Section 1

Vector spaces (Revision)

Definition and examples of vector spaces

Being "linear" will boil down to "preserving addition" and "preserving scalar multiplication." Our first step is to specify the scalars with which we intend to work.

Definition 1.1 A *field* is a set F together with two binary operations

$$\begin{aligned} F \times F &\to F \\ (\alpha, \beta) &\mapsto \alpha + \beta \end{aligned} \qquad\qquad \begin{aligned} F \times F &\to F \\ (\alpha, \beta) &\mapsto \alpha\beta \end{aligned}$$

called *addition* and *multiplication*, respectively, such that

(i) $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in F$;

(ii) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for all $\alpha, \beta, \gamma \in F$;

(iii) there exists an element 0 in F such that $\alpha + 0 = \alpha$ for all $\alpha \in F$;

(iv) for each $\alpha \in F$, there exists an element $-\alpha$ in F such that $\alpha + (-\alpha) = 0$;

(v) $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in F$;

(vi) $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ for all $\alpha, \beta, \gamma \in F$;

(vii) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ for all $\alpha, \beta, \gamma \in F$;

(viii) there exists an element 1 in F such that $1 \neq 0$ and $1\alpha = \alpha$ for all $\alpha \in F$;

(ix) for each $\alpha \in F$ with $\alpha \neq 0$, there exists an element α^{-1} (or $1/\alpha$) in F such that $\alpha\alpha^{-1} = 1$.

Although a full set of axioms have been provided, we are not going to examine them in detail nor spend time developing the theory of fields. Instead, one should simply note that in a field one may add, subtract, multiply and divide (by *non-zero* scalars) and that all normal rules of arithmetic hold. This is illustrated by our examples.

Example 1.2 The following are examples of fields:

(i) $\mathbb{Q} = \{\, m/n \mid m, n \in \mathbb{Z},\ n \neq 0 \,\}$

(ii) $\mathbb{R}$

(iii) $\mathbb{C} = \{\, x + iy \mid x, y \in \mathbb{R} \,\}$

with all three possessing the usual addition and multiplication.

(iv) $\mathbb{Z}/p\mathbb{Z} = \{0, 1, \dots, p-1\}$, where p is a prime number, with addition and multiplication being performed modulo p.

The latter example is important in the context of pure mathematics and it is for this reason that I mention it. For the purposes of this module and for many applications of linear algebra in applied mathematics and the physical sciences, the examples $\mathbb{R}$ and $\mathbb{C}$ are the most important and it is safe to think of them as the typical examples of a field throughout. For those of a pure mathematical bent, however, it is worth noting that much of what is done in this course will work over an arbitrary field.

Definition 1.3 Let F be a field. A *vector space* over F is a set V together with the following operations

$$\begin{aligned} V \times V &\to V \\ (u, v) &\mapsto u + v \end{aligned} \qquad\qquad \begin{aligned} F \times V &\to V \\ (\alpha, v) &\mapsto \alpha v, \end{aligned}$$

called *addition* and *scalar multiplication*, respectively, such that

(i) $u + v = v + u$ for all $u, v \in V$;

(ii) $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$;

(iii) there exists a vector $\mathbf{0}$ in V such that $v + \mathbf{0} = v$ for all $v \in V$;

(iv) for each $v \in V$, there exists a vector $-v$ in V such that $v + (-v) = \mathbf{0}$;

(v) $\alpha(u + v) = \alpha u + \alpha v$ for all $u, v \in V$ and $\alpha \in F$;

(vi) $(\alpha + \beta)v = \alpha v + \beta v$ for all $v \in V$ and $\alpha, \beta \in F$;

(vii) $(\alpha\beta)v = \alpha(\beta v)$ for all $v \in V$ and $\alpha, \beta \in F$;

(viii) $1v = v$ for all $v \in V$.

Comments:

(i) A vector space then consists of a collection of *vectors* which we are permitted to add and which we may multiply by *scalars* from our base field. These operations behave in a natural way.

(ii) One aspect which requires some care is that the field contains the number 0 while the vector space contains the zero vector $\mathbf{0}$. The latter will be denoted by boldface in notes and on the slides. On the board boldface is unavailable, so although the difference is usually clear from the context (we can multiply vectors by scalars, but cannot multiply vectors, and we can add two vectors but cannot add a vector to a scalar) we shall use $\underset{\sim}{0}$ to denote the zero vector on that medium.

(iii) We shall use the term *real vector space* to refer to a vector space over the field $\mathbb{R}$ and *complex vector space* to refer to one over the field $\mathbb{C}$. Almost all examples in this module will be either real or complex vector spaces.

(iv) We shall sometimes refer simply to a vector space V without specifying the base field F. Nevertheless, there is always such a field F in the background and we will then use the term *scalar* to refer to the elements of this field when we fail to actually name it.

To illustrate why these are important objects, we shall give a number of examples, many of which should be familiar (not least from MT2501).

Example 1.4 (i) Let n be a positive integer and let F^n denote the set of column vectors of length n with entries from the field F:

$$F^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; x_1, x_2, \dots, x_n \in F \right\}.$$

This is an example of a vector space over F. Addition in F^n is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

while scalar multiplication is similarly given by

$$\alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

The zero vector is

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and

$$-\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$$

specifies the negative of a vector.

We could also consider the set of row vectors of length n as a vector space over F. This is sometimes also denoted F^n and has the advantage of being more easily written on the page! However, column vectors turn out to be slightly more natural than row vectors when we consider the matrix of a linear transformation later in the module.

As a further comment about distinguishing between scalars and vectors, we shall follow the usual convention of using boldface (or writing something such as $\underset{\sim}{v}$ on the board) to denote column vectors in the vector space F^n. We shall, however, usually not use boldface letters when referring to vectors in an abstract vector space (since in an actual example, they could become genuine column vectors, but also possibly matrices, polynomials, functions, etc.).

(ii) The complex numbers $\mathbb{C}$ can be viewed as a vector space over $\mathbb{R}$. Addition is the usual addition of complex numbers:

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2);$$

while scalar multiplication is given by

$$\alpha(x + iy) = (\alpha x) + i(\alpha y) \qquad (\text{for } \alpha \in \mathbb{R}).$$

The zero vector is the element $0 = 0 + i0 \in \mathbb{C}$.

(iii) A *polynomial* over the field F is an expression of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m,$$

for some $m \geqslant 0$, where $a_0, a_1, \dots, a_m \in F$ and where we ignore terms with 0 as the coefficient. The set of all polynomials over F is usually denoted by $F[x]$. If necessary we can "pad" such an expression for a polynomial using 0 as the coefficient for the extra terms to increase its length. Thus to add $f(x)$ above to another polynomial $g(x)$, we may assume they are represented by expressions of the same length, say

$$g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_mx^m.$$

Then

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_m + b_m)x^m.$$

Scalar multiplication is straightforward:

$$\alpha f(x) = (\alpha a_0) + (\alpha a_1)x + (\alpha a_2)x^2 + \cdots + (\alpha a_m)x^m$$

for $f(x)$ as above and $\alpha \in F$. The vector space axioms are pretty much straightforward to verify. The zero vector is the polynomial with all coefficients 0:

$$0 = 0 + 0x + 0x^2 + \cdots + 0x^m$$

(for any choice of m) and

$$-f(x) = (-a_0) + (-a_1)x + (-a_2)x^2 + \cdots + (-a_m)x^m.$$

(iv) The final example is heavily related to the example of polynomials, which are after all special types of functions. Let $\mathcal{F}_\mathbb{R}$ denote the set of all functions $f \colon \mathbb{R} \to \mathbb{R}$. Define the addition of two functions f and g by

$$(f + g)(x) = f(x) + g(x) \qquad (\text{for } x \in \mathbb{R})$$

and scalar multiplication of f by $\alpha \in \mathbb{R}$ by

$$(\alpha f)(x) = \alpha \cdot f(x) \qquad (\text{for } x \in \mathbb{R}).$$

Then $\mathcal{F}_\mathbb{R}$ is a real vector space with

$$(-f)(x) = -f(x)$$

and $\mathbf{0}$ being the function given by $x \mapsto 0$ for all $x \in \mathbb{R}$.

These examples illustrate that vector spaces occur in numerous situations (functions covering a large class of mathematical objects for a start) and so the study of linear algebra is of considerable importance. We shall spend some time developing (and reviewing the development) of the theory of vector spaces.

Basic properties of vector spaces

In order to work with the vectors in a vector space, we record the basic properties we need.

Proposition 1.5 *Let V be a vector space over a field F. Let $v \in V$ and $\alpha \in F$. Then*

(i) $\alpha \mathbf{0} = \mathbf{0}$*;*

(ii) $0v = \mathbf{0}$*;*

(iii) *if* $\alpha v = \mathbf{0}$*, then either* $\alpha = 0$ *or* $v = \mathbf{0}$*;*

(iv) $(-\alpha)v = -\alpha v = \alpha(-v)$*.*

PROOF: [Omitted in lectures — appears in MT2501]

(i) Use condition (v) of Definition 1.3 to give

$$\alpha(\mathbf{0} + \mathbf{0}) = \alpha\mathbf{0} + \alpha\mathbf{0};$$

that is,

$$\alpha\mathbf{0} = \alpha\mathbf{0} + \alpha\mathbf{0}.$$

Now add $-\alpha\mathbf{0}$ to both sides to yield

$$\mathbf{0} = \alpha\mathbf{0}.$$

(ii) Use condition (vi) of Definition 1.3 to give

$$(0 + 0)v = 0v + 0v;$$

that is,

$$0v = 0v + 0v.$$

Now add $-0v$ to deduce $\mathbf{0} = 0v$.

(iii) Suppose $\alpha v = \mathbf{0}$, but that $\alpha \neq 0$. Then F contains the scalar α^{-1} and multiplying by this gives

$$\alpha^{-1}(\alpha v) = \alpha^{-1}\mathbf{0} = \mathbf{0} \qquad \text{(using (i))}.$$

Therefore

$$1v = (\alpha^{-1} \cdot \alpha)v = \mathbf{0}.$$

Condition (viii) of Definition 1.3 then shows $v = \mathbf{0}$. Hence if $\alpha v = \mathbf{0}$, either $\alpha = 0$ or $v = \mathbf{0}$.

(iv)

$$\alpha v + (-\alpha)v = (\alpha + (-\alpha))v = 0v = \mathbf{0},$$

so $(-\alpha)v$ is the vector which when added to αv yields $\mathbf{0}$; that is, $(-\alpha)v = -\alpha v$.

Similarly,

$$\alpha v + \alpha(-v) = \alpha(v + (-v)) = \alpha\mathbf{0} = \mathbf{0}$$

and we deduce that $\alpha(-v)$ must be the vector $-\alpha v$. □

Subspaces

Although linear algebra is a branch of mathematics that is used throughout the whole spectrum of pure and applied mathematics, it is nonetheless a branch of algebra. As a consequence, we should expect to do the sort of thing that is done throughout algebra, namely examine substructures and structure preserving maps. For the former, we make the following definition.

Definition 1.6 Let V be a vector space over a field F. A *subspace* W of V is a non-empty subset such that

(i) if $u, v \in W$, then $u + v \in W$, and

(ii) if $v \in W$ and $\alpha \in F$, then $\alpha v \in W$.

Thus a subspace W is a non-empty subset of the vector space V such that W is closed under vector addition and scalar multiplication by any scalar from the field F. The following basic properties hold:

Lemma 1.7 *Let V be a vector space and let W be a subspace of V. Then*

(i) $\mathbf{0} \in W$*;*

(ii) *if $v \in W$, then $-v \in W$.*

Proof: (i) Since W is non-empty, there exists at least one vector $u \in W$. Now W is closed under scalar multiplication, so $0u \in W$; that is, $\mathbf{0} \in W$ (by Proposition 1.5(i)).

(ii) Let v be any vector in W. Then W contains

$$(-1)v = -1v = -v.$$

$\square$

Consequence: If W is a subspace of V (over a field F), then W is also a vector space of F: if u, v are elements of W and $\alpha \in F$, then $u + v$, $-v$ and αv are defined elements of W. We have a zero vector $\mathbf{0} \in W$ and the axioms are all inherited from the fact that they hold universally on all vectors in V.

Example 1.8 Many examples of subspaces were presented in MT2501. We list a few here with full details, but these details will probably be omitted during the lectures.

(i) Let $V = \mathbb{R}^3$, the real vector space of column vectors of length 3. Consider

$$W = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \;\middle|\; x, y \in \mathbb{R} \right\} \subseteq \mathbb{R}^3;$$

so W consists of all vectors with zero in the last entry. We check

$$\begin{pmatrix} x_1 \\ y_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ 0 \end{pmatrix} \in W$$

and

$$\alpha \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha x \\ \alpha y \\ 0 \end{pmatrix} \in W \qquad (\text{for } \alpha \in \mathbb{R}).$$

Thus W is closed under sums and scalar multiplication; that is, W is a subspace of $\mathbb{R}^3$.

(ii) Let $\mathcal{F}_{\mathbb{R}}$ be the set of all functions $f \colon \mathbb{R} \to \mathbb{R}$, which forms a real vector space under

$$(f+g)(x) = f(x) + g(x); \qquad (\alpha f)(x) = \alpha \cdot f(x).$$

Let $\mathcal{P}$ denote the set of polynomial functions; i.e., each $f \in \mathcal{P}$ has the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

for some $m \geqslant 0$ and $a_0, a_1, \dots, a_m \in \mathbb{R}$. Then $\mathcal{P} \subseteq \mathcal{F}_{\mathbb{R}}$ and, since the sum of two polynomials is a polynomial and a scalar multiple of a polynomial is a polynomial, $\mathcal{P}$ is a subspace of $\mathcal{F}_{\mathbb{R}}$.

We shall meet a generic way of constructing subspaces in a short while. The following specifies basic ways of manipulating subspaces.

Definition 1.9 Let V be a vector space and let U and W be subspaces of V.

(i) The *intersection* of U and W is

$$U \cap W = \{\, v \mid v \in U \text{ and } v \in W \,\}.$$

(ii) The *sum* of U and W is

$$U + W = \{\, u + w \mid u \in U, w \in W \,\}.$$

Since V is a vector space, addition of a vector $u \in U \subseteq V$ and $w \in W \subseteq V$ makes sense. Thus the sum $U + W$ is a sensible collection of vectors in V.

Proposition 1.10 *Let V be a vector space and let U and W be subspaces of V. Then*

(i) *$U \cap W$ is a subspace of V;*

(ii) *$U + W$ is a subspace of V.*

Proof: (i) First note that Lemma 1.7(i) tells us that $\mathbf{0}$ lies in both U and W. Hence $\mathbf{0} \in U \cap W$, so this intersection is non-empty. Let $u, v \in U \cap W$ and α be a scalar from the base field. Then U is a subspace containing u and v, so $u + v \in U$ and $\alpha v \in U$. Equally, $u, v \in W$ so we deduce $u + v \in W$ and $\alpha v \in W$. Hence $u + v \in U \cap W$ and $\alpha v \in U \cap W$. This shows $U \cap W$ is a subspace of V.

(ii) Using the fact that $\mathbf{0}$ lies in U and W, we see $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U + W$. Hence $U + W$ is non-empty. Now let $v_1, v_2 \in U + W$, say $v_1 = u_1 + w_1$ and $v_2 = u_2 + w_2$ where $u_1, u_2 \in U$ and $w_1, w_2 \in W$. Then

$$v_1 + v_2 = (u_1 + w_1) + (u_2 + w_2) = (u_1 + u_2) + (w_1 + w_2) \in U + W$$

and if α is a scalar then

$$\alpha v_1 = \alpha(u_1 + w_1) = (\alpha u_1) + (\alpha w_1) \in U + W.$$

Hence $U + W$ is a subspace of V. □

A straightforward induction argument then establishes:

Corollary 1.11 *Let V be a vector space and let U_1, U_2, …, U_k be subspaces of V. Then*

$$U_1 + U_2 + \cdots + U_k = \{\, u_1 + u_2 + \cdots + u_k \mid u_i \in U_i \text{ for each } i \,\}$$

is a subspace of V.

Spanning sets

We have defined earlier what is meant by a subspace. We shall now describe a good way (indeed, probably the canonical way) to specify subspaces.

Definition 1.12 Let V be a vector space over a field F and suppose that $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ is a set of vectors in V. A *linear combination* of these vectors is a vector of the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

for some $\alpha_1, \alpha_2, \dots, \alpha_k \in F$. The set of all such linear combinations is called the *span* of the vectors v_1, v_2, …, v_k and is denoted by $\mathrm{Span}(v_1, v_2, \dots, v_k)$ or by $\mathrm{Span}(\mathscr{A})$.

Remarks

(i) We shall often use the familiar summation notation to abbreviate a linear combination:

$$\sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k.$$

(ii) In some settings, one might (and certainly some authors do) write $\langle \mathscr{A} \rangle$ or $\langle v_1, v_2, \dots, v_k \rangle$ for the span of $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$. We will not do so in this course, since we wish to reserve angled brackets for inner products later in the course.

(iii) When $\mathscr{A}$ is an infinite set of vectors in a vector space V (over a field F), we need to apply a little care when defining $\mathrm{Span}(\mathscr{A})$. It does not make sense to add together infinitely many vectors: our addition only allows us to combine two vectors at a time. Consequently for an arbitrary set $\mathscr{A}$ of vectors we make the following definition:

$$\mathrm{Span}(\mathscr{A}) = \left\{ \sum_{i=1}^{k} \alpha_i v_i \;\middle|\; v_1, v_2, \dots, v_k \in \mathscr{A},\ \alpha_1, \alpha_2, \dots, \alpha_k \in F \right\}.$$

Thus $\mathrm{Span}(\mathscr{A})$ is the set of all linear combinations formed by selecting finitely many vectors from $\mathscr{A}$. When $\mathscr{A}$ is finite, this coincides with Definition 1.12.

Proposition 1.13 *Let $\mathscr{A}$ be a set of vectors in the vector space V. Then* $\mathrm{Span}(\mathscr{A})$ *is a subspace of V.*

PROOF: [Omitted in lectures — appears in MT2501]

We prove the proposition for the case when $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ is a finite set of vectors from V. The case when $\mathscr{A}$ is infinite requires few changes.

First note that, taking $\alpha_i = 0$ for each i, we see

$$\mathbf{0} = \sum_{i=1}^{k} 0 v_i \in \mathrm{Span}(\mathscr{A}).$$

Let $u, v \in \mathrm{Span}(\mathscr{A})$, say

$$u = \sum_{i=1}^{k} \alpha_i v_i \qquad \text{and} \qquad v = \sum_{i=1}^{k} \beta_i v_i$$

where the α_i and β_i are scalars. Then

$$u + v = \sum_{i=1}^{k} (\alpha_i + \beta_i) v_i \in \mathrm{Span}(\mathscr{A})$$

and if γ is a further scalar then

$$\gamma v = \sum_{i=1}^{k} (\gamma \alpha_i) v_i \in \operatorname{Span}(\mathscr{A}).$$

Thus $\operatorname{Span}(\mathscr{A})$ is a non-empty subset of V which is closed under addition and scalar multiplication; that is, it is a subspace of V. $\square$

It is fairly easy to see that if W is a subspace of a vector space V, then $W = \operatorname{Span}(\mathscr{A})$ for some choice of $\mathscr{A} \subseteq W$. Indeed, the fact that W is closed under addition and scalar multiplication ensures that linear combinations of its elements are again in W and hence $W = \operatorname{Span}(W)$. However, what we will typically want to do is seek sets $\mathscr{A}$ which span particular subspaces where $\mathscr{A}$ can be made reasonably small.

Definition 1.14 A *spanning set* for a subspace W is a set $\mathscr{A}$ of vectors such that $\operatorname{Span}(\mathscr{A}) = W$.

Thus if $\operatorname{Span}(\mathscr{A}) = W$, then each element of W can be written in the form

$$v = \sum_{i=1}^{k} \alpha_i v_i$$

where $v_1, v_2, \dots, v_k \in \mathscr{A}$ and $\alpha_1, \alpha_2, \dots, \alpha_k$ are scalars from the base field. We seek to find efficient choices of spanning sets $\mathscr{A}$; i.e., make $\mathscr{A}$ as small as possible.

Example 1.15 (i) Since every vector in $\mathbb{R}^3$ has the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

we conclude that

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a spanning set for $\mathbb{R}^3$. However, note that although this is probably the most natural spanning set, it is not the only one. For example,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{x+y}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{x-y}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

so

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is also a spanning set.

We can also add vectors to a set that already spans and produce yet another spanning set. (Though there will inevitably be a level of redundancy to this and this relates to the concept of linear independence which we address next.) For example,

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\}$$

is a spanning set for $\mathbb{R}^3$, since every vector in $\mathbb{R}^3$ can be written as a linear combination of the vectors appearing (in multiple ways); for example,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(x + \frac{z}{2}\right) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (y+z) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{z}{2} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{z}{2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

(ii) Recall the vector space $F[x]$ of polynomials over the field F; its elements have the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m.$$

We can therefore write

$$f(x) = a_0 f_0(x) + a_1 f_1(x) + a_2 f_2(x) + \cdots + a_m f_m(x)$$

where $f_i(x) = x^i$ for $i = 0, 1, 2, \ldots$. Hence the set

$$\mathscr{M} = \{1, x, x^2, x^3, \ldots\}$$

of all *monomials* is a spanning set for $F[x]$.

Linear independent elements and bases

We described spanning sets in the previous section. When seeking to make these as efficient as possible, we will want to use linear independent spanning sets.

Definition 1.16 Let V be a vector space over a field F. A set $\mathscr{A} = \{v_1, v_2, \ldots, v_k\}$ is called *linearly independent* if the only solution to the equation

$$\sum_{i=1}^{k} \alpha_i v_i = \mathbf{0}$$

(with $\alpha_i \in F$) is $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.

If $\mathscr{A}$ is not linearly independent, we shall call it *linearly dependent.*

The method to check whether a set of vectors is linearly independent was covered in MT2501 and is fairly straightforward. Take the set of vectors $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ and consider the equation $\sum_{i=1}^{k} \alpha_i v_i = \mathbf{0}$. This will usually convert to a system of linear equations in the variables α_i. The usual method of solving systems of linear equations (such as applying row operations to the matrix associated to the system) determines whether $\alpha_i = 0$ for all i is the only solution or whether there are non-trivial solutions.

Example 1A *Determine whether the set $\{x + x^2, 1 - 2x^2, 3 + 6x\}$ is linearly independent in the vector space $\mathcal{P}$ of all real polynomials.*

Solution: We solve

$$\alpha(x + x^2) + \beta(1 - 2x^2) + \gamma(3 + 6x) = 0;$$

that is,

$$(\beta + 3\gamma) + (\alpha + 6\gamma)x + (\alpha - 2\beta)x^2 = 0. \tag{1.1}$$

Equating coefficients yields the system of equations

$$\begin{aligned} \beta + 3\gamma &= 0 \\ \alpha \qquad + 6\gamma &= 0 \\ \alpha - 2\beta \qquad &= 0; \end{aligned}$$

that is,

$$\begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 6 \\ 1 & -2 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A sequence of row operations (Check!) converts this to

$$\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence the original equation (1.1) is equivalent to

$$\begin{aligned} \alpha - 2\beta \qquad &= 0 \\ \beta + 3\gamma &= 0. \end{aligned}$$

Since there are fewer equations remaining than the number of variables, we have enough freedom to produce a non-zero solution. For example, if we set $\gamma = 1$, then $\beta = -3\gamma = -3$ and $\alpha = 2\beta = -6$. Hence the set $\{x + x^2, 1 - 2x^2, 3 + 6x\}$ is *linearly dependent*. □

In general, if $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ is linearly dependent, then we have

$$\sum_{i=1}^{k} \alpha_i v_i = \mathbf{0}$$

with not all $\alpha_i \in F$ equal to zero. Suppose $\alpha_j \neq 0$ and rearrange the equation to

$$\alpha_j v_j = -(\alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \cdots + \alpha_k v_j).$$

Therefore

$$\begin{aligned} v_j = (-\alpha_1/\alpha_j)v_1 + \cdots + (-\alpha_{j-1}/\alpha_j)v_{j-1} \\ + (-\alpha_{j+1}/\alpha_j)v_{j+1} + \cdots + (-\alpha_k/\alpha_j)v_k, \end{aligned}$$

so v_j can be expressed as a linear combination of the other vectors in $\mathscr{A}$. Equally such an expression can be rearranged into an equation of linear dependence for $\mathscr{A}$. Hence we have the following important observation:

Lemma 1.17 *Let $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ be a set of vectors in the vector space V. Then $\mathscr{A}$ is linearly independent if and only if no vector in $\mathscr{A}$ can be expressed as a linear combination of the others.* $\square$

Suppose that $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ is a finite set of vectors which is *linearly dependent* and let $W = \mathrm{Span}(\mathscr{A})$. Now as $\mathscr{A}$ is linear dependent, Lemma 1.17 says that one of the vectors in $\mathscr{A}$ is a linear combination of the others. Let us suppose without loss of generality (and for notational convenience only) that v_k is such a vector. Say

$$v_k = \sum_{i=1}^{k-1} \alpha_i v_i$$

for some scalars α_i. Put $\mathscr{B} = \{v_1, v_2, \dots, v_{k-1}\} \subseteq \mathscr{A}$.

Claim: $\mathrm{Span}(\mathscr{B}) = \mathrm{Span}(\mathscr{A}) = W$.

PROOF: Since $\mathscr{B} \subseteq \mathscr{A}$, it is clear that $\mathrm{Span}(\mathscr{B}) \subseteq \mathrm{Span}(\mathscr{A})$: any linear combination of the vectors in $\mathscr{B}$ is a linear combination of those in $\mathscr{A}$ by taking 0 as the coefficient for v_k.

Conversely, if $w \in \mathrm{Span}(\mathscr{A})$, then

$$w = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k$$

for some scalars β_1, β_2, …, β_k. Hence

$$w = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_{k-1} v_{k-1} + \beta_k \sum_{i=1}^{k-1} \alpha_i v_i$$

$$
\begin{aligned}
&= \sum_{i=1}^{k-1} (\beta_i + \beta_k \alpha_i) v_i \\
&\in \operatorname{Span}(\mathscr{B}).
\end{aligned}
$$

This proves the claim and we have the following result. $\square$

Lemma 1.18 *Let $\mathscr{A}$ be a linear dependent set of vectors belonging to a vector space V. Then there exists some vector v in $\mathscr{A}$ such that $\mathscr{A} \setminus \{v\}$ spans the same subspace as $\mathscr{A}$.*

As we started with a finite set, repeating this process eventually stops, at which point we must have produced a linearly independent set. Hence, we conclude:

Theorem 1.19 *Let V be a vector space. If $\mathscr{A}$ is a finite subset of V and $W = \operatorname{Span}(\mathscr{A})$, then there exists a linearly independent subset $\mathscr{B}$ with $\mathscr{B} \subseteq \mathscr{A}$ and $\operatorname{Span}(\mathscr{B}) = W$.* $\square$

Thus we can pass from a finite spanning set to a linearly independent spanning set simply by omitting the correct choice of vectors. (The above tells us that we want to omit a vector that can be expressed as a linear combination of the others, and then repeat.) In particular, if the vector space V possesses a finite spanning set, then it possesses a linearly independent spanning set. Accordingly, we make the following definition:

Definition 1.20 Let V be a vector space over the field F. A *basis* for V is a linearly independent spanning set. We say that V is *finite-dimensional* if it possesses a finite spanning set; that is, if V possesses a finite basis. The *dimension* of V is the size of any basis for V and is denoted by $\dim V$.

Example 1.21 (i) The set

$$
\mathscr{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}
$$

is a basis for $V = F^n$. We shall call it the *standard basis* for F^n. Hence $\dim F^n = n$ (as one would probably expect). Throughout we shall write

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the 1 is in the ith position, so that $\mathscr{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$.

[Verification (omitted in lectures): If $\boldsymbol{v}$ is an arbitrary vector in V, say

$$\boldsymbol{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$

(where $x_i \in F$). Thus $\mathscr{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ spans V. Suppose there exist scalars α_1, α_2, $\dots$, α_n such that

$$\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = \mathbf{0};$$

that is,

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hence $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. Thus $\mathscr{B}$ is linearly independent.]

(ii) Let $\mathcal{P}_n$ be the set of polynomials over the field F of degree at most n:

$$\mathcal{P}_n = \{\, f(x) \mid f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \quad \text{for some } a_i \in F \,\}.$$

It is easy to check $\mathcal{P}_n$ is closed under sums and scalar multiples, so $\mathcal{P}_n$ forms a vector subspace of the space $F[x]$ of all polynomials. The set of monomials $\{1, x, x^2, \dots, x^n\}$ is a basis for $\mathcal{P}_n$. Hence $\dim \mathcal{P}_n = n + 1$.

In these examples we have referred to dimension as though it is uniquely determined. In Corollary 1.24 we shall show that it is. Beforehand, however, we shall observe how bases are efficient as spanning sets, since they produce a uniqueness to the linear combinations required.

Lemma 1.22 *Let V be a vector space of dimension n and suppose that $\mathscr{B} = \{v_1, v_2, \dots, v_n\}$ is a basis for V. Then every vector in V can be expressed as a linear combination of the vectors in $\mathscr{B}$ in a unique way.*

PROOF: Let $v \in V$. Since $\mathscr{B}$ is a spanning set for V, we can certainly express v as a linear combination of the vectors in $\mathscr{B}$. Suppose we have two expressions for v:

$$v = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i v_i.$$

Hence

$$\sum_{i=1}^{n} (\alpha_i - \beta_i) v_i = \mathbf{0}.$$

Since the set $\mathscr{B}$ is linearly independent, we deduce $\alpha_i - \beta_i = 0$ for all i; that is, $\alpha_i = \beta_i$ for all i. Hence our linear combination expression for v is indeed unique. $\square$

Theorem 1.23 *Let V be a finite-dimensional vector space. Suppose that $\{v_1, v_2, \dots, v_m\}$ is a linearly independent set of vectors and $\{w_1, w_2, \dots, w_n\}$ is a spanning set for V. Then*

$$m \leqslant n.$$

The proof of this theorem appears only in the lecture notes, but it will not be presented in the actual lectures. Before giving the proof, we first record the important consequence which we referred to above.

Corollary 1.24 *Let V be a finite-dimensional vector space. Then any two bases for V have the same size and consequently $\dim V$ is uniquely determined.*

PROOF: If $\{v_1, v_2, \dots, v_m\}$ and $\{w_1, w_2, \dots, w_n\}$ are bases for V, then they are both linearly independent and spanning sets for V, so Theorem 1.23 applied twice gives

$$m \leqslant n \qquad \text{and} \qquad n \leqslant m.$$

Hence $m = n$. $\square$

Let us now turn to the proof of the main theorem about linearly independent sets and spanning sets.

PROOF OF THEOREM 1.23: [Omitted in lectures]

Let V be a finite-dimensional vector space over a field F. We shall assume that $\{v_1, v_2, \dots, v_m\}$ is a linearly independent set in V and $\{w_1, w_2, \dots, w_n\}$

is a spanning set for V. In particular, there exist $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$ such that

$$v_1 = \sum_{i=1}^{n} \alpha_i w_i.$$

Since $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, certainly $v_1 \neq \mathbf{0}$. Thus some of the α_i are non-zero. By re-arranging the w_i, there is no loss of generality in assuming $\alpha_1 \neq 0$. Then

$$w_1 = \frac{1}{\alpha_1}\left(v_1 - \sum_{i=2}^{n} \alpha_i w_i\right)$$

and this enables us to replace w_1 in any expression by a linear combination of v_1 and $w_2, \ldots, w_n$. Since $V = \operatorname{Span}(w_1, w_2, \ldots, w_n)$, we now deduce

$$V = \operatorname{Span}(v_1, w_2, \ldots, w_n).$$

Suppose that we manage to show

$$V = \operatorname{Span}(v_1, v_2, \ldots, v_j, w_{j+1}, \ldots, w_n)$$

for some value of j where $j < m, n$. Then v_{j+1} is a vector in V, so can be expressed as a linear combination of v_1, v_2, …, v_j, w_{j+1}, …, w_n, say

$$v_{j+1} = \sum_{i=1}^{j} \beta_i v_i + \sum_{i=j+1}^{n} \beta_i w_i$$

for some scalars $\beta_i \in F$. Now if $\beta_{j+1} = \cdots = \beta_n = 0$, then we would have

$$v_{j+1} = \sum_{i=1}^{j} \beta_i v_i,$$

which would contradict the set $\{v_1, v_2, \ldots, v_n\}$ being linearly independent (see Lemma 1.17). Hence some β_i, with $i \geqslant j+1$, is non-zero. Re-arranging the w_i (again), we can assume that it is $\beta_{j+1} \neq 0$. Hence

$$w_{j+1} = \frac{1}{\beta_{j+1}}\left(v_{j+1} - \sum_{i=1}^{j} \beta_i v_i - \sum_{i=j+2}^{n} \beta_i w_i\right).$$

Consequently, we can replace w_{j+1} by a linear combination of the vectors v_1, v_2, …, v_{j+1}, w_{j+2}, …, w_n. Therefore

$$\begin{aligned} V &= \operatorname{Span}(v_1, v_2, \ldots, v_j, w_{j+1}, w_{j+2}, \ldots, w_n) \\ &= \operatorname{Span}(v_1, v_2, \ldots, v_{j+1}, w_{j+2}, \ldots, w_n). \end{aligned}$$

If it were the case that $m > n$, then this process stops when we have replaced all the w_i by v_i and have

$$V = \mathrm{Span}(v_1, v_2, \dots, v_n).$$

But then v_{n+1} is a linear combination of v_1, v_2, ..., v_n, and this contradicts $\{v_1, v_2, \dots, v_m\}$ being linearly independent.

Consequently $m \leqslant n$, as required. $\square$

In two examples that finish this section, we shall illustrate two ways to build bases for subspaces. The first is the following:

Example 1.25 *Let*

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 3 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 3 \\ 1 \\ -6 \end{pmatrix}, \quad \boldsymbol{v}_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \quad \boldsymbol{v}_5 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

and let U be the subspace of $\mathbb{R}^4$ spanned by the set $\mathscr{A} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4, \boldsymbol{v}_5\}$. Find a basis $\mathscr{B}$ for U and hence determine the dimension of U.

Solution: Since $\dim \mathbb{R}^4 = 4$, the maximum size for a linearly independent set is 4 (by Theorem 1.23). This tells us that $\mathscr{A}$ is certainly not linearly independent; we need to find a linearly independent subset $\mathscr{B}$ of $\mathscr{A}$ that also spans U (see Theorem 1.19). This is done by finding which vectors in $\mathscr{A}$ can be written as a linear combination of the other vectors in $\mathscr{A}$ (see Lemma 1.17). We solve

$$\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2 + \gamma \boldsymbol{v}_3 + \delta \boldsymbol{v}_4 + \varepsilon \boldsymbol{v}_5 = \boldsymbol{0}; \tag{1.2}$$

that is,

$$\alpha \begin{pmatrix} 1 \\ -1 \\ 0 \\ 3 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 3 \\ 1 \\ -6 \end{pmatrix} + \delta \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} + \varepsilon \begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

or, equivalently,

$$\begin{pmatrix} 1 & 2 & 0 & 0 & -1 \\ -1 & 1 & 3 & 1 & 1 \\ 0 & 1 & 1 & 0 & -1 \\ 3 & 0 & -6 & -1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We apply row operation as follows to the appropriate augmented matrix:

$$
\left(\begin{array}{ccccc|c}
1 & 2 & 0 & 0 & -1 & 0 \\
-1 & 1 & 3 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 \\
3 & 0 & -6 & -1 & 0 & 0
\end{array}\right)
$$

$$
\longrightarrow \left(\begin{array}{ccccc|c}
1 & 2 & 0 & 0 & -1 & 0 \\
0 & 3 & 3 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 \\
0 & -6 & -6 & -1 & 3 & 0
\end{array}\right) \qquad \begin{array}{l} r_2 \mapsto r_2 + r_1 \\ r_4 \mapsto r_4 - 3r_1 \end{array}
$$

$$
\longrightarrow \left(\begin{array}{ccccc|c}
1 & 2 & 0 & 0 & -1 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 \\
0 & 3 & 3 & 1 & 0 & 0 \\
0 & -6 & -6 & -1 & 3 & 0
\end{array}\right) \qquad r_2 \longleftrightarrow r_3
$$

$$
\longrightarrow \left(\begin{array}{ccccc|c}
1 & 2 & 0 & 0 & -1 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 3 & 0 \\
0 & 0 & 0 & -1 & -3 & 0
\end{array}\right) \qquad \begin{array}{l} r_3 \mapsto r_3 - 3r_2 \\ r_4 \mapsto r_4 + 6r_2 \end{array}
$$

$$
\longrightarrow \left(\begin{array}{ccccc|c}
1 & 2 & 0 & 0 & -1 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right) \qquad r_4 \mapsto r_4 + r_3
$$

So our equation (1.2) is equivalent to:

$$
\begin{aligned}
\alpha + 2\beta \qquad\quad - \varepsilon &= 0 \\
\beta + \gamma \quad - \varepsilon &= 0 \\
\delta + 3\varepsilon &= 0
\end{aligned}
$$

Given arbitrary γ and ε, we can read off α, β and δ that solve the equation. Taking $\gamma = 1$, $\varepsilon = 0$ and $\gamma = 0$, $\varepsilon = 1$ tells us that the vectors $\boldsymbol{v}_3$ and $\boldsymbol{v}_5$ in $\mathscr{A}$ can be written as a linear combination of the others, as we shall now observe.

If $\gamma = 1$ and $\varepsilon = 0$, then the above tells us:

$$
\begin{aligned}
\delta &= -3\varepsilon = 0 \\
\beta &= -\gamma + \varepsilon = -1 \\
\alpha &= -2\beta + \varepsilon = 2
\end{aligned}
$$

Hence

$$2\boldsymbol{v}_1 - \boldsymbol{v}_2 + \boldsymbol{v}_3 = \mathbf{0},$$

so

$$\boldsymbol{v}_3 = -2\boldsymbol{v}_1 + \boldsymbol{v}_2. \tag{1.3}$$

If $\gamma = 0$ and $\varepsilon = 1$, then the above tells us:

$$\begin{aligned}\delta &= -3\varepsilon = -3\\ \beta &= -\gamma + \varepsilon = 1\\ \alpha &= -2\beta + \varepsilon = -1.\end{aligned}$$

Hence

$$-\boldsymbol{v}_1 + \boldsymbol{v}_2 - 3\boldsymbol{v}_4 + \boldsymbol{v}_5 = \mathbf{0},$$

so

$$\boldsymbol{v}_5 = \boldsymbol{v}_1 - \boldsymbol{v}_2 + 3\boldsymbol{v}_4. \tag{1.4}$$

Equations (1.3) and (1.4) tell us $\boldsymbol{v}_3, \boldsymbol{v}_5 \in \operatorname{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_4)$. Therefore any linear combination of the vectors in $\mathscr{A}$ can also be written as a linear combination of $\mathscr{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_4\}$ (using (1.3) and (1.4) to achieve this). Hence

$$U = \operatorname{Span}(\mathscr{A}) = \operatorname{Span}(\mathscr{B}).$$

We finish by observing that $\mathscr{B}$ is linearly independent. Solve

$$\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2 + \gamma\boldsymbol{v}_4 = \mathbf{0};$$

that is,

$$\alpha \begin{pmatrix} 1 \\ -1 \\ 0 \\ 3 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

or

$$\begin{aligned}\alpha + 2\beta \qquad &= 0\\ -\alpha \ + \beta + \gamma &= 0\\ \beta \qquad &= 0\\ 3\alpha \qquad - \gamma &= 0\end{aligned}$$

In this case, we can automatically read off the solution:

$$\beta = 0, \qquad \alpha = -2\beta = 0, \qquad \gamma = 3\alpha = 0.$$

Hence $\mathscr{B}$ is linearly independent. It follows that $\mathscr{B}$ is the required basis for U. Then

$$\dim U = |\mathscr{B}| = 3.$$

□

Before the final example, we shall make some important observations concerning the creation of bases for finite-dimensional vector spaces.

Suppose that V is a finite-dimensional vector space, say $\dim V = n$, and suppose that we already have a linearly independent set of vectors, say $\mathscr{A} = \{v_1, v_2, \dots, v_m\}$. If $\mathscr{A}$ happens to span V, then it is a basis for V (and consequently $m = n$).

If not, there exists some vector, which we shall call v_{m+1}, such that $v_{m+1} \notin \operatorname{Span}(\mathscr{A})$. Consider the set

$$\mathscr{A}' = \{v_1, v_2, \dots, v_m, v_{m+1}\}.$$

Claim: $\mathscr{A}'$ is linearly independent.

PROOF: Suppose

$$\sum_{i=1}^{m+1} \alpha_i v_i = \mathbf{0}.$$

If $\alpha_{m+1} \neq 0$, then

$$v_{m+1} = -\frac{1}{\alpha_{m+1}} \sum_{i=0}^{m} \alpha_i v_i \in \operatorname{Span}(\mathscr{A}),$$

which is a contradiction. Thus $\alpha_{m+1} = 0$, so

$$\sum_{i=1}^{m} \alpha_i v_i = \mathbf{0},$$

which implies $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$, as $\mathscr{A}$ is linearly independent. $\square$

Hence, if $\mathscr{A}$ is a linearly independent subset which does not span V then we can adjoin another vector to produce a larger linearly independent set.

Let us now repeat the process. This cannot continue forever, since Theorem 1.23 says there is a maximum size for a linearly independent set, namely $n = \dim V$. Hence we must eventually reach a linearly independent set containing $\mathscr{A}$ that does span V. This proves:

Proposition 1.26 *Let V be a finite-dimensional vector space. Then every linearly independent set of vectors in V can be extended to a basis for V by adjoining a finite number of vectors.* $\square$

Corollary 1.27 *Let V be a vector space of finite dimension n. If $\mathscr{A}$ is a linearly independent set containing n vectors, then $\mathscr{A}$ is a basis for V.*

PROOF: By Proposition 1.26, we can extend $\mathscr{A}$ to a basis $\mathscr{B}$ for V. But by Corollary 1.24, $\mathscr{B}$ contains $\dim V = n$ vectors. Hence we cannot have introduced any new vectors and so $\mathscr{B} = \mathscr{A}$ is the basis we have found. $\square$

Example 1.28 *Let $V = \mathbb{R}^4$. Show that the set*

$$\mathscr{A} = \left\{ \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix} \right\}$$

is a linearly independent set of vectors. Find a basis for $\mathbb{R}^4$ containing $\mathscr{A}$.

Solution: To show $\mathscr{A}$ is linearly independent, we suppose

$$\alpha_1 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

This yields four equations:

$$3\alpha_1 + \alpha_2 = 0, \quad \alpha_1 = 0, \quad 3\alpha_2 = 0, \quad 4\alpha_2 = 0.$$

Hence $\alpha_1 = \alpha_2 = 0$. Thus $\mathscr{A}$ is linearly independent.

We now seek to extend $\mathscr{A}$ to a basis of $\mathbb{R}^4$. We do so by first attempting to add the first vector of the standard basis for $\mathbb{R}^4$ to $\mathscr{A}$: Set

$$\mathscr{B} = \left\{ \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Suppose

$$\alpha_1 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore

$$3\alpha_1 + \alpha_2 + \alpha_3 = 0, \quad \alpha_1 = 0, \quad 3\alpha_2 = 0, \quad 4\alpha_2 = 0.$$

So $\alpha_1 = \alpha_2 = 0$ (from the second and third equations) and we deduce $\alpha_3 = -3\alpha_1 - \alpha_2 = 0$. Hence our new set $\mathscr{B}$ is linearly independent.

If we now attempt to adjoin $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ to $\mathscr{B}$ and repeat the above, we would find that we were unable to prove the corresponding α_i are non-zero. Indeed,

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \operatorname{Span}(\mathscr{B}).$$

Thus there is no need to adjoin the second standard basis vector to $\mathscr{B}$.

Now let us attempt to adjoin $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ to $\mathscr{B}$:

$$\mathscr{C} = \left\{ \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Suppose

$$\alpha_1 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence

$$\begin{aligned} 3\alpha_1 + \alpha_2 + \alpha_3 \quad &= 0 \\ \alpha_1 \qquad\qquad\quad &= 0 \\ 3\alpha_2 \quad + \alpha_4 &= 0 \\ 4\alpha_2 \qquad\quad &= 0 \end{aligned}$$

Therefore $\alpha_1 = \alpha_2 = 0$, from which we deduce $\alpha_3 = \alpha_4 = 0$. Thus we have produced a linearly independent set $\mathscr{C}$ of size 4. But $\dim \mathbb{R}^4 = 4$ and hence $\mathscr{C}$ must now be a basis for $\mathbb{R}^4$. $\square$