3 Vector spaces and linear transformations

Chapter contents

3.1 Vectors in $\mathbf{R}^n$

Since Descartes, the space we live in is described mathematically as

$$\mathbf{R}^3 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \ : \ x_1, x_2, x_3 \in \mathbf{R} \right\}.$$

But we work with other spaces as well, for instance we graph real-valued functions of one real variable in

These have the common (and obvious) generalisation for $n \in \mathbf{N}$

We refer to $\mathbf{R}^n$ as *n-space* (or *n-dimensional space*), and to the elements of $\mathbf{R}^n$ as *vectors*.

Arithmetic operations on vectors

Scalar multiplication

$$\lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

(Two nonzero vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ are said to be *parallel* if $\mathbf{u} = \lambda \mathbf{v}$ for some scalar $\lambda \in \mathbf{R}$.)

Vector addition

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

Some properties of vector arithmetic

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$
- $\lambda(\mathbf{u} + \mathbf{v}) = \lambda \mathbf{u} + \lambda \mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ and all $\lambda \in \mathbf{R}$
- $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in \mathbf{R}^n$, etc.

3.2 Axiomatic definition of vector spaces

If you try to describe properties that are common to the spaces $\mathbf{R}^n$ that we have been working with, you may eventually end up with the following.

Fix a field $\mathbf{F}$. (For us, this means: either $\mathbf{R}$, $\mathbf{Q}$, $\mathbf{C}$, or $\mathbf{F}_2$.)

A *vector space V over $\mathbf{F}$* is a set with two operations

- addition $V \times V \longrightarrow V$, $(\mathbf{u}, \mathbf{v}) \longmapsto \mathbf{u} + \mathbf{v}$
- scalar multiplication $\mathbf{F} \times V \longrightarrow V$, $(\lambda, \mathbf{u}) \longmapsto \lambda\mathbf{u}$

satisfying the axioms

(a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in V$.

(b) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(c) There exists an element $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.

(d) For all $\mathbf{v} \in V$ there exists an element denoted $-\mathbf{v}$ with the property that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

(e) $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$ for all $\mathbf{v} \in V$ and all $\lambda \in \mathbf{F}$.

(f) $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$ for all $\mathbf{v} \in V$ and all $\lambda, \mu \in \mathbf{F}$.

(g) $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in V$ and all $\lambda \in \mathbf{F}$.

(h) $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.

Example 3.1 (Real n-dimensional space)**.** For any $n \in \mathbf{N}$, the space $\mathbf{R}^n$ is a vector space over $\mathbf{R}$.

Yes, even $\mathbf{R}^0 = \{\mathbf{0}\}$ (a single point).

Example 3.2 (Spaces of polynomials).

$$\mathcal{P}(\mathbf{R}) = \left\{a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0 : \; a_j \in \mathbf{R}, \; m \in \mathbf{N}\right\}.$$

$$\mathcal{P}_m(\mathbf{R}) = \left\{a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0 : \; a_j \in \mathbf{R}\right\}, \qquad m \in \mathbf{N} \text{ fixed.}$$

Example 3.3 (Space of functions)**.** Fix a set S and consider

$$\mathcal{F}(S, \mathbf{R}) = \{f : \; S \longrightarrow \mathbf{R}\}.$$

Example 3.4 (Space of $m \times n$ matrices)**.** Let $M_{m\times n}(\mathbf{R})$ denote the set of all $m \times n$ matrices with real entries.

All the examples we have seen have direct analogues over other fields $\mathbf{F}$ such as $\mathbf{Q}$, $\mathbf{C}$, or $\mathbf{F}_2$.

For instance, $\mathbf{C}^3$, $\mathcal{P}_5(\mathbf{F}_2)$, $M_{2\times 3}(\mathbf{Q})$.

3.3 Linear transformations

A guiding principle in mathematics is that, once you have a structure on a set, you should study functions that preserve this structure. We do this now for vector spaces, which are sets with the added structure of an addition and a scalar multiplication.

Let V and W be vector spaces over the same field of scalars $\mathbf{F}$. A function $T: V \longrightarrow W$ is called a *linear transformation* if it satisfies

(a) $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ for all $\mathbf{v}_1, \mathbf{v}_2 \in V$

(b) $T(\lambda \mathbf{v}) = \lambda T(\mathbf{v})$ for all $\mathbf{v} \in V$ and all $\lambda \in \mathbf{F}$.

Since linear transformations are functions, they can be injective, surjective, bijective, or none of the above.

A bijective linear transformation is also called an *isomorphism*; we then say that the spaces V and W are *isomorphic*.

Example 3.5. Fix m, n, and a matrix $A \in M_{m\times n}(\mathbf{F})$. Consider the function $T:\ \mathbf{F}^n \longrightarrow \mathbf{F}^m$ defined by

$$T(\mathbf{x}) = A\mathbf{x}.$$

Proposition 3.6. *Fix m, n, and a linear transformation $T : \mathbf{F}^n \longrightarrow \mathbf{F}^m$. Then there exists a unique matrix $A \in M_{m\times n}(\mathbf{F})$ such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbf{F}^n$.*

Corollary 3.7. *Fix m and n. There is a bijective function*

$$\{\textit{linear transformations } \mathbf{F}^n \longrightarrow \mathbf{F}^m\} \longrightarrow M_{m\times n}(\mathbf{F}).$$

Linear operations on functions

You have encountered linear transformations on spaces of functions, for instance

$$D_0 : \ \mathrm{Diff}(\mathbf{R}, \mathbf{R}) \longrightarrow \mathbf{R}$$

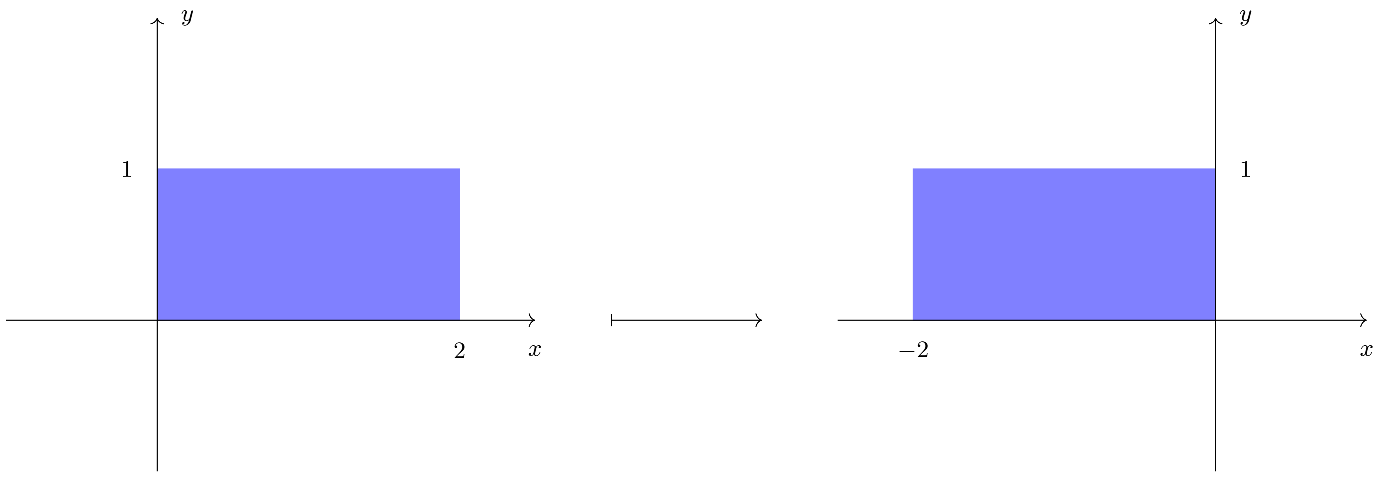
or

$$I_{[a,b]} : \ \mathrm{Cont}\big([a,b], \mathbf{R}\big) \longrightarrow \mathbf{R}$$

You also know many geometric transformations on $\mathbf{R}^2$ that are linear. In this simple setting, a linear transformation

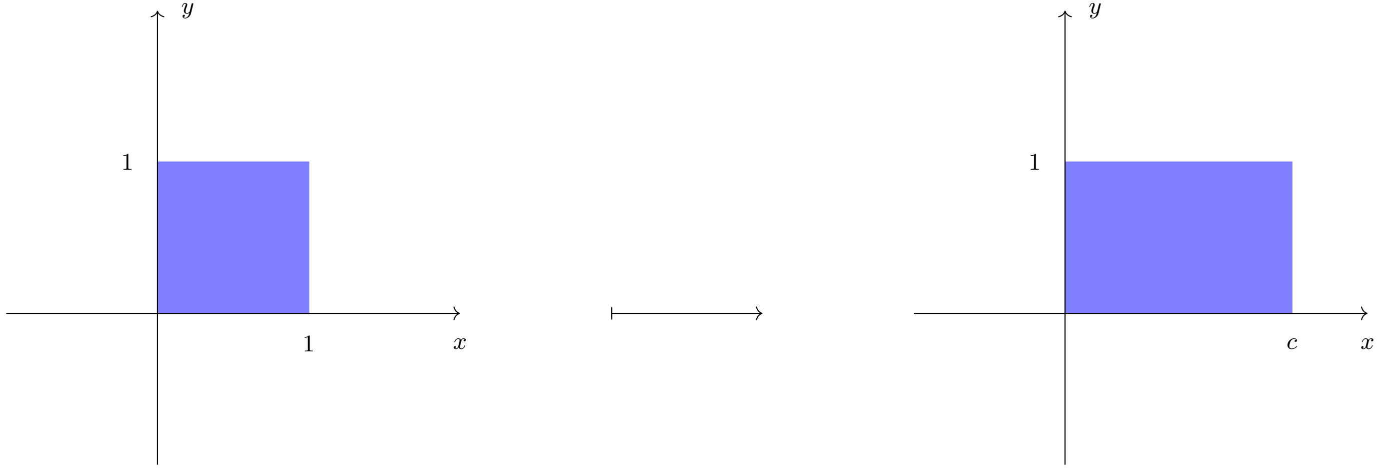
- sends the origin to the origin
- sends lines through the origin to lines through the origin;
- sends parallelograms with one vertex at the origin to parallellograms with one vertex at the origin (not necessarily congruent).

Reflections

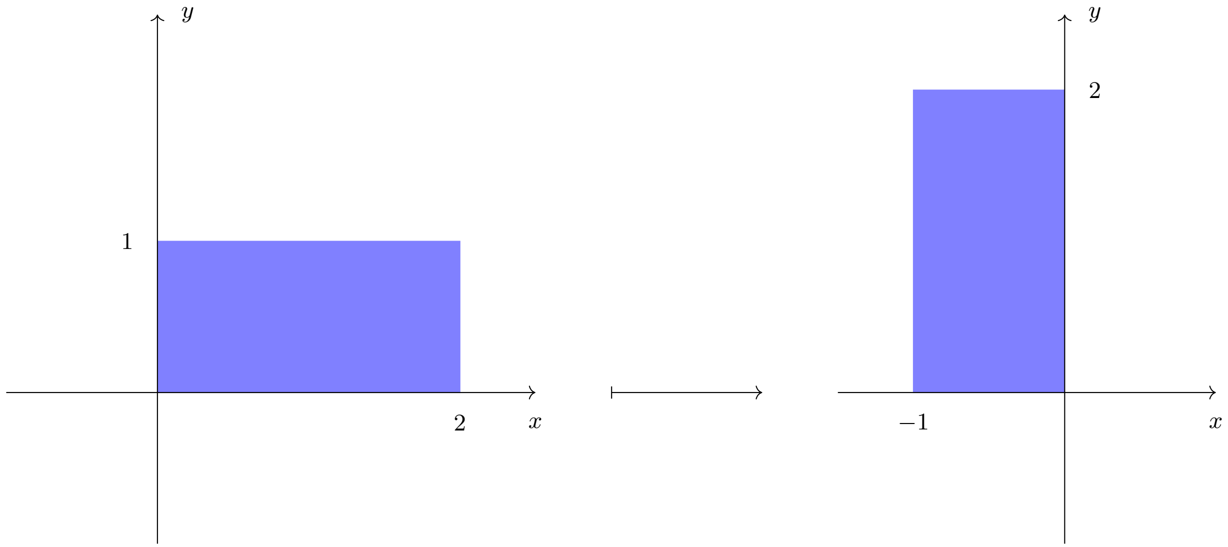


More generally, reflection in any line passing through the origin is a linear transformation.

Scaling: dilations/contractions

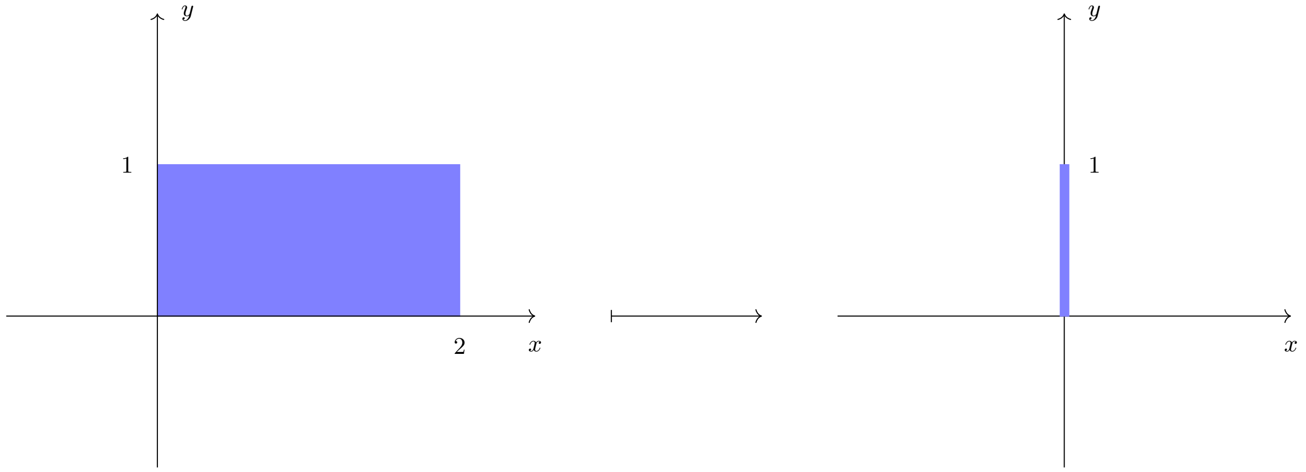


Rotations



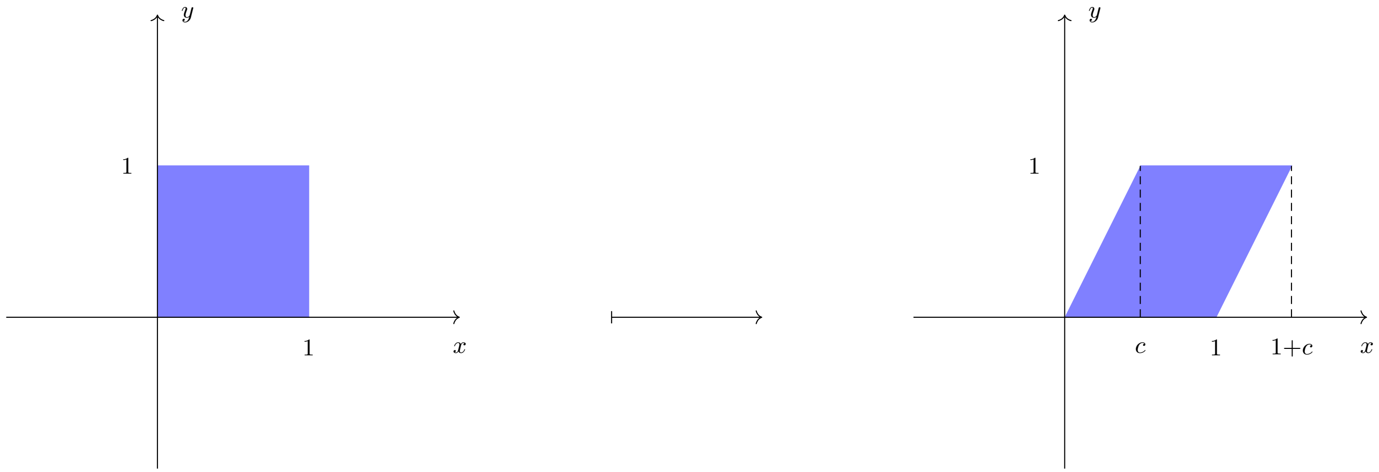
More generally, rotation around the origin by any angle θ is a linear transformation.

Orthogonal projections

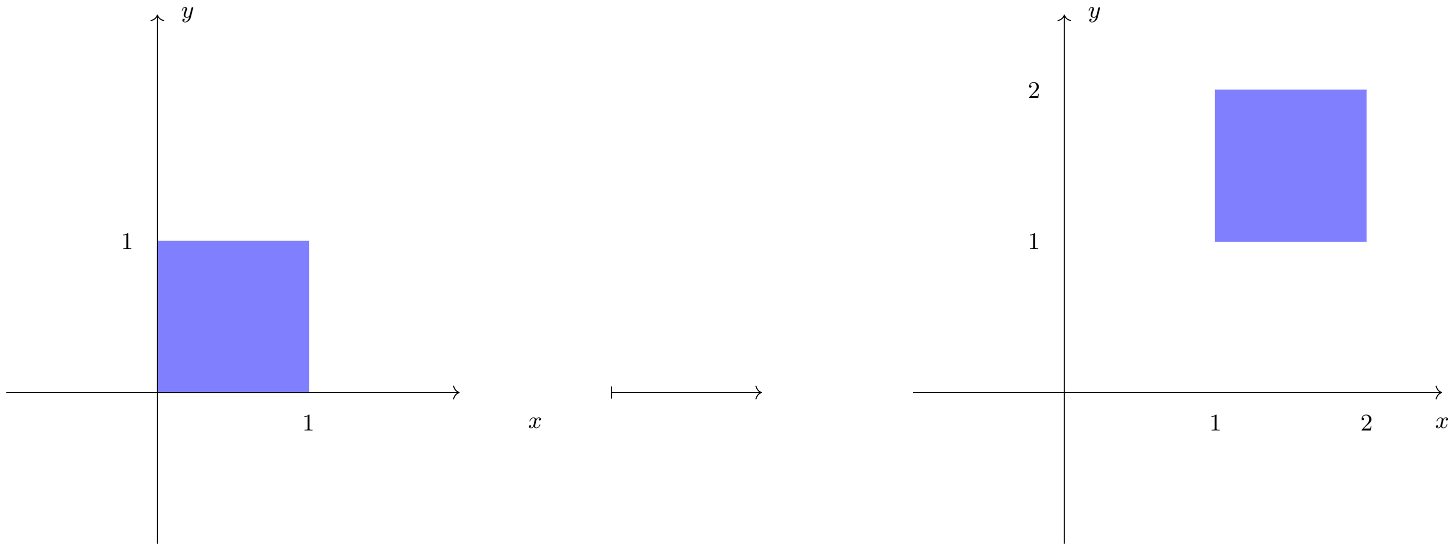


More generally, orthogonal projection onto any line passing through the origin is a linear transformation.

Shears



Translations



Translations are basically never linear transformations (the only one that is linear is translation by

More generally, any $S: \ V \longrightarrow W$ such that $S(\mathbf{0}_V) \neq \mathbf{0}_W$ is not linear.

3.4 Subspaces

Let V be a vector space and let U be a subset of V. We say that U is a *subspace* of V if U is itself a vector space under the operations inherited from V.

The Subspace Theorem. A subset U of a vector space V is a subspace if and only if it satisfies the three properties

(a) $\mathbf{0}_V \in U$;

(b) if $\mathbf{u}_1, \mathbf{u}_2 \in U$, then $\mathbf{u}_1 + \mathbf{u}_2 \in U$;

(c) if $\lambda \in \mathbf{F}$ and $\mathbf{u} \in U$, then $\lambda\mathbf{u} \in U$.

Example 3.8. Let $V = \mathbf{R}^2$ and

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 : \; y = 2x \right\}.$$

Example 3.9. Let $V = \mathbf{R}^2$ and

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 : \; y = x^2 \right\}.$$

Example 3.10. Let $V = \mathbf{R}^2$ and

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 : \; xy = 0 \right\}.$$

Example 3.11. Let $V = \mathcal{F}(\mathbf{R}, \mathbf{R})$ and

$$U = \{f \in V : \ f \text{ is continuous}\}.$$

Subsets that **are** subspaces tend to arise in one of a small number of natural constructions; this allows us to prove that they are subspaces without going through the Subspace Theorem.

The *kernel* (or *nullspace*) of a linear transformation $T: V \longrightarrow W$ is

$$\ker(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}_W\}.$$

Proposition 3.12. $\ker(T)$ *is a subspace of* V*.*

Example 3.13 (Example 3.8 revisited)**.** Let $V = \mathbf{R}^2$ and

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 : \; y = 2x \right\}.$$

Example 3.14. Let $V = \mathcal{F}(\mathbf{R}, \mathbf{R})$ and

$$U = \{f \in V : \ f(-3) = 0 \text{ and } f(1) = 0\}.$$

Example 3.15. Let $V = \mathbf{F}_2^4$ and

$$U = \left\{ \begin{bmatrix} \hat{0} \\ \hat{0} \\ \hat{0} \\ \hat{0} \end{bmatrix}, \begin{bmatrix} \hat{1} \\ \hat{1} \\ \hat{0} \\ \hat{0} \end{bmatrix}, \begin{bmatrix} \hat{1} \\ \hat{0} \\ \hat{1} \\ \hat{0} \end{bmatrix}, \begin{bmatrix} \hat{1} \\ \hat{0} \\ \hat{0} \\ \hat{1} \end{bmatrix}, \begin{bmatrix} \hat{0} \\ \hat{1} \\ \hat{1} \\ \hat{0} \end{bmatrix}, \begin{bmatrix} \hat{0} \\ \hat{1} \\ \hat{0} \\ \hat{1} \end{bmatrix}, \begin{bmatrix} \hat{0} \\ \hat{0} \\ \hat{1} \\ \hat{1} \end{bmatrix}, \begin{bmatrix} \hat{1} \\ \hat{1} \\ \hat{1} \\ \hat{1} \end{bmatrix} \right\}.$$

The *image* of a linear transformation $T: V \longrightarrow W$ is defined to be the range of T, that is

$$\operatorname{im}(T) = \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}.$$

Proposition 3.16. $\operatorname{im}(T)$ *is a subspace of* W*.*

3.5 Linear combinations, spanning sets, and linear independence

Let V be a vector space and $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$. A *linear combination* of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a vector of the form

$$\mathbf{w} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m$$

with $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$, the field of scalars of V.

The set of **all** linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is called the *span* of these m vectors:

$$\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = \{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m : \ \lambda_1, \ldots, \lambda_m \in \mathbf{F}\}.$$

By convention, if $m = 0$, we put $\operatorname{Span}(\varnothing) = \{\mathbf{0}_V\}$.

Example 3.17. Is the vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in the span of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$?

More generally, if $V = \mathbf{F}^n$ and we want to determine whether $\mathbf{w}$ is in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_m$, we form the augmented matrix

$$\left[\begin{array}{ccc|c} | & & | & | \\ \mathbf{v}_1 & \ldots & \mathbf{v}_m & \mathbf{w} \\ | & & | & | \end{array}\right]$$

and see whether the system has at least one solution (then $\mathbf{w}$ is in the span) or no solutions (then $\mathbf{w}$ is not in the span).

Proposition 3.18. *The span of a set of vectors in a vector space V is a subspace of V.*

Example 3.19. Let $V = \mathbf{F}_2^4$ and

$$W = \left\{ \begin{bmatrix} \hat{0} \\ \hat{0} \\ \hat{0} \\ \hat{0} \end{bmatrix}, \begin{bmatrix} \hat{1} \\ \hat{0} \\ \hat{0} \\ \hat{1} \end{bmatrix}, \begin{bmatrix} \hat{0} \\ \hat{1} \\ \hat{1} \\ \hat{0} \end{bmatrix}, \begin{bmatrix} \hat{1} \\ \hat{1} \\ \hat{1} \\ \hat{1} \end{bmatrix} \right\}.$$

Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$. We say that these vectors *span V* if

$$\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = V,$$

in other words every vector in V is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$. This set of vectors is then called a *spanning set of V*.

Example 3.20. Do the following vectors span $\mathbf{R}^3$?

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Linear (in)dependence

A vector space has many spanning sets. For instance $\mathbf{R}^3$ can be spanned any of the following:

- $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$
- $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} \right\}$
- $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1, -\mathbf{e}_2, \mathbf{e}_1 + 3\mathbf{e}_3\}$
- $\mathbf{R}^3$
- and many others.

Clearly some are more economical than others. To make this precise, we want to define the notion of redundancy in a set of vectors. This is called *linear dependence*.

A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ is *linearly dependent* if there exist scalars $a_1, \ldots, a_m \in \mathbf{F}$, not all zero, such that

$$a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}_V.$$

(We say that there is a *nontrivial linear relation* between the vectors.)

A consequence of linear dependence is that we can express some vectors as linear combinations of the others; let i be such that $a_i \neq 0$, then

$$\mathbf{v}_i = -\frac{a_1}{a_i}\mathbf{v}_1 - \cdots - \frac{a_{i-1}}{a_i}\mathbf{v}_{i-1} - \frac{a_{i+1}}{a_i}\mathbf{v}_{i+1} - \cdots - \frac{a_m}{a_i}\mathbf{v}_m.$$

We say that the vector $\mathbf{v}_i$ is *redundant*.

A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ is *linearly independent* if it is not linearly dependent; more precisely, if the only scalars $a_1, \ldots, a_m \in \mathbf{F}$ such that

$$a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}_V$$

are $a_1 = \cdots = a_m = 0$.

Example 3.21.

- $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$

- $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 6 \end{bmatrix}$

- $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

More generally, if $V = \mathbf{F}^n$ and we want to determine whether the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly independent, we form the augmented matrix

$$\left[\begin{array}{ccc|c} | & & | & | \\ \mathbf{v}_1 & \ldots & \mathbf{v}_m & \mathbf{0}_V \\ | & & | & | \end{array}\right]$$

and see whether the system has a unique solution (then the vectors are linearly independent) or infinitely many solutions (then the vectors are linearly dependent).

Example 3.22. Are the vectors $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ linearly independent?

Example 3.23. Are the polynomials

$$x(x-1), \qquad (x-1)(x-2), \qquad x^2-1$$

linearly independent in $\mathcal{P}_2(\mathbf{R})$?

Does the constant polynomial 1 lie in their span?

3.6 Bases and dimension

Let V be a vector space. A subset S of V is a *basis of V* if it spans V and is linearly independent.

Convention: $\{\mathbf{0}\}$ has the unique basis $\varnothing$.

Example 3.24 (Standard basis of $\mathbf{F}^n$)**.**

Example 3.25. Is $\left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right\}$ a basis of $\mathbf{R}^2$?

Executive summary: Let V be a vector space.

- V has a basis.
- Any two bases of V have the same cardinality. (In the sense we discussed already, i.e. there is a bijection from one basis to the other.) If this cardinality is finite, it is called the *dimension of V*, and V is said to be *finite-dimensional*.
  We will focus almost exclusively on this situation. MAST30026 **Metric and Hilbert spaces** treats infinite-dimensional vector spaces.
- Any spanning set of V contains a basis of V.
- Any linearly independent subset of V can be extended to a basis of V.
- If V has dimension n, then
  - Any subset S with $\#S < n$ is **not** spanning V.
  - Any subset S with $\#S > n$ is **not** linearly independent.

In order to avoid circularity in the logic, we give a slightly different initial definition of finite-dimensionality:
A vector space V is *finite-dimensional* if there exists a finite spanning subset S.

Proposition 3.26. *Given any linearly independent subset* $\mathbf{v}_1, \ldots, \mathbf{v}_m$ *of* V *and any finite spanning set* $\mathbf{w}_1, \ldots, \mathbf{w}_n$ *of* V*, we have* $m \leqslant n$*.*

Proposition 3.27. *Let V be a finite-dimensional vector space. Every subspace of V is finite-dimensional.*

Proposition 3.28. *Any finite-dimensional vector space V has a finite basis.*

Note that the proof showed something of independent interest, which we record here:

Proposition 3.29. *Any finite spanning set S of V contains a basis $\mathcal{B}$ of V.*

Proposition 3.30. *Let V be a finite-dimensional vector space. Any two bases of V have the same size.*

The size of any basis of V is called the *dimension* of V.

The zero vector space has dimension $\#\varnothing = 0$.

Example 3.31. The standard basis of the space $M_{m\times n}(\mathbf{F})$ is

Example 3.32. The standard basis of the space $\mathcal{P}_n(\mathbf{F})$ is

Proposition 3.33. *Let V be a vector space and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$. Then S is ...*

(a) ... a spanning set of V if and only if every $\mathbf{v} \in V$ has at least one expression of the form (*),

(b) ... a linearly independent subset of V if and only if every $\mathbf{v} \in V$ has at most expression of the form (*),

(c) ... a basis of V if and only if every $\mathbf{v} \in V$ has a unique expression of the form (*),

where

$$(*) \qquad \mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n, \qquad a_i \in \mathbf{F}.$$

A linear transformation can be defined (and is completely determined) by its values on a basis of the domain:

Proposition 3.34. *Let* V, W *be vector spaces,* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *a basis of* V*, and* $\mathbf{w}_1, \ldots, \mathbf{w}_n$ *a collection of* n *(not necessarily distinct) vectors in* W*. Then there exists a unique linear transformation* $T: V \longrightarrow W$ *with the property that*

$$T(\mathbf{v}_i) = \mathbf{w}_i \qquad \textit{for } i = 1, 2, \ldots, n.$$

We use linear transformations to clarify the effect of choosing a basis on a vector space.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis of a vector space V. Choose an ordering on these vectors, say $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$. We will refer to $\mathcal{B}$ as an *ordered basis* of V.

Example 3.35. We know that $\{1, x, x^2\}$ is a linearly independent and spanning set for the vector space $\mathcal{P}_2$. So $\mathcal{B} = (1, x, x^2)$ is an ordered basis.

We will use this to construct a linear transformation $T: \mathcal{P}_2 \longrightarrow \mathbf{R}^3$ as follows: by Proposition 3.34 we can require

This is a general phenomenon: Let $\mathcal{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an ordered basis of a vector space V.

Let $\varphi_{\mathcal{B}} : V \longrightarrow \mathbf{F}^n$ denote the linear transformation obtained by requiring

$$\varphi_{\mathcal{B}}(\mathbf{v}_i) = \mathbf{e}_i \qquad \text{for } i = 1, \ldots, n.$$

To see what this does to an arbitrary $\mathbf{v} \in V$, write $\mathbf{v}$ uniquely as a linear combination

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n,$$

then

$$\varphi_{\mathcal{B}}(\mathbf{v}) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = [\mathbf{v}]_{\mathcal{B}},$$

the *coordinate vector of* $\mathbf{v}$ *with respect to* $\mathcal{B}$.

Proposition 3.36. *The linear transformation* $\varphi_{\mathcal{B}} : V \longrightarrow \mathbf{F}^n$ *is invertible.*

So a choice of ordered basis on V defines an *isomorphism* between V and $\mathbf{F}^n$.

Example 3.37 ($\mathcal{P}_n$)**.**

Example 3.38. Consider the ordered basis $\mathcal{B} = (x^2 + x + 1, x + 1, 1)$ of $\mathcal{P}_2$.

3.7 Finding bases of subspaces

The approach depends on the manner in which the subspace is given to us.

Solution space of a homogeneous system

A homogeneous system with m equations and n variables takes the form

$$A\mathbf{x} = \mathbf{0}, \tag{1}$$

where A is an $m \times n$ matrix.

The matrix A defines a linear transformation $T : \mathbf{F}^n \longrightarrow \mathbf{F}^m$ by $T(\mathbf{x}) = A\mathbf{x}$. The solutions of the system (1) are precisely the elements of $\ker(T)$.

Example 3.39. Find a basis for the solution space of the system

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 = 0 \\ 3x_1 + 6x_2 + 4x_3 + x_4 = 0 \end{cases}$$

Note that the dimension of the solution space is equal to the number of free parameters of the system. Things work the same over other fields of scalars than **R**.

Example 3.40. Find a basis for the solution space of the system over $\mathbf{F}_2$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = \hat{0} \\ x_1 \qquad + x_3 \qquad = \hat{0} \end{cases}$$

Span of a set of vectors

Suppose $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of vectors in $\mathbf{F}^m$ and we are interested in the subspace

$$W = \mathrm{Span}(S) = \{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n : \ \lambda_i \in \mathbf{F}\}.$$

Recall (from our discussion of matrix multiplication a long time ago) that

$$A \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n,$$

where A is the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Therefore we can identify $\mathrm{Span}(S)$ with the image of the linear transformation $T : \ \mathbf{F}^n \longrightarrow \mathbf{F}^m$ given by $T(\mathbf{v}) = A\mathbf{v}$:

$$\mathrm{im}(T) = \{\mathbf{w} \in \mathbf{F}^m : \ \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in \mathbf{F}^n\} = \mathrm{Span}(S).$$

Example 3.41. Find a basis for Span(S), where $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} \right\} \subseteq \mathbf{R}^3$.

I claim that the vectors in S corresponding to the columns with leading ones in the RREF form a basis for $\mathrm{Span}(S)$:

This observation holds in general, and it stems from:

Proposition 3.42. *Let A be a matrix and suppose B can be obtained from A by applying one elementary row operation. Then the set of linear relations between the columns of A is equal to the set of linear relations between the columns of B; in other words, for any $c_1, \ldots, c_n \in \mathbf{F}$:*

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

if and only if

$$c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n = \mathbf{0},$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_n$ denote the columns of A and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ denote the columns of B.

This leads to *the column method* for finding a basis of $\mathrm{Span}(S)$, for finite $S \subseteq \mathbf{F}^n$:

(a) Construct a matrix A whose columns are the elements of S.

(b) Compute a row echelon form of A.

(c) The vectors of S corresponding to the columns with leading entries form a basis.

Note that the basis we obtain is a subset of the S we started with.

Span of a set of vectors, take two

Instead of building a matrix A whose **columns** are the vectors in S, we build a matrix B whose **rows** are the vectors in S.

Example 3.43. Find a basis for $\mathrm{Span}(S)$, where $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} \right\} \subseteq \mathbf{R}^3$.

I claim that the nonzero rows in the RREF of B form a basis for $\text{Span}(S)$:

This leads to *the row method* for finding a basis of $\mathrm{Span}(S)$, for finite $S \subseteq \mathbf{R}^n$:

(a) Construct a matrix B whose rows are the elements of S.

(b) Compute a row echelon form C of B.

(c) The nonzero rows of C form a basis of $\mathrm{Span}(S)$.

Note that the basis we obtain is in most cases **not a subset** of the S we started with.

Let C be an $m \times n$ matrix.

The subspace of $\mathbf{F}^m$ spanned by the columns of C is called *the column space of* C.

The subspace of $\mathbf{F}^n$ spanned by the rows of C is called *the row space of* C.

Proposition 3.44. $\dim \textit{row space}(C) = \dim \textit{column space}(C) = \operatorname{rank}(C)$.

Rank-nullity Theorem, Version I. For any matrix C,

$$\dim \text{column space}(C) + \dim \ker(C) = \text{number of columns of } C.$$

Corollary 3.45. *If A and B are $n \times n$ matrices such that $AB = I_n$, then $BA = I_n$.*

Example 3.46. Given the matrix

$$C = \begin{bmatrix} 1 & -1 & 2 & -2 \\ 2 & 0 & 1 & 0 \\ 5 & -3 & 7 & -6 \\ 1 & 1 & -1 & 3 \end{bmatrix},$$

find a basis and the dimension of the

(a) column space of C

(b) row space of C

(c) solution space of C.

3.8 Matrix representations and change of basis

Recall that an ordered basis $\mathcal{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ of a vector space V produces an isomorphism

$$\varphi_{\mathcal{B}} : V \longrightarrow \mathbf{F}^n$$

defined by taking coordinates with respect to $\mathcal{B}$: $\varphi_{\mathcal{B}}(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$.

Suppose we have a linear transformation $T : V \longrightarrow W$ and an ordered basis $\mathcal{C}$ of W, giving rise to a diagram:

The matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ is called the *matrix representation of the linear transformation T* with respect to the basis $\mathcal{B}$ of V and the basis $\mathcal{C}$ of W.

It is uniquely determined by the relation:

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} \qquad \text{for all } \mathbf{v} \in V,$$

and its entries can be described explicitly as:

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} | & & | \\ [T(\mathbf{v}_1)]_{\mathcal{C}} & \dots & [T(\mathbf{v}_n)]_{\mathcal{C}} \\ | & & | \end{bmatrix},$$

where $\mathcal{B} = \big(\mathbf{v}_1, \dots, \mathbf{v}_n\big)$.

Example 3.47. Let $V = \mathcal{P}_2(\mathbf{R})$ with ordered basis $\mathcal{B} = (1, x, x^2)$ and $W = \mathcal{P}_1(\mathbf{R})$ with ordered basis $\mathcal{C} = (1, x)$. Let $D : V \longrightarrow W$ be the derivative linear operator: $D(f) = \frac{df}{dx}$.

The matrix of D with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ is

For instance, the derivative of $2 - 5x + x^2$ is $-5 + 2x$, and indeed:

Rank-nullity Theorem, Version II. For any linear transformation $T: V \longrightarrow W$, we have

$$\operatorname{rank}(T) + \dim\ker(T) = \dim(V).$$

Corollary 3.48. *A linear transformation $f: V \longrightarrow V$ is injective if and only if it is surjective.*

For a linear transformation $T: V \longrightarrow V$, we often choose the same ordered basis $\mathcal{B}$ for both the domain V and the codomain V, and get a matrix representation that is simply denoted $[T]_{\mathcal{B}}$.

Example 3.49. The rotation about the origin through an angle of θ in $\mathbf{R}^2$ is a linear transformation.

Composition of linear transformations and matrix multiplication

If $T_1 : U \longrightarrow V$ and $T_2 : V \longrightarrow W$ are linear transformations, with $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ bases for U, V, W, then

$$[T_2 \circ T_1]_{\mathcal{D}\leftarrow\mathcal{B}} = [T_2]_{\mathcal{D}\leftarrow\mathcal{C}}[T_1]_{\mathcal{C}\leftarrow\mathcal{B}}.$$

In other words, the matrix representation of the composition $T_2 \circ T_1$ is the product of the matrix representation of T_2 and the matrix representation of T_1.

Important special case: Effect of change of basis on coordinates

The construction of the matrix representation applies to any linear transformation between finite-dimensional vector spaces.

In particular, we can consider the identity linear transformation $\mathrm{id}_V : V \longrightarrow V$.

After choosing two ordered bases $\mathcal{B}$ and $\mathcal{C}$ for V, we get a matrix representation

This is called the *change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$* and is uniquely determined by the relation

$$[\mathbf{v}]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} \qquad \text{for all } \mathbf{v} \in V.$$

Example 3.50. In $V = \mathbf{R}^2$, write down the change of basis matrix from $\mathcal{B}$ to $\mathcal{S}$, where

$$\mathcal{B} = \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \qquad \text{and} \qquad \mathcal{S} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right),$$

and use it to compute $[\mathbf{v}]_{\mathcal{S}}$, given that $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Example 3.51. Let $V = \mathcal{P}_2(\mathbf{R})$, $\mathcal{B} = (1, x, x^2)$ and $\mathcal{C} = (1, 1 + x, 1 + x + x^2)$.

Suppose that we have three ordered bases $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ of V.

What do you think the matrix product $P_{\mathcal{D}\leftarrow\mathcal{C}}P_{\mathcal{C}\leftarrow\mathcal{B}}$ is?

As a special case, we have

$$P_{\mathcal{B}\leftarrow\mathcal{C}}P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{B}\leftarrow\mathcal{B}} = I.$$

Effect of change of basis on matrix representations

Suppose V is a vector space and $T : V \longrightarrow V$ is a linear transformation. Given an ordered basis $\mathcal{B}$, we get a matrix representation $[T]_{\mathcal{B}}$. Given another ordered basis $\mathcal{C}$, we get another matrix representation $[T]_{\mathcal{C}}$.

To relate these two matrices, consider the following diagram:

Example 3.52. Consider the linear transformation $T:\ \mathbf{R}^3 \longrightarrow \mathbf{R}^3$ given by the matrix

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 3 \\ 2 & 0 & 2 \end{bmatrix}$$

and the ordered bases $\mathcal{S}$ (the standard basis) and $\mathcal{C} = \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right)$.

Matrix representations for geometric transformations of $\mathbb{R}^2$

Example 3.53 (Reflections)**.** Consider $T : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ given by reflection in the x-axis.

What if we change T to be reflection in the line $y = mx$?

Example 3.54 (Orthogonal projections)**.** Consider $T: \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ given by orthogonal projection onto the x-axis.

If we change T to be orthogonal projection onto the line $y = mx$, then

3.9 Eigenvectors and eigenvalues

We focus our attention on the study of linear transformations $T: V \longrightarrow V$.

As we have glimpsed in Examples 3.53 and 3.54, it is useful to identify vectors $\mathbf{v} \in V$ for which the image $T(\mathbf{v})$ is particularly easy to describe.

Example 3.55. Consider

$$A = \begin{bmatrix} -2 & 5 & 0 & 0 \\ 0 & -3 & -1 & 0 \\ 0 & 1 & 8 & 0 \\ 0 & 9 & -4 & 0 \end{bmatrix}.$$

A **nonzero** vector $\mathbf{v} \in V$ with the property that $T(\mathbf{v}) = \lambda\mathbf{v}$ for some $\lambda \in \mathbf{F}$ is called an *eigenvector* of T. The corresponding scalar λ is called an *eigenvalue* of T.

Example 3.56. Consider the linear transformation $T: \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ given by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

What is the effect on the vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$?

Finding eigenvalues

Proposition 3.57. *A scalar* λ *is an eigenvalue of the matrix* A *if and only if*

$$\det(A - \lambda I) = 0.$$

Example 3.58. Find the eigenvalues of $\begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix}$.

If A is an $n \times n$ matrix, then the expression $\det(A - xI)$ is a polynomial of degree n in x, called the *characteristic polynomial of A*.

The equation $\det(A - xI) = 0$ is called the *characteristic equation of A*.

When A has real entries, it is possible that some of its eigenvalues are complex (and not real).

When A has entries in $\mathbf{F}_2$, it is possible that some of its eigenvalues are not in $\mathbf{F}_2$ (exactly where they are is a matter for MAST30005 Algebra).

Example 3.59. Find the eigenvalues of $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

Finding eigenvectors

If λ is an eigenvalue of A, then the corresponding eigenvectors are the solutions of the vector equation

$$(A - \lambda I)\mathbf{v} = \mathbf{0}.$$

In other words, these are the nonzero elements in the kernel of $A - \lambda I$. In this setting, the kernel of $A - \lambda I$ is called the *λ-eigenspace of A*.

We solve this equation as usual, via row reduction.

Example 3.60. Find eigenvectors for the eigenvalues obtained in Example 3.58.

Example 3.61. Find the eigenspaces of the matrix from Example 3.59.

Proposition 3.62. *Let A be an $n \times n$ matrix and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be the result of finding a basis for each eigenspace of A and putting these bases together. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent.*

Corollary 3.63. *If A is an $n \times n$ matrix and the dimensions of the eigenspaces of A add up to n, then A has n linearly independent eigenvectors.*

Example 3.64. In Example 3.61,

Example 3.65. How many linearly independent eigenvectors does $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ have?

The Cayley–Hamilton Theorem

Theorem 3.66. *Let A be an $n \times n$ matrix and let*

$$\det(A - xI) = a_0 + a_1 x + \cdots + a_n x^n$$

be the characteristic polynomial of A. Then

$$a_0 I + a_1 A + \cdots + a_n A^n = \mathbf{0}.$$

Example 3.67. Consider $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$.

We can make use of the characteristic equation to compute powers of A:

Diagonalisation

Suppose you are given a linear transformation $T:\ V \longrightarrow V$. We have already seen that some choices of ordered basis of V yield simpler matrices than other choices.

Two natural questions present themselves:

(a) How simple can we make the matrix representation?

(b) How can we compute an ordered basis that achieves such a simple matrix representation?

The ultimate answer to these questions depends on the field of scalars of V, and takes us to MAST20022 **Group Theory and Linear Algebra**, MAST30005 **Algebra**, and beyond.

We will be focussing on a special case that is very useful in practice, that where the matrix representation is diagonal.

We say that a linear transformation $T: V \longrightarrow V$ is *diagonalisable* if there exists an ordered basis $\mathcal{B}$ such that the matrix $[T]_{\mathcal{B}}$ is diagonal.

We say that an $n \times n$ matrix A is *diagonalisable* if the associated linear transformation $\mathbf{R}^n \longrightarrow \mathbf{R}^n$ given by $\mathbf{v} \longmapsto A\mathbf{v}$ is diagonalisable.

Theorem 3.68. *A linear transformation $T: V \longrightarrow V$ is diagonalisable if and only if V has a basis $\mathcal{B}$ consisting of eigenvectors for T.*

Example 3.69. Consider $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.

Example 3.70. Consider $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

Corollary 3.71. *An $n \times n$ matrix A is diagonalisable if and only if there exist a diagonal $n \times n$ matrix D and an invertible $n \times n$ matrix P such that*

$$P^{-1}AP = D.$$

Moreover, the columns of P are linearly independent eigenvectors of A and the diagonal entries of D are the corresponding eigenvalues.

Example 3.72. Consider $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

Proposition 3.73. *If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalisable.*

Example 3.74. Consider $\begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix}$.

Matrix powers

Suppose we have to compute a high power of a matrix A. There is one situation when this is very easy, namely for a diagonal matrix

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \qquad \Rightarrow \qquad D^k = \begin{bmatrix} d_1^k & & & \\ & d_2^k & & \\ & & \ddots & \\ & & & d_n^k \end{bmatrix}.$$

The next best thing is when A is diagonalisable: