\langle v_3, v_3 \rangle} v_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{4/3} \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1/4 \\ -1/4 \\ 1/4 \\ 1 \end{bmatrix}.$$

In conclusion, we have

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \\ 0 \end{bmatrix} \qquad v_4 = \begin{bmatrix} -1/4 \\ 1/4 \\ -1/4 \\ 1/4 \\ 1 \end{bmatrix}.$$

A polynomial example

A polynomial example

Example ??:

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$.

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V.

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\, dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with
$v_1 = u_1 = 1$

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with
$v_1 = u_1 = 1$

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\, dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with $v_1 = u_1 = 1$, and note that $\langle v_1, v_1 \rangle = \int_{-1}^{1} 1\, dx = 2$.

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with $v_1 = u_1 = 1$, and note that $\langle v_1, v_1 \rangle = \int_{-1}^{1} 1\,dx = 2$. We also have $\langle x, v_1 \rangle = \int_{-1}^{1} x\,dx = \left[x^2/2\right]_{-1}^{1} = 0$, so x is already orthogonal to v_1.

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with $v_1 = u_1 = 1$, and note that $\langle v_1, v_1 \rangle = \int_{-1}^{1} 1\,dx = 2$. We also have $\langle x, v_1 \rangle = \int_{-1}^{1} x\,dx = \left[x^2/2\right]_{-1}^{1} = 0$, so x is already orthogonal to v_1. It follows that

$$v_2 = x - \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = x,$$

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with $v_1 = u_1 = 1$, and note that $\langle v_1, v_1 \rangle = \int_{-1}^{1} 1\,dx = 2$. We also have $\langle x, v_1 \rangle = \int_{-1}^{1} x\,dx = \left[x^2/2\right]_{-1}^{1} = 0$, so x is already orthogonal to v_1. It follows that

$$v_2 = x - \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = x,$$

and thus that $\langle v_2, v_2 \rangle = \int_{-1}^{1} x^2\,dx = \left[x^3/3\right]_{-1}^{1} = 2/3$.

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with $v_1 = u_1 = 1$, and note that $\langle v_1, v_1 \rangle = \int_{-1}^{1} 1\,dx = 2$. We also have $\langle x, v_1 \rangle = \int_{-1}^{1} x\,dx = \left[x^2/2\right]_{-1}^{1} = 0$, so x is already orthogonal to v_1. It follows that

$$v_2 = x - \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = x,$$

and thus that $\langle v_2, v_2 \rangle = \int_{-1}^{1} x^2\,dx = \left[x^3/3\right]_{-1}^{1} = 2/3$. We also have

$$\langle x^2, v_1 \rangle = \int_{-1}^{1} x^2\,dx = 2/3 \qquad \langle x^2, v_2 \rangle = \int_{-1}^{1} x^3\,dx = 0$$

A polynomial example

Example ??: Consider the space $V = \mathbb{R}[x]_{\leq 2}$ with the inner product $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,dx$. We will apply the Gram-Schmidt procedure to the usual basis $1, x, x^2$ to get an orthonormal basis for V. We start with $v_1 = u_1 = 1$, and note that $\langle v_1, v_1\rangle = \int_{-1}^{1} 1\,dx = 2$. We also have $\langle x, v_1\rangle = \int_{-1}^{1} x\,dx = \left[x^2/2\right]_{-1}^{1} = 0$, so x is already orthogonal to v_1. It follows that

$$v_2 = x - \frac{\langle x, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = x,$$

and thus that $\langle v_2, v_2\rangle = \int_{-1}^{1} x^2\,dx = \left[x^3/3\right]_{-1}^{1} = 2/3$. We also have

$$\langle x^2, v_1\rangle = \int_{-1}^{1} x^2\,dx = 2/3 \qquad \langle x^2, v_2\rangle = \int_{-1}^{1} x^3\,dx = 0$$

so

$$v_3 = x^2 - \frac{\langle x^2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle x^2, v_2\rangle}{\langle v_2, v_2\rangle} v_2 = x^2 - \frac{2/3}{2} 1 = x^2 - 1/3.$$

A polynomial example

$$v_1 = 1 \qquad v_2 = x \qquad v_3 = x^2 - 1/3$$

We find that

$$\langle v_3, v_3 \rangle = \int_{-1}^{1} (x^2 - 1/3)^2 \, dx = \int_{-1}^{1} x^4 - \tfrac{2}{3}x^2 + \tfrac{1}{9} \, dx = \left[\tfrac{1}{5}x^5 - \tfrac{2}{9}x^3 + \tfrac{1}{9}x \right]_{-1}^{1} = 8/45.$$

The required orthonormal basis is thus given by

$$\begin{aligned} \hat{v}_1 &= v_1/\|v_1\| = 1/\sqrt{2} \\ \hat{v}_2 &= v_2/\|v_2\| = \sqrt{3/2}x \\ \hat{v}_3 &= v_3/\|v_3\| = \sqrt{45/8}(x^2 - 1/3). \end{aligned}$$

A matrix example

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero.

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

The general form of a matrix in V is $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$.

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

The general form of a matrix in V is $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$. Thus, if we put

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

The general form of a matrix in V is $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$. Thus, if we put

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

we see that an arbitrary element $A \in V$ can be written uniquely as $aA_1 + bA_2 + cA_3 + dA_4 + eA_5$

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

The general form of a matrix in V is $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$. Thus, if we put

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

we see that an arbitrary element $A \in V$ can be written uniquely as $aA_1 + bA_2 + cA_3 + dA_4 + eA_5$, so $A_1, \ldots, A_5$ is a basis for V.

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

The general form of a matrix in V is $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$. Thus, if we put

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

we see that an arbitrary element $A \in V$ can be written uniquely as $aA_1 + bA_2 + cA_3 + dA_4 + eA_5$, so $A_1, \dots, A_5$ is a basis for V.

It is not far from being an orthonormal basis: we have $\langle A_i, A_i \rangle = 2$ for all i

A matrix example

Example ??: Consider $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and let V be the space of 3×3 symmetric matrices of trace zero. We will find the matrix $Q \in V$ closest to P.

The general form of a matrix in V is $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & -a-d \end{bmatrix}$. Thus, if we put

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

we see that an arbitrary element $A \in V$ can be written uniquely as $aA_1 + bA_2 + cA_3 + dA_4 + eA_5$, so $A_1, \dots, A_5$ is a basis for V.

It is not far from being an orthonormal basis: we have $\langle A_i, A_i \rangle = 2$ for all i, and when $i \neq j$ we have $\langle A_i, A_j \rangle = 0$ except for the case $\langle A_1, A_4 \rangle = 1$.

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$=$$

$$=$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$B_1 = A_1$$

$$=$$

$$=$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 - \frac{\langle A_2, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 \\ &= \\ &= \end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$B_1 = A_1$$

$$B_2 = A_2 - \frac{\langle A_2, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 = A_2$$

$$=$$

$$=$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 - \frac{\langle A_2 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 = A_2 \\ B_3 &= A_3 - \frac{\langle A_3 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_3 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 \\ &= \\ &= \end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_3, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 = A_3 \\
&= \\
&=
\end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_3, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 = A_3 \\
B_4 &= A_4 - \frac{\langle A_4, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_4, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 - \frac{\langle A_4, B_3 \rangle}{\langle B_3, B_3 \rangle} B_3 \\
&=
\end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_3, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 = A_3 \\
B_4 &= A_4 - \frac{\langle A_4, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_4, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 - \frac{\langle A_4, B_3 \rangle}{\langle B_3, B_3 \rangle} B_3 = A_4 - \frac{1}{2} B_1 \\
\\
&=
\end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_3 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 = A_3 \\
B_4 &= A_4 - \frac{\langle A_4 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_4 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 - \frac{\langle A_4 , B_3 \rangle}{\langle B_3 , B_3 \rangle} B_3 = A_4 - \frac{1}{2} B_1 \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\
&=
\end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_3 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 = A_3 \\
B_4 &= A_4 - \frac{\langle A_4 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_4 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 - \frac{\langle A_4 , B_3 \rangle}{\langle B_3 , B_3 \rangle} B_3 = A_4 - \frac{1}{2} B_1 \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} \\
&=
\end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_3, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 = A_3 \\
B_4 &= A_4 - \frac{\langle A_4, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_4, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 - \frac{\langle A_4, B_3 \rangle}{\langle B_3, B_3 \rangle} B_3 = A_4 - \frac{1}{2} B_1 \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} \\
B_5 &= A_5 - \frac{\langle A_5, B_1 \rangle}{\langle B_1, B_1 \rangle} B_1 - \frac{\langle A_5, B_2 \rangle}{\langle B_2, B_2 \rangle} B_2 - \frac{\langle A_5, B_3 \rangle}{\langle B_3, B_3 \rangle} B_3 - \frac{\langle A_5, B_4 \rangle}{\langle B_4, B_4 \rangle} B_4
\end{aligned}$$

A matrix example

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The Gram-Schmidt procedure works out as follows:

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_2 - \frac{\langle A_2 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 = A_2 \\
B_3 &= A_3 - \frac{\langle A_3 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_3 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 = A_3 \\
B_4 &= A_4 - \frac{\langle A_4 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_4 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 - \frac{\langle A_4 , B_3 \rangle}{\langle B_3 , B_3 \rangle} B_3 = A_4 - \frac{1}{2} B_1 \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} \\
B_5 &= A_5 - \frac{\langle A_5 , B_1 \rangle}{\langle B_1 , B_1 \rangle} B_1 - \frac{\langle A_5 , B_2 \rangle}{\langle B_2 , B_2 \rangle} B_2 - \frac{\langle A_5 , B_3 \rangle}{\langle B_3 , B_3 \rangle} B_3 - \frac{\langle A_5 , B_4 \rangle}{\langle B_4 , B_4 \rangle} B_4 = A_5 .
\end{aligned}$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We have $\|B_4\| = \sqrt{3/2}$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We have $\|B_4\| = \sqrt{3/2}$ and $\|B_i\| = \sqrt{2}$ for all other i.

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We have $\|B_4\| = \sqrt{3/2}$ and $\|B_i\| = \sqrt{2}$ for all other i. After noting that $(1/2)/\sqrt{3/2} = 1/\sqrt{6}$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We have $\|B_4\| = \sqrt{3/2}$ and $\|B_i\| = \sqrt{2}$ for all other i. After noting that $(1/2)/\sqrt{3/2} = 1/\sqrt{6}$, it follows that the following matrices give an orthonormal basis for V:

$$\widehat{B}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \widehat{B}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \widehat{B}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We have $\|B_4\| = \sqrt{3/2}$ and $\|B_i\| = \sqrt{2}$ for all other i. After noting that $(1/2)/\sqrt{3/2} = 1/\sqrt{6}$, it follows that the following matrices give an orthonormal basis for V:

$$\widehat{B}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \widehat{B}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \widehat{B}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\widehat{B}_4 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

We have $\|B_4\| = \sqrt{3/2}$ and $\|B_i\| = \sqrt{2}$ for all other i. After noting that $(1/2)/\sqrt{3/2} = 1/\sqrt{6}$, it follows that the following matrices give an orthonormal basis for V:

$$\widehat{B}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \widehat{B}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \widehat{B}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\widehat{B}_4 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \widehat{B}_5 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$ and $\langle P, B_4 \rangle = -1/2$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$ and $\langle P, B_4 \rangle = -1/2$ and $\langle P, B_5 \rangle = 0$.

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$ and $\langle P, B_4 \rangle = -1/2$ and $\langle P, B_5 \rangle = 0$.

Also $\langle B_1, B_1 \rangle = \langle B_2, B_2 \rangle = \langle B_3, B_3 \rangle = 2$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$ and $\langle P, B_4 \rangle = -1/2$ and $\langle P, B_5 \rangle = 0$.

Also $\langle B_1, B_1 \rangle = \langle B_2, B_2 \rangle = \langle B_3, B_3 \rangle = 2$ and $\langle B_4, B_4 \rangle = 3/2$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$ and $\langle P, B_4 \rangle = -1/2$ and $\langle P, B_5 \rangle = 0$.

Also $\langle B_1, B_1 \rangle = \langle B_2, B_2 \rangle = \langle B_3, B_3 \rangle = 2$ and $\langle B_4, B_4 \rangle = 3/2$, so

$$Q = \frac{1}{2}(B_1 + B_2 + B_3) + \frac{-1}{2}\frac{2}{3}B_4$$

A matrix example

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B_4 = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \quad B_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The closest point in V to P is $Q = \sum_{i=1}^{5} \frac{\langle P, B_i \rangle}{\langle B_i, B_i \rangle} B_i$.

The relevant inner products are $\langle P, B_1 \rangle = \langle P, B_2 \rangle = \langle P, B_3 \rangle = 1$ and $\langle P, B_4 \rangle = -1/2$ and $\langle P, B_5 \rangle = 0$.

Also $\langle B_1, B_1 \rangle = \langle B_2, B_2 \rangle = \langle B_3, B_3 \rangle = 2$ and $\langle B_4, B_4 \rangle = 3/2$, so

$$Q = \frac{1}{2}(B_1 + B_2 + B_3) + \frac{-1}{2}\frac{2}{3}B_4 = \begin{bmatrix} 2/3 & 1/2 & 1/2 \\ 1/2 & -1/3 & 0 \\ 1/2 & 0 & -1/3 \end{bmatrix}$$

Hermitian forms

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??:

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$.

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v \rangle \in \mathbb{C}$ for each $u, v \in V$

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v \rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v \rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v \rangle = t\langle u, v \rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

(d) For all $u \in V$ we have $\langle u, u\rangle \geq 0$

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

(d) For all $u \in V$ we have $\langle u, u\rangle \geq 0$ (which is meaningful because $\langle u, u\rangle \in \mathbb{R}$)

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

(d) For all $u \in V$ we have $\langle u, u\rangle \geq 0$ (which is meaningful because $\langle u, u\rangle \in \mathbb{R}$), and $\langle u, u\rangle = 0$ iff $u = 0$.

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

(d) For all $u \in V$ we have $\langle u, u\rangle \geq 0$ (which is meaningful because $\langle u, u\rangle \in \mathbb{R}$), and $\langle u, u\rangle = 0$ iff $u = 0$.

Note that (b) and (c) together imply that $\langle u, tv\rangle = \overline{t}\langle u, v\rangle$.

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

(d) For all $u \in V$ we have $\langle u, u\rangle \geq 0$ (which is meaningful because $\langle u, u\rangle \in \mathbb{R}$), and $\langle u, u\rangle = 0$ iff $u = 0$.

Note that (b) and (c) together imply that $\langle u, tv\rangle = \overline{t}\langle u, v\rangle$.
Given an inner product, we will write $\|u\| = \sqrt{\langle u, u\rangle}$, and call this the *norm* of u.

Hermitian forms

We now discuss the analogue of inner products for complex vector spaces.

Given $z = x + iy \in \mathbb{C}$, we write $\overline{z}$ for the complex conjugate, which is $x - iy$.

Definition ??: Let V be a vector space over $\mathbb{C}$. A *Hermitian form* on V is a rule that gives a number $\langle u, v\rangle \in \mathbb{C}$ for each $u, v \in V$, such that:

(a) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.

(b) $\langle tu, v\rangle = t\langle u, v\rangle$ for all $u, v \in V$ and $t \in \mathbb{C}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$. In particular, by taking $v = u$ we see that $\langle u, u\rangle = \overline{\langle u, u\rangle}$, so $\langle u, u\rangle$ is real.

(d) For all $u \in V$ we have $\langle u, u\rangle \geq 0$ (which is meaningful because $\langle u, u\rangle \in \mathbb{R}$), and $\langle u, u\rangle = 0$ iff $u = 0$.

Note that (b) and (c) together imply that $\langle u, tv\rangle = \overline{t}\langle u, v\rangle$.
Given an inner product, we will write $\|u\| = \sqrt{\langle u, u\rangle}$, and call this the *norm* of u. We say that u is a *unit vector* if $\|u\| = 1$.

The standard Hermitian form on $\mathbb{C}^n$

The standard Hermitian form on $\mathbb{C}^n$

Example ??: We can define a Hermitian form on $\mathbb{C}^n$ by

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 \overline{v_1} + \cdots + u_n \overline{v_n}.$$

The standard Hermitian form on $\mathbb{C}^n$

Example ??: We can define a Hermitian form on $\mathbb{C}^n$ by

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 \overline{v_1} + \cdots + u_n \overline{v_n}.$$

This gives

$$\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle = |u_1|^2 + \cdots + |u_n|^2.$$

Hermitian adjoints

Hermitian adjoints

Definition ??: For any $n \times m$ matrix A over $\mathbb{C}$, we let $A^\dagger$ be the complex conjugate of the transpose of A

Hermitian adjoints

Definition ??: For any $n \times m$ matrix A over $\mathbb{C}$, we let $A^\dagger$ be the complex conjugate of the transpose of A, so for example

$$\begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \end{bmatrix}^\dagger = \begin{bmatrix} 1-i & 4-i \\ 2-i & 5-i \\ 3-i & 6-i \end{bmatrix}.$$

Hermitian adjoints

Definition ??: For any $n \times m$ matrix A over $\mathbb{C}$, we let $A^\dagger$ be the complex conjugate of the transpose of A, so for example

$$\begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \end{bmatrix}^\dagger = \begin{bmatrix} 1-i & 4-i \\ 2-i & 5-i \\ 3-i & 6-i \end{bmatrix}.$$

The above Hermitian form on $\mathbb{C}^n$ can then be rewritten as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^\dagger \mathbf{u} = \overline{\mathbf{u}^\dagger \mathbf{v}}.$$

Hermitian examples

Example ??: We can define a Hermitian form on $\mathbb{C}[t]$ by

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

Hermitian examples

Example ??: We can define a Hermitian form on $\mathbb{C}[t]$ by

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

This gives

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 |f(t)|^2\, dt.$$

Hermitian examples

Example ??: We can define a Hermitian form on $\mathbb{C}[t]$ by

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt.$$

This gives

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 |f(t)|^2\,dt.$$

Example ??: We can define a Hermitian form on $M_n\mathbb{C}$ by $\langle A, B\rangle = \mathrm{trace}(B^\dagger A)$.

Hermitian examples

Example ??: We can define a Hermitian form on $\mathbb{C}[t]$ by

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt.$$

This gives

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 |f(t)|^2\,dt.$$

Example ??: We can define a Hermitian form on $M_n\mathbb{C}$ by $\langle A, B\rangle = \text{trace}(B^\dagger A)$. If we identify $M_n\mathbb{C}$ with $\mathbb{C}^{n^2}$ in the usual way, then this is just the same as the Hermitian form in Example **??**.

Results about Hermitian forms

Results about Hermitian forms

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\| \, \|w\|$,

Results about Hermitian forms

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent over $\mathbb{C}$. $\square$

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\|\,\|w\|$, with equality iff v and w are linearly dependent over $\mathbb{C}$. $\square$

Lemma ??: Let $v_1, \ldots, v_n$ be an orthogonal sequence in V, and put $v = v_1 + \cdots + v_n$.

Results about Hermitian forms

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent over $\mathbb{C}$. $\square$

Lemma ??: Let $v_1, \ldots, v_n$ be an orthogonal sequence in V, and put $v = v_1 + \cdots + v_n$. Then $\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}$. $\square$

Results about Hermitian forms

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent over $\mathbb{C}$. $\square$

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence in V, and put $v = v_1 + \cdots + v_n$. Then $\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}$. $\square$

Proposition ??: Let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V.

Results about Hermitian forms

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent over $\mathbb{C}$. $\square$

Lemma ??: Let $v_1, \dots, v_n$ be an orthogonal sequence in V, and put $v = v_1 + \cdots + v_n$. Then $\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}$. $\square$

Proposition ??: Let $\mathcal{W} = w_1, \dots, w_p$ be an orthonormal sequence in V. Then for any $v \in V$ we have

$$\|v\|^2 \geq \sum_{i=1}^{p} |\langle v, w_i\rangle|^2.$$

Results about Hermitian forms

Let V be a vector space over $\mathbb{C}$ with a Hermitian form.

Theorem ?? (The Cauchy-Schwartz inequality):
For $v, w \in V$ we have $|\langle v, w\rangle| \leq \|v\| \, \|w\|$, with equality iff v and w are linearly dependent over $\mathbb{C}$. $\square$

Lemma ??: Let $v_1, \ldots, v_n$ be an orthogonal sequence in V, and put $v = v_1 + \cdots + v_n$. Then $\|v\| = \sqrt{\|v_1\|^2 + \cdots + \|v_n\|^2}$. $\square$

Proposition ??: Let $\mathcal{W} = w_1, \ldots, w_p$ be an orthonormal sequence in V. Then for any $v \in V$ we have

$$\|v\|^2 \geq \sum_{i=1}^{p} |\langle v, w_i\rangle|^2.$$

Moreover, this inequality is actually an equality iff $v \in \operatorname{span}(\mathcal{W})$. $\square$

Adjoints of linear maps

Definition ??:

Definition ??: Let V and W be real vector spaces with inner products

Definition ??: Let V and W be real vector spaces with inner products (or complex vector spaces with Hermitian forms).

Adjoints of linear maps

Definition ??: Let V and W be real vector spaces with inner products (or complex vector spaces with Hermitian forms). Let $\phi\colon V \to W$ and $\psi\colon W \to V$ be linear maps (over $\mathbb{R}$ or $\mathbb{C}$ as appropriate).

Adjoints of linear maps

Definition ??: Let V and W be real vector spaces with inner products (or complex vector spaces with Hermitian forms). Let $\phi\colon V \to W$ and $\psi\colon W \to V$ be linear maps (over $\mathbb{R}$ or $\mathbb{C}$ as appropriate). We say that ϕ is *adjoint* to ψ if we have

$$\langle \phi(v), w\rangle = \langle v, \psi(w)\rangle$$

for all $v \in V$ and $w \in W$.

Adjoints for matrices

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??:

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$.

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A.

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**.

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle$$

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T \mathbf{v}$$

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v}$$

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T \mathbf{v} \rangle$$

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} \colon \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T\mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T\mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle$$

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T\mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T\mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle,$$

as required.

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T \mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle,$$

as required.

Example ??:

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T\mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T\mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle,$$

as required.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{C}$

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} : \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T \mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle,$$

as required.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{C}$, giving a linear map $\phi_A \colon \mathbb{C}^m \to \mathbb{C}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$.

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} \colon \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T\mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T\mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle,$$

as required.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{C}$, giving a linear map $\phi_A \colon \mathbb{C}^m \to \mathbb{C}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. Let $A^\dagger$ be the complex conjugate of A^T.

Adjoints for matrices

This is essentially a basis-free formulation of the operation of transposing a matrix, as we see from the following example.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{R}$, giving a linear map $\phi_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. The transpose of A is then an $m \times n$ matrix A^T, giving a linear map $\phi_{A^T} \colon \mathbb{R}^n \to \mathbb{R}^m$. We claim that ϕ_{A^T} is adjoint to ϕ_A. This is easy to see using the formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ as in Remark **??**. Indeed, we have

$$\langle \phi_A(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u})^T\mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \langle \mathbf{u}, A^T\mathbf{v} \rangle = \langle \mathbf{u}, \phi_{A^T}(\mathbf{v}) \rangle,$$

as required.

Example ??: Let A be an $n \times m$ matrix over $\mathbb{C}$, giving a linear map $\phi_A \colon \mathbb{C}^m \to \mathbb{C}^n$ by $\phi_A(\mathbf{v}) = A\mathbf{v}$. Let $A^\dagger$ be the complex conjugate of A^T. Then $\phi_{A^\dagger}$ is adjoint to ϕ_A.

Cross products are anti self adjoint

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$.

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array}$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}]$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha \colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{matrix} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{matrix} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}] = -\det[\mathbf{a}|\mathbf{y}|\mathbf{x}]$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}] = -\det[\mathbf{a}|\mathbf{y}|\mathbf{x}] = -\langle \alpha(\mathbf{y}), \mathbf{x} \rangle$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}] = -\det[\mathbf{a}|\mathbf{y}|\mathbf{x}] = -\langle \alpha(\mathbf{y}), \mathbf{x} \rangle = \langle \mathbf{x}, -\alpha(\mathbf{y}) \rangle$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}] = -\det[\mathbf{a}|\mathbf{y}|\mathbf{x}] = -\langle \alpha(\mathbf{y}), \mathbf{x} \rangle = \langle \mathbf{x}, -\alpha(\mathbf{y}) \rangle$$

so $\alpha^T = -\alpha$.

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{matrix} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{matrix} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}] = -\det[\mathbf{a}|\mathbf{y}|\mathbf{x}] = -\langle \alpha(\mathbf{y}), \mathbf{x} \rangle = \langle \mathbf{x}, -\alpha(\mathbf{y}) \rangle$$

so $\alpha^T = -\alpha$. Alternatively, we have $\alpha = \phi_A$, where A is as found below:

$$\alpha(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \qquad \alpha(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \qquad \alpha(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

Cross products are anti self adjoint

Fix a vector $\mathbf{a} = [a_1, a_2, a_3]^T \in \mathbb{R}^3$, and define $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $\alpha(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. Then

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \left\langle \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right\rangle = \begin{array}{l} a_2x_3y_1 - a_3x_2y_1 + \\ a_3x_1y_2 - a_1x_3y_2 + \\ a_1x_2y_3 - a_2x_1y_3 \end{array} = \det \begin{bmatrix} a_1 & x_1 & y_1 \\ a_2 & x_2 & y_2 \\ a_3 & x_3 & y_3 \end{bmatrix}$$

It follows that

$$\langle \alpha(\mathbf{x}), \mathbf{y} \rangle = \det[\mathbf{a}|\mathbf{x}|\mathbf{y}] = -\det[\mathbf{a}|\mathbf{y}|\mathbf{x}] = -\langle \alpha(\mathbf{y}), \mathbf{x} \rangle = \langle \mathbf{x}, -\alpha(\mathbf{y}) \rangle$$

so $\alpha^T = -\alpha$. Alternatively, we have $\alpha = \phi_A$, where A is as found below:

$$\alpha(\mathbf{e}_1) = \begin{bmatrix} 0 \\ a_3 \\ -a_2 \end{bmatrix} \qquad \alpha(\mathbf{e}_2) = \begin{bmatrix} -a_3 \\ 0 \\ a_1 \end{bmatrix} \qquad \alpha(\mathbf{e}_3) = \begin{bmatrix} a_2 \\ -a_1 \\ 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

It follows that $\alpha^T = \phi_{A^T} = \phi_{-A} = -\alpha$.

Differentiation is anti self adjoint

Example ??:

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial.

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**.

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$.

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$.

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$.

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$. This is equivalent to the statement that for all f and g in V, we have $\langle D(f), g\rangle + \langle f, D(g)\rangle = 0$.

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$. This is equivalent to the statement that for all f and g in V, we have $\langle D(f), g\rangle + \langle f, D(g)\rangle = 0$. This is true because

$$\langle f', g\rangle + \langle f, g'\rangle = \int_{-\infty}^{\infty} f'(x)g(x) + f(x)g'(x)\,dx$$

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$. This is equivalent to the statement that for all f and g in V, we have $\langle D(f), g\rangle + \langle f, D(g)\rangle = 0$. This is true because

$$\langle f', g\rangle + \langle f, g'\rangle = \int_{-\infty}^{\infty} f'(x)g(x) + f(x)g'(x)\,dx = \int_{-\infty}^{\infty} \frac{d}{dx}(f(x)g(x))\,dx$$

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$. This is equivalent to the statement that for all f and g in V, we have $\langle D(f),g\rangle + \langle f,D(g)\rangle = 0$. This is true because

$$\langle f',g\rangle + \langle f,g'\rangle = \int_{-\infty}^{\infty} f'(x)g(x) + f(x)g'(x)\,dx = \int_{-\infty}^{\infty} \frac{d}{dx}(f(x)g(x))\,dx = [f(x)g(x)]_{-\infty}^{\infty}$$

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$. This is equivalent to the statement that for all f and g in V, we have $\langle D(f), g\rangle + \langle f, D(g)\rangle = 0$. This is true because

$$\begin{aligned}\langle f', g\rangle + \langle f, g'\rangle &= \int_{-\infty}^{\infty} f'(x)g(x) + f(x)g'(x)\,dx = \int_{-\infty}^{\infty} \frac{d}{dx}(f(x)g(x))\,dx = [f(x)g(x)]_{-\infty}^{\infty} \\ &= \lim_{x\to+\infty} f(x)g(x) - \lim_{x\to-\infty} f(x)g(x).\end{aligned}$$

Differentiation is anti self adjoint

Example ??: Let V be the set of functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is a polynomial. We use the inner product $\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx$, as in Example **??**. If we have a function $f(x) = p(x)e^{-x^2/2}$ in V, we note that

$$f'(x) = p'(x)e^{-x^2/2} + p(x).(-x).e^{-x^2/2} = (p'(x) - x\,p(x))e^{-x^2/2},$$

and $p'(x) - x\,p(x)$ is again a polynomial, so $f'(x) \in V$. We can thus define a linear map $D\colon V \to V$ by $D(f) = f'$. We claim that D is adjoint to $-D$. This is equivalent to the statement that for all f and g in V, we have $\langle D(f), g\rangle + \langle f, D(g)\rangle = 0$. This is true because

$$\begin{aligned}\langle f', g\rangle + \langle f, g'\rangle &= \int_{-\infty}^{\infty} f'(x)g(x) + f(x)g'(x)\,dx = \int_{-\infty}^{\infty} \frac{d}{dx}(f(x)g(x))\,dx = [f(x)g(x)]_{-\infty}^{\infty} \\ &= \lim_{x\to+\infty} f(x)g(x) - \lim_{x\to-\infty} f(x)g(x).\end{aligned}$$

Both limits here are zero, because the very rapid decrease of $e^{-x^2/2}$ wipes out the much slower increase of the polynomial terms.

An example

Example ??:

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$)

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product).

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi\begin{bmatrix} p \\ q \end{bmatrix} = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ.

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ. To check, consider $f(x) = ax^2 + bx + c \in \mathbb{R}[x]_{\leq 2}$ and $\mathbf{v} = \left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] \in \mathbb{R}^2$.

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ. To check, consider $f(x) = ax^2 + bx + c \in \mathbb{R}[x]_{\leq 2}$ and $\mathbf{v} = \left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] \in \mathbb{R}^2$. Note that $f(0) = c$ and $f(1) = a + b + c$, so $\phi(f) = \left[\begin{smallmatrix} c \\ a+b+c \end{smallmatrix}\right]$.

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi\begin{bmatrix} p \\ q \end{bmatrix} = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ. To check, consider $f(x) = ax^2 + bx + c \in \mathbb{R}[x]_{\leq 2}$ and $\mathbf{v} = \begin{bmatrix} p \\ q \end{bmatrix} \in \mathbb{R}^2$. Note that $f(0) = c$ and $f(1) = a + b + c$, so $\phi(f) = \begin{bmatrix} c \\ a+b+c \end{bmatrix}$. We must show that $\langle f, \psi(\mathbf{v})\rangle = \langle \phi(f), \mathbf{v}\rangle$

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ. To check, consider $f(x) = ax^2 + bx + c \in \mathbb{R}[x]_{\leq 2}$ and $\mathbf{v} = \left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] \in \mathbb{R}^2$. Note that $f(0) = c$ and $f(1) = a + b + c$, so $\phi(f) = \left[\begin{smallmatrix} c \\ a+b+c \end{smallmatrix}\right]$. We must show that $\langle f, \psi(\mathbf{v})\rangle = \langle \phi(f), \mathbf{v}\rangle$, or in other words that

$$\int_0^1 (ax^2 + bx + c)((30p + 30q)x^2 - (36p + 24q)x + (9p + 3q))\,dx = p\,f(0) + q\,f(1) = pc + q(a + b + c).$$

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad\qquad \psi\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ. To check, consider $f(x) = ax^2 + bx + c \in \mathbb{R}[x]_{\leq 2}$ and $\mathbf{v} = \left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] \in \mathbb{R}^2$. Note that $f(0) = c$ and $f(1) = a + b + c$, so $\phi(f) = \left[\begin{smallmatrix} c \\ a+b+c \end{smallmatrix}\right]$. We must show that $\langle f, \psi(\mathbf{v})\rangle = \langle \phi(f), \mathbf{v}\rangle$, or in other words that

$$\int_0^1 (ax^2 + bx + c)((30p + 30q)x^2 - (36p + 24q)x + (9p + 3q))\,dx = p\,f(0) + q\,f(1) = pc + q(a + b + c).$$

This can be done with Maple

An example

Example ??: Consider the vector spaces $\mathbb{R}[x]_{\leq 2}$ (with inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$) and $\mathbb{R}^2$ (with the usual inner product). Define maps $\phi\colon \mathbb{R}[x]_{\leq 2} \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}[x]_{\leq 2}$ by

$$\phi(f) = \begin{bmatrix} f(0) \\ f(1) \end{bmatrix} \qquad \psi \left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] = (30p + 30q)x^2 - (36p + 24q)x + (9p + 3q).$$

Claim: ϕ is adjoint to ψ. To check, consider $f(x) = ax^2 + bx + c \in \mathbb{R}[x]_{\leq 2}$ and $\mathbf{v} = \left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right] \in \mathbb{R}^2$. Note that $f(0) = c$ and $f(1) = a + b + c$, so $\phi(f) = \left[\begin{smallmatrix} c \\ a+b+c \end{smallmatrix}\right]$. We must show that $\langle f, \psi(\mathbf{v}) \rangle = \langle \phi(f), \mathbf{v} \rangle$, or in other words that

$$\int_0^1 (ax^2 + bx + c)((30p + 30q)x^2 - (36p + 24q)x + (9p + 3q))\,dx = p\,f(0) + q\,f(1) = pc + q(a + b + c).$$

This can be done with Maple: entering

```
expand(int((a*x^2+b*x+c)*((30*p+30*q)*x^2 - (36*p+24*q)*x + (9*p+3*q)),x=0..1));
```

gives $cp + aq + bq + cq$, as required.

Existence of adjoints

Proposition ??:

Proposition ??: Let V and W be finite-dimensional real vector spaces with inner products

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional real vector spaces with inner products (or complex vector spaces with Hermitian forms).

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional real vector spaces with inner products (or complex vector spaces with Hermitian forms). Let $\phi\colon V \to W$ be a linear map (over $\mathbb{R}$ or $\mathbb{C}$ as appropriate).

Then there is a unique map $\psi\colon W \to V$ that is adjoint to ϕ.

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional real vector spaces with inner products (or complex vector spaces with Hermitian forms). Let $\phi\colon V \to W$ be a linear map (over $\mathbb{R}$ or $\mathbb{C}$ as appropriate).

Then there is a unique map $\psi\colon W \to V$ that is adjoint to ϕ. (We write $\psi = \phi^*$ in the real case, or $\psi = \phi^\dagger$ in the complex case.)

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional real vector spaces with inner products (or complex vector spaces with Hermitian forms). Let $\phi\colon V \to W$ be a linear map (over $\mathbb{R}$ or $\mathbb{C}$ as appropriate).

Then there is a unique map $\psi\colon W \to V$ that is adjoint to ϕ. (We write $\psi = \phi^*$ in the real case, or $\psi = \phi^\dagger$ in the complex case.)

We will prove the complex case; the real case is similar but slightly easier.

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof:

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint.

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger : W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ, so

$$\langle v, \psi(w)\rangle = \langle \phi(v), w\rangle = \langle v, \psi'(w)\rangle$$

for all $v \in V$ and $w \in W$.

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ, so

$$\langle v, \psi(w)\rangle = \langle \phi(v), w\rangle = \langle v, \psi'(w)\rangle$$

for all $v \in V$ and $w \in W$. This means that $\langle v, \psi(w) - \psi'(w)\rangle = 0$ for all v and w.

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ, so

$$\langle v, \psi(w)\rangle = \langle \phi(v), w\rangle = \langle v, \psi'(w)\rangle$$

for all $v \in V$ and $w \in W$. This means that $\langle v, \psi(w) - \psi'(w)\rangle = 0$ for all v and w. In particular, we can take $v = \psi(w) - \psi'(w)$

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ, so

$$\langle v, \psi(w)\rangle = \langle \phi(v), w\rangle = \langle v, \psi'(w)\rangle$$

for all $v \in V$ and $w \in W$. This means that $\langle v, \psi(w) - \psi'(w)\rangle = 0$ for all v and w. In particular, we can take $v = \psi(w) - \psi'(w)$, and we find that

$$\|\psi(w) - \psi'(w)\|^2 = \langle \psi(w) - \psi'(w), \psi(w) - \psi'(w)\rangle = 0,$$

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^{\dagger}\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ, so

$$\langle v, \psi(w)\rangle = \langle \phi(v), w\rangle = \langle v, \psi'(w)\rangle$$

for all $v \in V$ and $w \in W$. This means that $\langle v, \psi(w) - \psi'(w)\rangle = 0$ for all v and w. In particular, we can take $v = \psi(w) - \psi'(w)$, and we find that

$$\|\psi(w) - \psi'(w)\|^2 = \langle \psi(w) - \psi'(w), \psi(w) - \psi'(w)\rangle = 0,$$

so $\psi(w) = \psi'(w)$ for all w

Existence of adjoints

Proposition ??: Let V and W be finite-dimensional complex vector spaces with Hermitian forms. Let $\phi\colon V \to W$ be a $\mathbb{C}$-linear maps.

Then there is a unique map $\psi = \phi^\dagger\colon W \to V$ that is adjoint to ϕ.

Proof: We first show that there is at most one adjoint. Suppose that ψ and ψ' are both adjoint to ϕ, so

$$\langle v, \psi(w)\rangle = \langle \phi(v), w\rangle = \langle v, \psi'(w)\rangle$$

for all $v \in V$ and $w \in W$. This means that $\langle v, \psi(w) - \psi'(w)\rangle = 0$ for all v and w. In particular, we can take $v = \psi(w) - \psi'(w)$, and we find that

$$\|\psi(w) - \psi'(w)\|^2 = \langle \psi(w) - \psi'(w), \psi(w) - \psi'(w)\rangle = 0,$$

so $\psi(w) = \psi'(w)$ for all w, so $\psi = \psi'$.

Existence of adjoints

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$.

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V, and define a linear map $\psi\colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle = \sum_j \langle \phi(v_j), w \rangle \langle v_i, v_j \rangle$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi\colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j)\rangle v_j.$$

Recall that $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$, and that $\overline{\langle x, y\rangle} = \langle y, x\rangle$. Using these rules we find that

$$\langle v_i, \psi(w)\rangle = \sum_j \langle v_i, \langle w, \phi(v_j)\rangle v_j\rangle = \sum_j \overline{\langle w, \phi(v_j)\rangle}\langle v_i, v_j\rangle = \sum_j \langle \phi(v_j), w\rangle \langle v_i, v_j\rangle = \langle \phi(v_i), w\rangle .$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j .$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle = \sum_j \langle \phi(v_j), w \rangle \langle v_i, v_j \rangle = \langle \phi(v_i), w \rangle .$$

More generally, any element $v \in V$ can be written as $\sum_i x_i v_i$ for some $x_1, \dots, x_n \in \mathbb{C}$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle = \sum_j \langle \phi(v_j), w \rangle \langle v_i, v_j \rangle = \langle \phi(v_i), w \rangle.$$

More generally, any element $v \in V$ can be written as $\sum_i x_i v_i$ for some $x_1, \ldots, x_n \in \mathbb{C}$, and then we have

$$\langle v, \psi(w) \rangle = \sum_i x_i \langle v_i, \psi(w) \rangle$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle = \sum_j \langle \phi(v_j), w \rangle \langle v_i, v_j \rangle = \langle \phi(v_i), w \rangle.$$

More generally, any element $v \in V$ can be written as $\sum_i x_i v_i$ for some $x_1, \dots, x_n \in \mathbb{C}$, and then we have

$$\langle v, \psi(w) \rangle = \sum_i x_i \langle v_i, \psi(w) \rangle = \sum_i x_i \langle \phi(v_i), w \rangle$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi\colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j)\rangle v_j.$$

Recall that $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$, and that $\overline{\langle x, y\rangle} = \langle y, x\rangle$. Using these rules we find that

$$\langle v_i, \psi(w)\rangle = \sum_j \langle v_i, \langle w, \phi(v_j)\rangle v_j\rangle = \sum_j \overline{\langle w, \phi(v_j)\rangle}\langle v_i, v_j\rangle = \sum_j \langle \phi(v_j), w\rangle \langle v_i, v_j\rangle = \langle \phi(v_i), w\rangle.$$

More generally, any element $v \in V$ can be written as $\sum_i x_i v_i$ for some $x_1, \dots, x_n \in \mathbb{C}$, and then we have

$$\langle v, \psi(w)\rangle = \sum_i x_i \langle v_i, \psi(w)\rangle = \sum_i x_i \langle \phi(v_i), w\rangle = \langle \phi\left(\sum_i x_i v_i\right), w\rangle$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle = \sum_j \langle \phi(v_j), w \rangle \langle v_i, v_j \rangle = \langle \phi(v_i), w \rangle .$$

More generally, any element $v \in V$ can be written as $\sum_i x_i v_i$ for some $x_1, \dots, x_n \in \mathbb{C}$, and then we have

$$\langle v, \psi(w) \rangle = \sum_i x_i \langle v_i, \psi(w) \rangle = \sum_i x_i \langle \phi(v_i), w \rangle = \langle \phi \left(\sum_i x_i v_i \right), w \rangle = \langle \phi(v), w \rangle .$$

Existence of adjoints

To show that there exists an adjoint, choose an orthonormal basis $\mathcal{V} = v_1, \dots, v_n$ for V, and define a linear map $\psi \colon W \to V$ by

$$\psi(w) = \sum_{j=1}^{n} \langle w, \phi(v_j) \rangle v_j.$$

Recall that $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$, and that $\overline{\langle x, y \rangle} = \langle y, x \rangle$. Using these rules we find that

$$\langle v_i, \psi(w) \rangle = \sum_j \langle v_i, \langle w, \phi(v_j) \rangle v_j \rangle = \sum_j \overline{\langle w, \phi(v_j) \rangle} \langle v_i, v_j \rangle = \sum_j \langle \phi(v_j), w \rangle \langle v_i, v_j \rangle = \langle \phi(v_i), w \rangle.$$

More generally, any element $v \in V$ can be written as $\sum_i x_i v_i$ for some $x_1, \dots, x_n \in \mathbb{C}$, and then we have

$$\langle v, \psi(w) \rangle = \sum_i x_i \langle v_i, \psi(w) \rangle = \sum_i x_i \langle \phi(v_i), w \rangle = \langle \phi \left(\sum_i x_i v_i \right), w \rangle = \langle \phi(v), w \rangle.$$

This shows that ψ is adjoint to ϕ, as required.

Periodic functions

Definition ??: We say that a function $f : \mathbb{R} \to \mathbb{C}$ is *periodic* if $f(t + 2\pi) = f(t)$ for all $t \in \mathbb{R}$.

Periodic functions

Definition ??: We say that a function $f\colon \mathbb{R} \to \mathbb{C}$ is *periodic* if $f(t+2\pi) = f(t)$ for all $t \in \mathbb{R}$. We let P be the set of all continuous periodic functions from $\mathbb{R}$ to $\mathbb{C}$, which is a vector space over $\mathbb{C}$.

Periodic functions

Definition ??: We say that a function $f \colon \mathbb{R} \to \mathbb{C}$ is *periodic* if $f(t + 2\pi) = f(t)$ for all $t \in \mathbb{R}$. We let P be the set of all continuous periodic functions from $\mathbb{R}$ to $\mathbb{C}$, which is a vector space over $\mathbb{C}$. We define a Hermitian form on P by

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}\, dt.$$

Definition ??: We say that a function $f\colon \mathbb{R} \to \mathbb{C}$ is *periodic* if $f(t+2\pi) = f(t)$ for all $t \in \mathbb{R}$. We let P be the set of all continuous periodic functions from $\mathbb{R}$ to $\mathbb{C}$, which is a vector space over $\mathbb{C}$. We define a Hermitian form on P by

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt.$$

Some important elements of P are the functions e_n (for $n \in \mathbb{Z}$), s_n (for $n > 0$) and c_n (for $n \geq 0$) defined as follows:

$$e_n(t) = \exp(int) \qquad s_n(t) = \sin(nt) \qquad c_n(t) = \cos(nt)$$

Definition ??: We say that a function $f\colon \mathbb{R} \to \mathbb{C}$ is *periodic* if $f(t+2\pi) = f(t)$ for all $t \in \mathbb{R}$. We let P be the set of all continuous periodic functions from $\mathbb{R}$ to $\mathbb{C}$, which is a vector space over $\mathbb{C}$. We define a Hermitian form on P by

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt.$$

Some important elements of P are the functions e_n (for $n \in \mathbb{Z}$), s_n (for $n > 0$) and c_n (for $n \geq 0$) defined as follows:

$$e_n(t) = \exp(int) \qquad s_n(t) = \sin(nt) \qquad c_n(t) = \cos(nt)$$

De Moivre's theorem tells us that

$$\begin{aligned} e_n &= c_n + i\,s_n \\ s_n &= (e_n - e_{-n})/(2i) \\ c_n &= (e_n + e_{-n})/2. \end{aligned}$$

Trigonometric polynomials

Trigonometric polynomials

Definition ??:

Definition ??: We put

$$T_n = \mathrm{span}(\{e_k \mid -n \leq k \leq n\}) \leq P$$

Trigonometric polynomials

Definition ??: We put

$$T_n = \mathrm{span}(\{e_k \mid -n \leq k \leq n\}) \leq P,$$

and note that $T_n \leq T_{n+1}$ for all n.

Definition ??: We put

$$T_n = \mathrm{span}(\{e_k \mid -n \leq k \leq n\}) \leq P,$$

and note that $T_n \leq T_{n+1}$ for all n. We also let T denote the span of all the e_k's, or equivalently, the union of all the sets T_n.

Definition ??: We put

$$T_n = \operatorname{span}(\{e_k \mid -n \leq k \leq n\}) \leq P,$$

and note that $T_n \leq T_{n+1}$ for all n. We also let T denote the span of all the e_k's, or equivalently, the union of all the sets T_n. The elements of T are the functions $f\colon \mathbb{R} \to \mathbb{C}$ that can be written in the form

$$f(t) = \sum_{k=-n}^{n} a_k e_k(t) = \sum_{k=-n}^{n} a_k \exp(ikt)$$

for some $n > 0$ and some coefficients $a_{-n}, \ldots, a_n \in \mathbb{C}$.

Trigonometric polynomials

Definition ??: We put

$$T_n = \operatorname{span}(\{e_k \mid -n \leq k \leq n\}) \leq P,$$

and note that $T_n \leq T_{n+1}$ for all n. We also let T denote the span of all the e_k's, or equivalently, the union of all the sets T_n. The elements of T are the functions $f\colon \mathbb{R} \to \mathbb{C}$ that can be written in the form

$$f(t) = \sum_{k=-n}^{n} a_k e_k(t) = \sum_{k=-n}^{n} a_k \exp(ikt)$$

for some $n > 0$ and some coefficients $a_{-n}, \ldots, a_n \in \mathbb{C}$. Functions of this form are called *trigonometric polynomials* or *finite Fourier series*.

An orthonormal basis

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof:

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\langle e_k , e_m \rangle = \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\langle e_k, e_m \rangle = \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_m(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(ikt)\exp(-imt)\,dt$$

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n + 1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt) \exp(-imt)\, dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k - m)t)\, dt\end{aligned}$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}\langle e_k, e_m\rangle &= \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_m(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(ikt)\exp(-imt)\,dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} \exp(i(k-m)t)\,dt = \frac{1}{2\pi}\left[\frac{\exp(i(k-m)t)}{i(k-m)}\right]_0^{2\pi}\end{aligned}$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_m(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(ikt)\exp(-imt)\,dt \\
&= \frac{1}{2\pi}\int_0^{2\pi} \exp(i(k-m)t)\,dt = \frac{1}{2\pi}\left[\frac{\exp(i(k-m)t)}{i(k-m)}\right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i}\left(e^{2(k-m)\pi i} - 1\right).
\end{aligned}$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt)\exp(-imt)\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k-m)t)\, dt = \frac{1}{2\pi} \left[\frac{\exp(i(k-m)t)}{i(k-m)} \right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i} \left(e^{2(k-m)\pi i} - 1 \right).
\end{aligned}$$

As $k - m$ is an integer, we have $e^{2(k-m)\pi i} = 1$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt)\exp(-imt)\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k-m)t)\, dt = \frac{1}{2\pi} \left[\frac{\exp(i(k-m)t)}{i(k-m)} \right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i} \left(e^{2(k-m)\pi i} - 1 \right).
\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$.

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt)\exp(-imt)\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k-m)t)\, dt = \frac{1}{2\pi} \left[\frac{\exp(i(k-m)t)}{i(k-m)} \right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i} \left(e^{2(k-m)\pi i} - 1 \right).
\end{aligned}$$

As $k - m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal.

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt)\exp(-imt)\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k-m)t)\, dt = \frac{1}{2\pi} \left[\frac{\exp(i(k-m)t)}{i(k-m)} \right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i} \left(e^{2(k-m)\pi i} - 1 \right).
\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal. We also have

$$\langle e_k, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_k(t)}\, dt$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_m(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(ikt)\exp(-imt)\,dt \\
&= \frac{1}{2\pi}\int_0^{2\pi} \exp(i(k-m)t)\,dt = \frac{1}{2\pi}\left[\frac{\exp(i(k-m)t)}{i(k-m)}\right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i}\left(e^{2(k-m)\pi i} - 1\right).
\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal. We also have

$$\langle e_k, e_k \rangle = \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_k(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(2k\pi it)\exp(-2k\pi it)\,dt$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt) \exp(-imt)\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k-m)t)\, dt = \frac{1}{2\pi} \left[\frac{\exp(i(k-m)t)}{i(k-m)} \right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i} \left(e^{2(k-m)\pi i} - 1 \right).
\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal. We also have

$$\langle e_k, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_k(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(2k\pi it) \exp(-2k\pi it)\, dt = \frac{1}{2\pi} \int_0^{2\pi} 1\, dt = 1.$$

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \dots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_m(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(ikt)\exp(-imt)\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \exp(i(k-m)t)\, dt = \frac{1}{2\pi} \left[\frac{\exp(i(k-m)t)}{i(k-m)} \right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i} \left(e^{2(k-m)\pi i} - 1 \right).
\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal. We also have

$$\langle e_k, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e_k(t)\overline{e_k(t)}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \exp(2k\pi it)\exp(-2k\pi it)\, dt = \frac{1}{2\pi} \int_0^{2\pi} 1\, dt = 1.$$

Our sequence is therefore orthonormal, and so linearly independent.

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}
\langle e_k, e_m \rangle &= \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_m(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(ikt)\exp(-imt)\,dt \\
&= \frac{1}{2\pi}\int_0^{2\pi} \exp(i(k-m)t)\,dt = \frac{1}{2\pi}\left[\frac{\exp(i(k-m)t)}{i(k-m)}\right]_0^{2\pi} \\
&= \frac{1}{2(k-m)\pi i}\left(e^{2(k-m)\pi i} - 1\right).
\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal. We also have

$$\langle e_k, e_k \rangle = \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_k(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(2k\pi it)\exp(-2k\pi it)\,dt = \frac{1}{2\pi}\int_0^{2\pi} 1\,dt = 1.$$

Our sequence is therefore orthonormal, and so linearly independent. It also spans T_n (by the definition of T_n)

An orthonormal basis

Proposition ??: The sequence $e_{-n}, e_{-n+1}, \ldots, e_{n-1}, e_n$ is an orthonormal basis for T_n (so $\dim(T_n) = 2n+1$).
Proof: For $m \neq k$ we have

$$\begin{aligned}\langle e_k, e_m \rangle &= \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_m(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(ikt)\exp(-imt)\,dt \\ &= \frac{1}{2\pi}\int_0^{2\pi} \exp(i(k-m)t)\,dt = \frac{1}{2\pi}\left[\frac{\exp(i(k-m)t)}{i(k-m)}\right]_0^{2\pi} \\ &= \frac{1}{2(k-m)\pi i}\left(e^{2(k-m)\pi i} - 1\right).\end{aligned}$$

As $k-m$ is an integer, we have $e^{2(k-m)\pi i} = 1$ and so $\langle e_k, e_m \rangle = 0$. This shows that the sequence of e_k's is orthogonal. We also have

$$\langle e_k, e_k \rangle = \frac{1}{2\pi}\int_0^{2\pi} e_k(t)\overline{e_k(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \exp(2k\pi it)\exp(-2k\pi it)\,dt = \frac{1}{2\pi}\int_0^{2\pi} 1\,dt = 1.$$

Our sequence is therefore orthonormal, and so linearly independent. It also spans T_n (by the definition of T_n), so it is a basis.

Projection onto T_n

Projection onto T_n

Definition ??: For any $f \in P$, let $\pi_n(f)$ be the orthogonal projection of f in T_n

Projection onto T_n

Definition ??: For any $f \in P$, let $\pi_n(f)$ be the orthogonal projection of f in T_n, so

$$\pi_n(f) = \sum_{m=-n}^{n} \langle f, e_m \rangle e_m.$$

Projection onto T_n

Definition ??: For any $f \in P$, let $\pi_n(f)$ be the orthogonal projection of f in T_n, so

$$\pi_n(f) = \sum_{m=-n}^{n} \langle f, e_m \rangle e_m.$$

We also put $\epsilon_n(f) = f - \pi_n(f)$

Projection onto T_n

Definition ??: For any $f \in P$, let $\pi_n(f)$ be the orthogonal projection of f in T_n, so

$$\pi_n(f) = \sum_{m=-n}^{n} \langle f, e_m \rangle e_m.$$

We also put $\epsilon_n(f) = f - \pi_n(f)$, so $f = \pi_n(f) + \epsilon_n(f)$, with $\pi_n(f) \in T_n$ and $\epsilon_n(f) \in T_n^{\perp}$ (by Proposition **??**).

Another orthogonal basis

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n.

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof:

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}.

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}. It follows that

$$\langle s_m, s_k \rangle = \langle s_m, c_k \rangle = \langle c_m, s_k \rangle = \langle c_m, c_k \rangle = 0.$$

Now suppose that $0 < m \leq n$, so c_m and s_m are both in $\mathcal{C}_n$.

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \dots, c_n, s_1, \dots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}. It follows that

$$\langle s_m, s_k \rangle = \langle s_m, c_k \rangle = \langle c_m, s_k \rangle = \langle c_m, c_k \rangle = 0.$$

Now suppose that $0 < m \leq n$, so c_m and s_m are both in $\mathcal{C}_n$. We have $\langle e_m, e_{-m} \rangle = 0$, and so

$$\langle s_m, c_m \rangle = \tfrac{1}{4i} \langle e_m - e_{-m}, e_m + e_{-m} \rangle$$

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \dots, c_n, s_1, \dots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}. It follows that

$$\langle s_m, s_k \rangle = \langle s_m, c_k \rangle = \langle c_m, s_k \rangle = \langle c_m, c_k \rangle = 0.$$

Now suppose that $0 < m \leq n$, so c_m and s_m are both in $\mathcal{C}_n$. We have $\langle e_m, e_{-m} \rangle = 0$, and so

$$\langle s_m, c_m \rangle = \frac{1}{4i} \langle e_m - e_{-m}, e_m + e_{-m} \rangle = \frac{1}{4i} (\langle e_m, e_m \rangle + \langle e_m, e_{-m} \rangle + -\langle e_{-m}, e_m \rangle - \langle e_{-m}, e_{-m} \rangle)$$

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}. It follows that

$$\langle s_m, s_k \rangle = \langle s_m, c_k \rangle = \langle c_m, s_k \rangle = \langle c_m, c_k \rangle = 0.$$

Now suppose that $0 < m \leq n$, so c_m and s_m are both in $\mathcal{C}_n$. We have $\langle e_m, e_{-m} \rangle = 0$, and so

$$\begin{aligned} \langle s_m, c_m \rangle &= \tfrac{1}{4i} \langle e_m - e_{-m}, e_m + e_{-m} \rangle = \tfrac{1}{4i} (\langle e_m, e_m \rangle + \langle e_m, e_{-m} \rangle + -\langle e_{-m}, e_m \rangle - \langle e_{-m}, e_{-m} \rangle) \\ &= \tfrac{1}{4i} (1 + 0 - 0 - 1) \end{aligned}$$

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \ldots, c_n, s_1, \ldots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}. It follows that

$$\langle s_m, s_k \rangle = \langle s_m, c_k \rangle = \langle c_m, s_k \rangle = \langle c_m, c_k \rangle = 0.$$

Now suppose that $0 < m \leq n$, so c_m and s_m are both in $\mathcal{C}_n$. We have $\langle e_m, e_{-m} \rangle = 0$, and so

$$\begin{aligned}\langle s_m, c_m \rangle &= \tfrac{1}{4i} \langle e_m - e_{-m}, e_m + e_{-m} \rangle = \tfrac{1}{4i} (\langle e_m, e_m \rangle + \langle e_m, e_{-m} \rangle + -\langle e_{-m}, e_m \rangle - \langle e_{-m}, e_{-m} \rangle) \\ &= \tfrac{1}{4i} (1 + 0 - 0 - 1) = 0.\end{aligned}$$

Another orthogonal basis

Proposition ??: The sequence $\mathcal{C}_n = c_0, c_1, \dots, c_n, s_1, \dots, s_n$ is another orthogonal basis for T_n. It is not orthonormal, but instead satisfies $\|s_k\|^2 = 1/2 = \|c_k\|^2$ for $k > 0$, and $\|c_0\|^2 = 1$.
Proof: We use $s_m = (e_m - e_{-m})/(2i)$ and $c_m = (e_m + e_{-m})/2$.

If $k \neq m$ (with $k, m \geq 0$) we see that e_k and e_{-k} are orthogonal to e_m and e_{-m}. It follows that

$$\langle s_m, s_k\rangle = \langle s_m, c_k\rangle = \langle c_m, s_k\rangle = \langle c_m, c_k\rangle = 0.$$

Now suppose that $0 < m \leq n$, so c_m and s_m are both in $\mathcal{C}_n$. We have $\langle e_m, e_{-m}\rangle = 0$, and so

$$\begin{aligned}\langle s_m, c_m\rangle &= \tfrac{1}{4i}\langle e_m - e_{-m}, e_m + e_{-m}\rangle = \tfrac{1}{4i}(\langle e_m, e_m\rangle + \langle e_m, e_{-m}\rangle + -\langle e_{-m}, e_m\rangle - \langle e_{-m}, e_{-m}\rangle) \\ &= \tfrac{1}{4i}(1 + 0 - 0 - 1) = 0.\end{aligned}$$

This shows that $\mathcal{C}_n$ is an orthogonal sequence.

Another orthogonal basis

Another orthogonal basis

For $k > 0$ we have

$$\langle s_k, s_k \rangle = \frac{1}{2i}\overline{\frac{1}{2i}}\langle e_k - e_{-k}, e_k - e_{-k} \rangle$$

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned}\langle s_k, s_k\rangle &= \frac{1}{2i}\overline{\frac{1}{2i}}\langle e_k - e_{-k}, e_k - e_{-k}\rangle \\ &= \tfrac{1}{4}(1 - 0 - 0 + 1)\end{aligned}$$

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned} \langle s_k, s_k \rangle &= \frac{1}{2i}\overline{\frac{1}{2i}}\langle e_k - e_{-k}, e_k - e_{-k} \rangle \\ &= \tfrac{1}{4}(1 - 0 - 0 + 1) = 1/2. \end{aligned}$$

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned}\langle s_k, s_k\rangle &= \frac{1}{2i}\overline{\frac{1}{2i}}\langle e_k - e_{-k}, e_k - e_{-k}\rangle \\ &= \tfrac{1}{4}(1 - 0 - 0 + 1) = 1/2.\end{aligned}$$

Similarly, we have $\langle c_k, c_k\rangle = 1/2$.

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned}\langle s_k, s_k\rangle &= \frac{1}{2i}\overline{\frac{1}{2i}}\langle e_k - e_{-k}, e_k - e_{-k}\rangle \\ &= \tfrac{1}{4}(1 - 0 - 0 + 1) = 1/2.\end{aligned}$$

Similarly, we have $\langle c_k, c_k\rangle = 1/2$. In the special case $k = 0$ we instead have $c_0(t) = 1$ for all t

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned}\langle s_k, s_k\rangle &= \frac{1}{2i}\frac{1}{\overline{2i}}\langle e_k - e_{-k}, e_k - e_{-k}\rangle \\ &= \tfrac{1}{4}(1 - 0 - 0 + 1) = 1/2.\end{aligned}$$

Similarly, we have $\langle c_k, c_k\rangle = 1/2$. In the special case $k = 0$ we instead have $c_0(t) = 1$ for all t, so $\langle c_0, c_0\rangle = (2\pi)^{-1}\int_0^{2\pi} 1\, dt = 1$.

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned}\langle s_k, s_k\rangle &= \frac{1}{2i}\frac{1}{\overline{2i}}\langle e_k - e_{-k}, e_k - e_{-k}\rangle \\ &= \tfrac{1}{4}(1-0-0+1) = 1/2.\end{aligned}$$

Similarly, we have $\langle c_k, c_k\rangle = 1/2$. In the special case $k = 0$ we instead have $c_0(t) = 1$ for all t, so $\langle c_0, c_0\rangle = (2\pi)^{-1}\int_0^{2\pi} 1\, dt = 1$. This completes the proof.

Another orthogonal basis

For $k > 0$ we have

$$\begin{aligned}\langle s_k, s_k \rangle &= \frac{1}{2i}\frac{1}{\overline{2i}}\langle e_k - e_{-k}, e_k - e_{-k}\rangle \\ &= \tfrac{1}{4}(1 - 0 - 0 + 1) = 1/2.\end{aligned}$$

Similarly, we have $\langle c_k, c_k \rangle = 1/2$. In the special case $k = 0$ we instead have $c_0(t) = 1$ for all t, so $\langle c_0, c_0 \rangle = (2\pi)^{-1}\int_0^{2\pi} 1\,dt = 1$. This completes the proof.

Corollary ??: Using Proposition **??**, we deduce that

$$\pi_n(f) = \langle f, c_0 \rangle c_0 + 2\sum_{k=1}^{n}\langle f, c_k \rangle c_k + 2\sum_{k=1}^{n}\langle f, s_k \rangle s_k.$$

The L^2 convergence theorem

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

(The proof is not examinable and will not be covered in lectures.)

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

(The proof is not examinable and will not be covered in lectures.)

Remark ??:

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

(The proof is not examinable and will not be covered in lectures.)

Remark ??: Recall that $\pi_n(f)$ is the closes point to f lying in T_n

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

(The proof is not examinable and will not be covered in lectures.)

Remark ??: Recall that $\pi_n(f)$ is the closes point to f lying in T_n, so the number $\|\epsilon_n(f)\| = \|f - \pi_n(f)\|$ can be regarded as the distance from f to T_n.

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

(The proof is not examinable and will not be covered in lectures.)

Remark ??: Recall that $\pi_n(f)$ is the closes point to f lying in T_n, so the number $\|\epsilon_n(f)\| = \|f - \pi_n(f)\|$ can be regarded as the distance from f to T_n. The theorem says that by taking n to be sufficiently large, we can make this distance as small as we like.

The L^2 convergence theorem

Theorem ??: For any $f \in P$ we have $\|\epsilon_n(f)\| \to 0$ as $n \to \infty$.

Proof: See Appendix **??**.

(The proof is not examinable and will not be covered in lectures.)

Remark ??: Recall that $\pi_n(f)$ is the closes point to f lying in T_n, so the number $\|\epsilon_n(f)\| = \|f - \pi_n(f)\|$ can be regarded as the distance from f to T_n. The theorem says that by taking n to be sufficiently large, we can make this distance as small as we like. In other words, f can be very well approximated by a trigonometric polynomial of sufficiently high degree.

Parseval's theorem

Parseval's theorem

Corollary ??: For any $f \in P$ we have

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k \rangle|^2$$

Parseval's theorem

Corollary ??: For any $f \in P$ we have

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k \rangle|^2 = |\langle f, c_0 \rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, c_k \rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, s_k \rangle|^2$$

Parseval's theorem

Corollary ??: For any $f \in P$ we have

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k \rangle|^2 = |\langle f, c_0 \rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, c_k \rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, s_k \rangle|^2$$

Proof:

Parseval's theorem

Corollary ??: For any $f \in P$ we have

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k \rangle|^2 = |\langle f, c_0 \rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, c_k \rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, s_k \rangle|^2$$

Proof: As $e_{-n}, \ldots, e_n$ is an orthonormal basis for T_n

Parseval's theorem

Corollary ??: For any $f \in P$ we have

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k\rangle|^2 = |\langle f, c_0\rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, c_k\rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, s_k\rangle|^2$$

Proof: As $e_{-n}, \dots, e_n$ is an orthonormal basis for T_n, we have

$$\|f\|^2 - \|\epsilon_n(f)\|^2 = \|\pi_n(f)\|^2 = \|\sum_{k=-n}^{n} \langle f, e_k\rangle e_k\|^2 = \sum_{k=-n}^{n} |\langle f, e_k\rangle|^2$$

Corollary ??: For any $f \in P$ we have

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k\rangle|^2 = |\langle f, c_0\rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, c_k\rangle|^2 + 2\sum_{k=1}^{\infty} |\langle f, s_k\rangle|^2$$

Proof: As $e_{-n}, \dots, e_n$ is an orthonormal basis for T_n, we have

$$\|f\|^2 - \|\epsilon_n(f)\|^2 = \|\pi_n(f)\|^2 = \|\sum_{k=-n}^{n} \langle f, e_k\rangle e_k\|^2 = \sum_{k=-n}^{n} |\langle f, e_k\rangle|^2$$

By taking limits as n tends to infinity, we see that $\|f\|^2 = \sum_{k=-\infty}^{\infty} |\langle f, e_k\rangle|^2$.

Parseval's theorem

Parseval's theorem

Similarly, using Corollary **??** and Proposition **??**, we see that

$$\|\pi_n(f)\|^2 = |\langle f, c_0\rangle|^2 \|c_0\|^2 + \sum_{k=1}^{n} 4|\langle f, c_k\rangle|^2 \|c_k\|^2 + \sum_{k=1}^{n} 4|\langle f, s_k\rangle|^2 \|s_k\|^2$$

Parseval's theorem

Similarly, using Corollary **??** and Proposition **??**, we see that

$$\begin{aligned}\|\pi_n(f)\|^2 &= |\langle f, c_0\rangle|^2\|c_0\|^2 + \sum_{k=1}^{n} 4|\langle f, c_k\rangle|^2\|c_k\|^2 + \sum_{k=1}^{n} 4|\langle f, s_k\rangle|^2\|s_k\|^2 \\ &= |\langle f, c_0\rangle|^2 + 2\sum_{k=1}^{n} |\langle f, c_k\rangle|^2 + 2\sum_{k=1}^{n} |\langle f, s_k\rangle|^2\end{aligned}$$

Parseval's theorem

Similarly, using Corollary **??** and Proposition **??**, we see that

$$\begin{aligned}\|\pi_n(f)\|^2 &= |\langle f, c_0\rangle|^2\|c_0\|^2 + \sum_{k=1}^{n} 4|\langle f, c_k\rangle|^2\|c_k\|^2 + \sum_{k=1}^{n} 4|\langle f, s_k\rangle|^2\|s_k\|^2 \\ &= |\langle f, c_0\rangle|^2 + 2\sum_{k=1}^{n}|\langle f, c_k\rangle|^2 + 2\sum_{k=1}^{n}|\langle f, s_k\rangle|^2\end{aligned}$$

We can again let n tend to infinity to see that

$$\|f\|^2 = |\langle f, c_0\rangle|^2 + 2\sum_{k=1}^{\infty}|\langle f, c_k\rangle|^2 + 2\sum_{k=1}^{\infty}|\langle f, s_k\rangle|^2.$$

Self-adjoint operators

Definition ??: Let V be a finite-dimensional vector space over $\mathbb{C}$. A *self-adjoint operator* on V is a linear map $\alpha\colon V \to V$ such that $\alpha^\dagger = \alpha$.

Eigenvalues are real

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof:

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$.

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle$$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle$$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^\dagger(v)\rangle$$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^\dagger(v)\rangle = \langle v, \alpha(v)\rangle$$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^\dagger(v)\rangle = \langle v, \alpha(v)\rangle = \langle v, \lambda v\rangle$$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^\dagger(v)\rangle = \langle v, \alpha(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle.$$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^{\dagger}(v)\rangle = \langle v, \alpha(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle.$$

As $v \neq 0$ we have $\langle v, v\rangle > 0$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^{\dagger}(v)\rangle = \langle v, \alpha(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle.$$

As $v \neq 0$ we have $\langle v, v\rangle > 0$, so we can divide by this to see that $\lambda = \overline{\lambda}$

Eigenvalues are real

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then every eigenvalue of α is real.

Proof: First suppose that λ is an eigenvalue of α, so there exists a nonzero vector $v \in V$ with $\alpha(v) = \lambda v$. We then have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle \alpha(v), v\rangle = \langle v, \alpha^{\dagger}(v)\rangle = \langle v, \alpha(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle.$$

As $v \neq 0$ we have $\langle v, v\rangle > 0$, so we can divide by this to see that $\lambda = \overline{\lambda}$, which means that λ is real.

The diagonalisation theorem

Theorem ??: If $\alpha\colon V \to V$ is a self-adjoint operator, then one can choose an orthonormal basis $\mathcal{V} = v_1, \ldots, v_n$ for V such that each v_i is an eigenvector of α.

A useful lemma

A useful lemma

Lemma ??:

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$

A useful lemma

Lemma ??: Let $\alpha \colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$).

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^{\perp}) \leq W^{\perp}$.

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^{\perp}) \leq W^{\perp}$.

Proof:

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^{\perp}) \leq W^{\perp}$.

Proof: Suppose that $v \in W^{\perp}$; we must show that $\alpha(v)$ is also in $W^{\perp}$.

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w\rangle = \langle v, \alpha^\dagger(w)\rangle$

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w\rangle = \langle v, \alpha^\dagger(w)\rangle = \langle v, \alpha(w)\rangle$

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w \rangle = \langle v, \alpha^\dagger(w) \rangle = \langle v, \alpha(w) \rangle$ (by the definition of adjoints and the fact that $\alpha^\dagger = \alpha$).

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w \rangle = \langle v, \alpha^\dagger(w) \rangle = \langle v, \alpha(w) \rangle$ (by the definition of adjoints and the fact that $\alpha^\dagger = \alpha$). As $\alpha(W) \leq W$ we see that $\alpha(w) \in W$

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w \rangle = \langle v, \alpha^\dagger(w) \rangle = \langle v, \alpha(w) \rangle$ (by the definition of adjoints and the fact that $\alpha^\dagger = \alpha$). As $\alpha(W) \leq W$ we see that $\alpha(w) \in W$, so $\langle v, \alpha(w) \rangle = 0$ (because $v \in W^\perp$).

A useful lemma

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^\perp) \leq W^\perp$.

Proof: Suppose that $v \in W^\perp$; we must show that $\alpha(v)$ is also in $W^\perp$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w\rangle = \langle v, \alpha^\dagger(w)\rangle = \langle v, \alpha(w)\rangle$ (by the definition of adjoints and the fact that $\alpha^\dagger = \alpha$). As $\alpha(W) \leq W$ we see that $\alpha(w) \in W$, so $\langle v, \alpha(w)\rangle = 0$ (because $v \in W^\perp$). We conclude that $\langle \alpha(v), w\rangle = 0$ for all $w \in W$

Lemma ??: Let $\alpha\colon V \to V$ be a self-adjoint operator, and let $W \leq V$ be a subspace such that $\alpha(W) \leq W$ (ie $\alpha(w) \in W$ for all $w \in W$). Then $\alpha(W^{\perp}) \leq W^{\perp}$.

Proof: Suppose that $v \in W^{\perp}$; we must show that $\alpha(v)$ is also in $W^{\perp}$. To see this, consider $w \in W$, and note that $\langle \alpha(v), w \rangle = \langle v, \alpha^{\dagger}(w) \rangle = \langle v, \alpha(w) \rangle$ (by the definition of adjoints and the fact that $\alpha^{\dagger} = \alpha$). As $\alpha(W) \leq W$ we see that $\alpha(w) \in W$, so $\langle v, \alpha(w) \rangle = 0$ (because $v \in W^{\perp}$). We conclude that $\langle \alpha(v), w \rangle = 0$ for all $w \in W$, so $\alpha(v) \in W^{\perp}$ as claimed.

Proof of Theorem ??

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n.

Proof of Theorem **??**

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.
Now suppose that $n > 1$.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra)

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.
Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$. Now put $V' = (\mathbb{C}v_1)^{\perp}$.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$. Now put $V' = (\mathbb{C}v_1)^\perp$. The lemma tells us that $\alpha(V') \leq V'$

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.
Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$. Now put $V' = (\mathbb{C}v_1)^\perp$. The lemma tells us that $\alpha(V') \leq V'$, so we can regard α as a self-adjoint operator on V'.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$. Now put $V' = (\mathbb{C}v_1)^\perp$. The lemma tells us that $\alpha(V') \leq V'$, so we can regard α as a self-adjoint operator on V'. Moreover, $\dim(V') = n - 1$, so our induction hypothesis applies.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$. Now put $V' = (\mathbb{C}v_1)^{\perp}$. The lemma tells us that $\alpha(V') \leq V'$, so we can regard α as a self-adjoint operator on V'. Moreover, $\dim(V') = n - 1$, so our induction hypothesis applies. This means that there is an orthonormal basis for V' (say $v_2, v_3, \ldots, v_n$) consisting of eigenvectors for α.

Proof of Theorem ??

Put $n = \dim(V)$; the proof is by induction on n. If $n = 1$ then we choose any unit vector $v_1 \in V$ and note that $\alpha(v_1) \in V = \mathbb{C}v_1$. This means that $\alpha(v_1) = \lambda_1 v_1$ for some $\lambda_1 \in \mathbb{C}$, so v_1 is an eigenvector, and this proves the theorem in the case $n = 1$.

Now suppose that $n > 1$. The characteristic polynomial of α is then a polynomial of degree n over $\mathbb{C}$, so it must have at least one root (by the fundamental theorem of algebra), say λ_1. We know that the roots of the characteristic polynomial are precisely the eigenvalues, so λ_1 is an eigenvalue, so we can find a nonzero vector $u_1 \in V$ with $\alpha(u_1) = \lambda_1 u_1$. We then put $v_1 = u_1/\|u_1\|$, so $\|v_1\| = 1$ and v_1 is still an eigenvector of eigenvalue λ_1, which implies that $\alpha(\mathbb{C}v_1) \leq \mathbb{C}v_1$. Now put $V' = (\mathbb{C}v_1)^\perp$. The lemma tells us that $\alpha(V') \leq V'$, so we can regard α as a self-adjoint operator on V'. Moreover, $\dim(V') = n - 1$, so our induction hypothesis applies. This means that there is an orthonormal basis for V' (say $v_2, v_3, \ldots, v_n$) consisting of eigenvectors for α. It follows that $v_1, v_2, \ldots, v_n$ is an orthonormal basis for V consisting of eigenvectors for α.

An example

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector.

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$.

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle$$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same.

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same. Now choose another unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same. Now choose another unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$.

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same. Now choose another unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. Then $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is an orthonormal basis for $\mathbb{R}^3$.

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi \colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same. Now choose another unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. Then $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is an orthonormal basis for $\mathbb{R}^3$. Moreover, we have $\pi(\mathbf{a}) = \mathbf{a}$ and $\pi(\mathbf{b}) = 0$ and $\pi(\mathbf{c}) = 0$

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same. Now choose another unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. Then $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is an orthonormal basis for $\mathbb{R}^3$. Moreover, we have $\pi(\mathbf{a}) = \mathbf{a}$ and $\pi(\mathbf{b}) = 0$ and $\pi(\mathbf{c}) = 0$, so $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are all eigenvectors for π

An example

Let $\mathbf{a} \in \mathbb{R}^3$ be a unit vector. Define $\pi\colon \mathbb{C}^3 \to \mathbb{C}^3$ by $\pi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}$. This is self-adjoint because

$$\langle \pi(\mathbf{v}), \mathbf{w} \rangle = \langle \langle \mathbf{v}, \mathbf{a} \rangle \mathbf{a}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{a} \rangle \langle \mathbf{a}, \mathbf{w} \rangle$$

whereas

$$\langle \mathbf{v}, \pi(\mathbf{w}) \rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{a} \rangle \mathbf{a} \rangle = \overline{\langle \mathbf{w}, \mathbf{a} \rangle} \langle \mathbf{v}, \mathbf{a} \rangle = \langle \mathbf{a}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{a} \rangle,$$

which is the same. Now choose another unit vector $\mathbf{b}$ orthogonal to $\mathbf{a}$ and put $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. Then $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is an orthonormal basis for $\mathbb{R}^3$. Moreover, we have $\pi(\mathbf{a}) = \mathbf{a}$ and $\pi(\mathbf{b}) = 0$ and $\pi(\mathbf{c}) = 0$, so $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are all eigenvectors for π (with eigenvalues 1, 0 and 0).

Another example

Another example

Let T_n be the set of trigonometric polynomials of degree at most n.

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$.

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$).

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle$$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\begin{aligned} &\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\ = &\int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt \end{aligned}$$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\, dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$
\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\, dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\, dt
\end{aligned}
$$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$
\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\
&= -\,i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt
\end{aligned}
$$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$
\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\
&= -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt = -i\,[f(2\pi)g(2\pi) - f(0)g(0)]
\end{aligned}
$$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\, dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\, dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\, dt \\
&= -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\, dt = -i\,[f(2\pi)g(2\pi) - f(0)g(0)] = 0
\end{aligned}$$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\
&= -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt = -i\left[f(2\pi)g(2\pi) - f(0)g(0)\right] = 0,
\end{aligned}$$

so δ is self-adjoint.

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\
&= -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt = -i\,[f(2\pi)g(2\pi) - f(0)g(0)] = 0,
\end{aligned}$$

so δ is self-adjoint. We have $e_k'(t) = ik\, e_k(t)$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\begin{aligned} &\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\ &= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\ &= -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt = -i\left[f(2\pi)g(2\pi) - f(0)g(0)\right] = 0, \end{aligned}$$

so δ is self-adjoint. We have $e_k'(t) = ik\, e_k(t)$, so $\delta(e_k) = -k\, e_k$

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$
\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
&= \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\
&= -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt = -i\,[f(2\pi)g(2\pi) - f(0)g(0)] = 0,
\end{aligned}
$$

so δ is self-adjoint. We have $e_k'(t) = ik\,e_k(t)$, so $\delta(e_k) = -k\,e_k$, so e_k is an eigenvector of eigenvalue $-k$.

Another example

Let T_n be the set of trigonometric polynomials of degree at most n. We use the usual inner product on T_n, given by $\langle f, g\rangle = \int_{t=0}^{2\pi} f(t)\overline{g(t)}\,dt$. Define $\delta\colon T_n \to T_n$ by $\delta(f) = if'$ (where $i = \sqrt{-1}$). We have

$$\begin{aligned}
&\langle f, \delta(g)\rangle - \langle \delta(f), g\rangle \\
={}& \int_0^{2\pi} f(t)\overline{ig'(t)} - if'(t)\overline{g(t)}\,dt = -i\int_0^{2\pi} f(t)\overline{g}'(t) + f'(t)\overline{g}(t)\,dt \\
={}& -i\int_0^{2\pi} \frac{d}{dt}(f(t)\overline{g}(t))\,dt = -i\,[f(2\pi)g(2\pi) - f(0)g(0)] = 0,
\end{aligned}$$

so δ is self-adjoint. We have $e_k'(t) = ik\,e_k(t)$, so $\delta(e_k) = -k\,e_k$, so e_k is an eigenvector of eigenvalue $-k$. These eigenvectors give us an orthonormal basis.

Another example

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$.

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle = \mathrm{trace}(\tau(X)\, Y^\dagger)$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle = \operatorname{trace}(\tau(X)\ Y^\dagger) = \operatorname{trace}(X^T \overline{Y}^T)$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T \overline{Y}^T) = \mathrm{trace}((\overline{Y} X)^T)$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T \overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

$$\langle X, \tau(Y) \rangle$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

$$\langle X, \tau(Y)\rangle = \mathrm{trace}(X\tau(Y)^\dagger)$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

$$\langle X, \tau(Y)\rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y})))$$

Another example

Define $\tau \colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T \overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

$$\langle X, \tau(Y) \rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y}))) = \mathrm{trace}(X\overline{Y})$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y \rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T \overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

$$\langle X, \tau(Y) \rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y}))) = \mathrm{trace}(X\overline{Y}) = \mathrm{trace}(\overline{Y}X).$$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$

$$\langle X, \tau(Y)\rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y}))) = \mathrm{trace}(X\overline{Y}) = \mathrm{trace}(\overline{Y}X).$$

This shows that τ is self-adjoint.

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$
$$\langle X, \tau(Y)\rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y}))) = \mathrm{trace}(X\overline{Y}) = \mathrm{trace}(\overline{Y}X).$$

This shows that τ is self-adjoint. The matrices

$$P_1 = \begin{bmatrix} 1&0&0\\0&0&0\\0&0&0 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 0&0&0\\0&1&0\\0&0&0 \end{bmatrix} \qquad P_3 = \begin{bmatrix} 0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$$

$$P_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\1&0&0\\0&0&0 \end{bmatrix} \qquad P_5 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\1&0&0 \end{bmatrix} \qquad P_6 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&1&0 \end{bmatrix}$$

$$P_7 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\-1&0&0\\0&0&0 \end{bmatrix} \qquad P_8 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\-1&0&0 \end{bmatrix} \qquad P_9 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&-1&0 \end{bmatrix}$$

give an orthonormal basis for $M_3\mathbb{R}$.

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \operatorname{trace}(\tau(X)\ Y^\dagger) = \operatorname{trace}(X^T\overline{Y}^T) = \operatorname{trace}((\overline{Y}X)^T) = \operatorname{trace}(\overline{Y}X)$$
$$\langle X, \tau(Y)\rangle = \operatorname{trace}(X\tau(Y)^\dagger) = \operatorname{trace}(X\tau(\tau(\overline{Y}))) = \operatorname{trace}(X\overline{Y}) = \operatorname{trace}(\overline{Y}X).$$

This shows that τ is self-adjoint. The matrices

$$P_1 = \begin{bmatrix} 1&0&0\\0&0&0\\0&0&0 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 0&0&0\\0&1&0\\0&0&0 \end{bmatrix} \qquad P_3 = \begin{bmatrix} 0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$$

$$P_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\1&0&0\\0&0&0 \end{bmatrix} \qquad P_5 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\1&0&0 \end{bmatrix} \qquad P_6 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&1&0 \end{bmatrix}$$

$$P_7 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\-1&0&0\\0&0&0 \end{bmatrix} \qquad P_8 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\-1&0&0 \end{bmatrix} \qquad P_9 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&-1&0 \end{bmatrix}$$

give an orthonormal basis for $M_3\mathbb{R}$. For $k \leq 6$ we have $\tau(P_k) = P_k$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \text{trace}(\tau(X)\ Y^\dagger) = \text{trace}(X^T\overline{Y}^T) = \text{trace}((\overline{Y}X)^T) = \text{trace}(\overline{Y}X)$$
$$\langle X, \tau(Y)\rangle = \text{trace}(X\tau(Y)^\dagger) = \text{trace}(X\tau(\tau(\overline{Y}))) = \text{trace}(X\overline{Y}) = \text{trace}(\overline{Y}X).$$

This shows that τ is self-adjoint. The matrices

$$P_1 = \begin{bmatrix} 1&0&0\\0&0&0\\0&0&0 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 0&0&0\\0&1&0\\0&0&0 \end{bmatrix} \qquad P_3 = \begin{bmatrix} 0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$$

$$P_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\1&0&0\\0&0&0 \end{bmatrix} \qquad P_5 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\1&0&0 \end{bmatrix} \qquad P_6 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&1&0 \end{bmatrix}$$

$$P_7 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\-1&0&0\\0&0&0 \end{bmatrix} \qquad P_8 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\-1&0&0 \end{bmatrix} \qquad P_9 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&-1&0 \end{bmatrix}$$

give an orthonormal basis for $M_3\mathbb{R}$. For $k \leq 6$ we have $\tau(P_k) = P_k$, so P_k is an eigenvector of τ with eigenvalue $+1$.

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$
$$\langle X, \tau(Y)\rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y}))) = \mathrm{trace}(X\overline{Y}) = \mathrm{trace}(\overline{Y}X).$$

This shows that τ is self-adjoint. The matrices

$$P_1 = \begin{bmatrix} 1&0&0\\0&0&0\\0&0&0 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 0&0&0\\0&1&0\\0&0&0 \end{bmatrix} \qquad P_3 = \begin{bmatrix} 0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$$

$$P_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\1&0&0\\0&0&0 \end{bmatrix} \qquad P_5 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\1&0&0 \end{bmatrix} \qquad P_6 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&1&0 \end{bmatrix}$$

$$P_7 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\-1&0&0\\0&0&0 \end{bmatrix} \qquad P_8 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\-1&0&0 \end{bmatrix} \qquad P_9 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&-1&0 \end{bmatrix}$$

give an orthonormal basis for $M_3\mathbb{R}$. For $k \leq 6$ we have $\tau(P_k) = P_k$, so P_k is an eigenvector of τ with eigenvalue $+1$. Similarly, for $k > 6$ we have $\tau(P_k) = -P_k$

Another example

Define $\tau\colon M_3\mathbb{C} \to M_3\mathbb{C}$ by $\tau(X) = X^T$. We find that

$$\langle \tau(X), Y\rangle = \mathrm{trace}(\tau(X)\ Y^\dagger) = \mathrm{trace}(X^T\overline{Y}^T) = \mathrm{trace}((\overline{Y}X)^T) = \mathrm{trace}(\overline{Y}X)$$
$$\langle X, \tau(Y)\rangle = \mathrm{trace}(X\tau(Y)^\dagger) = \mathrm{trace}(X\tau(\tau(\overline{Y}))) = \mathrm{trace}(X\overline{Y}) = \mathrm{trace}(\overline{Y}X).$$

This shows that τ is self-adjoint. The matrices

$$P_1 = \begin{bmatrix} 1&0&0\\0&0&0\\0&0&0 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 0&0&0\\0&1&0\\0&0&0 \end{bmatrix} \qquad P_3 = \begin{bmatrix} 0&0&0\\0&0&0\\0&0&1 \end{bmatrix}$$

$$P_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\1&0&0\\0&0&0 \end{bmatrix} \qquad P_5 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\1&0&0 \end{bmatrix} \qquad P_6 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&1&0 \end{bmatrix}$$

$$P_7 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&1&0\\-1&0&0\\0&0&0 \end{bmatrix} \qquad P_8 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&1\\0&0&0\\-1&0&0 \end{bmatrix} \qquad P_9 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0&0&0\\0&0&1\\0&-1&0 \end{bmatrix}$$

give an orthonormal basis for $M_3\mathbb{R}$. For $k \leq 6$ we have $\tau(P_k) = P_k$, so P_k is an eigenvector of τ with eigenvalue $+1$. Similarly, for $k > 6$ we have $\tau(P_k) = -P_k$, so P_k is an eigenvector of τ with eigenvalue -1.