Chapter 3

Linear Independence and Bases

Spanning sets

In a vector space, we can multiply vectors by scalars and then add them. We shall investigate what happens if we apply such operations to some fixed set of vectors.

Throughout the following discussion, we fix a vector space V over a field F and let $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ be some fixed set of vectors in V.

Definition 3.1 A vector v is a *linear combination* of the vectors in $\mathscr{A}$ if there are scalars α_1, α_2, ..., α_k in F such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k.$$

The set of all linear combinations of the vectors in $\mathscr{A}$ is called the *span* of $\mathscr{A}$. We denote this by $\mathrm{Span}(\mathscr{A})$ or $\mathrm{Span}(v_1, v_2, \dots, v_k)$.

Theorem 3.2 *Let $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ be a set of vectors in a vector space V (over a field F). Then the span of $\mathscr{A}$ is a subspace of V.*

PROOF: Write $W = \mathrm{Span}(\mathscr{A})$. First taking $\alpha_i = 0$ for all i, we see that W contains

$$0v_1 + 0v_2 + \cdots + 0v_k = \mathbf{0}$$

(by Proposition 2.12(ii)). Hence at least W is non-empty.

Now let $v, w \in W$, so

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \quad \text{and} \quad w = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k$$

for some scalars α_i and β_i. Hence

$$v + w = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \cdots + (\alpha_k + \beta_k)v_k \in W$$

and if α is any scalar in F then

$$\alpha v = (\alpha\alpha_1)v_1 + (\alpha\alpha_2)v_2 + \cdots + (\alpha\alpha_k)v_k \in W.$$

Hence W is a subspace of V. $\square$

If a subspace W of V is expressible as $W = \text{Span}(v_1, v_2, \ldots, v_k)$, then we say that $\{v_1, v_2, \ldots, v_k\}$ is a *spanning set* or *generating set* for W, and that the set $\{v_1, v_2, \ldots, v_k\}$ spans W.

Example 3.3 *Define*

$$\mathbf{e}_1 = (1,0,0), \qquad \mathbf{e}_2 = (0,1,0), \qquad \mathbf{e}_3 = (0,0,1).$$

Show that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *is a spanning set for* $\mathbb{R}^3$.

We refer to this as the *standard* spanning set. It is easy to see that this generalizes: $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard spanning set for $\mathbb{R}^n$, where $\mathbf{e}_i$ is an n-tuple which has 1 in the i-th position and all other entries 0.

Solution: For $x, y, z, \in \mathbb{R}$,

$$x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 = x(1,0,0) + y(0,1,0) + z(0,0,1) = (x,y,z).$$

So any $(x,y,z) \in \mathbb{R}^3$ is a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and any element of $\text{Span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is in $\mathbb{R}^3$. Hence $\text{Span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \mathbb{R}^3$.

Note that, for any proper subspace W of $\mathbb{R}^3$ (i.e. subspace of $\mathbb{R}^3$ which is not $\mathbb{R}^3$ itself), the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ *do not* form a spanning set of W, even though every vector in W is a linear combination of $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. This is because $\text{Span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \mathbb{R}^3 \neq W$; specifically, there will be linear combinations of $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ which do not lie in W (including one of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$).

Example 3.4 *Consider the vector space* $M_{2\times 2}(F)$ *over* F. *Determine the subspace of* $M_{2\times 2}(F)$ *spanned by:*

- *the set* $\mathscr{A} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$;
- *the set* $\mathscr{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$.

Solution:

- A typical element in the space spanned by $\mathscr{A}$ has the form

  $$\alpha \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

  i.e. a 2×2 diagonal matrix. Any 2×2 diagonal matrix is clearly expressible as a linear combination of elements of $\mathscr{A}$, and so $\text{Span}(\mathscr{A})$ is the space of all diagonal 2×2 matrices.
- A typical element in the space spanned by $\mathscr{B}$ has the form

  $$\alpha \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

  i.e. an arbitrary element of $M_{2\times 2}(F)$, and any element of $M_{2\times 2}(F)$ is so expressible. Hence $\text{Span}(\mathscr{B})$ is $M_{2\times 2}(F)$ itself.

Notice that every element of $\mathrm{Span}(\mathscr{A})$ is a linear combination of vectors from $\mathscr{B}$ but $\mathrm{Span}(\mathscr{B}) \neq \mathrm{Span}(\mathscr{A})$.

Example 3.5 *Find a spanning set for the subspace of* $\mathbb{R}^3$ *given by*

$$W = \{(a, b, a+b) : a, b \in \mathbb{R}\}.$$

Why is $\{(1,0,0),(0,0,1),(0,1,1)\}$ *not a spanning set for* W*?*

Solution: We take a typical element $(a, b, a+b)$ of W and write it in the form $a(\ldots,\ldots,\ldots) + b(\ldots,\ldots,\ldots)$. We obtain

$$(a, b, a+b) = a(1,0,1) + b(0,1,1).$$

Now, every element of W is a linear combination of $\{(1,0,1),(0,1,1)\}$ and every such linear combination of these 2 vectors lies in W. So $W = \mathrm{Span}((1,0,1),(0,1,1))$.
For the second part, observe that we can write any element $(a, b, a+b) \in W$ as a linear combination of these 3 vectors:

$$(a, b, a+b) = a(1,0,0) + a(0,0,1) + b(0,1,1).$$

But we can also obtain the linear combination $a(1,0,0) + c(0,0,1) + b(0,1,1) = (a, b, b+c) \notin W$ if $a \neq c$.

Example 3.6 *The complex numbers* $\mathbb{C}$ *form a vector space over* $\mathbb{R}$. *Show that* $\{1, i\}$ *is a spanning set for* $\mathbb{C}$.

Solution: Every complex number can be written in the form $a + bi$ $(a, b \in \mathbb{R})$, and every such expression is a complex number.

The following example will have a special role to play in this course.

Definition 3.7 Let A be an $m \times n$ matrix with entries from a field F.

(i) The rows of A can be viewed as (row) vectors in F^n. The subspace of F^n spanned by these rows is called the *row-space* of A.

(ii) The columns of A can be viewed as (column) vectors in F^m. The subspace of F^m spanned by these columns is called the *column-space* of A.

Example 3.8 *Consider the* 3×3 *identity matrix* I. *Its row-space is spanned by its three rows*

$$(1,0,0),(0,1,0),(0,0,1)$$

(viewed as elements of $\mathbb{R}^3$*). But this is the standard spanning set for the whole of* $\mathbb{R}^3$, *as seen in Example 3.3, and so the row-space of* I *is* $\mathbb{R}^3$ *itself. Analogously, its column-space is also* $\mathbb{R}^3$.

Proposition 3.9 *Let* $S = \{v_1, \ldots, v_i, \ldots, v_j, \ldots, v_s\}$ *be a set of vectors in vector space* V. *Suppose we alter* S *by:*

- *re-ordering the vectors: let* $S_1 := \{v_1, \ldots, v_j, \ldots, v_i, \ldots, v_s\}$;

- *multiplying a vector by a non–zero scalar: let* $S_2 = \{v_1, \ldots, \beta v_i, \ldots, v_j, \ldots, v_s\}$ *(*$\beta \neq 0$*);*
- *modifying a vector by adding to it a scalar multiple of another vector: let* $S_3 = \{v_1, \ldots, v_i + \gamma v_j, \ldots, v_j, \ldots, v_s\}$ *(*$\gamma \neq 0$*).*

Then each of S_1, S_2, S_3 *spans the same space as* S*.*

PROOF: Exercise. □

For example, we saw in Example 1.34 that $\begin{pmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix}$ is row-equivalent to the identity I, so $\mathrm{Span}((1,3,3),(1,4,3),(1,3,4)) = \mathrm{Span}((1,0,0),(0,1,0),(0,0,1)) = \mathbb{R}^3$.

Theorem 3.10 *(i) Two matrices that are row-equivalent to each other have the same row-space.*

(ii) Two matrices that are column-equivalent to each other have the same column-space.

PROOF: Apply Proposition 3.9. □

Linear independence

Definition 3.11 Suppose that V is a vector space over a field F. A set $\mathscr{A} = \{v_1, v_2, \ldots, v_k\}$ of vectors is called *linearly independent* if the only solution to the equation

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \mathbf{0}$$

(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*. We shall often abbreviate linearly independent to LI, and linearly dependent to LD.

Note that: if a set $\{v_1, \ldots, v_k\}$ of vectors is linearly independent then none is the zero vector. For, if $v_i = \mathbf{0}$, then we can write

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_k = \mathbf{0}.$$

Proposition 3.12 *Let* $\mathscr{A}$ *be a set of vectors in a vector space* V*.*

(i) *The set* $\mathscr{A}$ *is linearly independent if and only if no vector in the set can be expressed as a linear combination of the others.*

(ii) *The set* $\mathscr{A}$ *is linearly dependent if and only if some vector in the set can be expressed as a linear combination of the others.*

PROOF: The two statements are equivalent. We shall prove the second.

Suppose $\mathscr{A} = \{v_1, v_2, \ldots, v_k\}$ is linearly dependent. This means that there exist scalars α_1, α_2, …, α_k, not all zero, such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \mathbf{0}.$$

Let us suppose that it is α_j that is non-zero. Rearrange the previous equation to

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

Therefore

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

Hence v_j is a linear combination of the other vectors $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k$ in $\mathscr{A}$.

Conversely, suppose one of the vectors in $\mathscr{A} = \{v_1, v_2, \ldots, v_k\}$ is a linear combination of the others, say

$$v_j = \beta_1 v_1 + \cdots + \beta_{j-1} v_{j-1} + \beta_{j+1} v_{j+1} + \cdots + \beta_k v_k.$$

Rearranging, we obtain

$$\beta_1 v_1 + \cdots + \beta_{j-1} v_{j-1} + (-1) v_j + \beta_{j+1} v_{j+1} + \cdots + \beta_k v_k = \mathbf{0}.$$

This is an equation expressing the linear dependence of the vectors in $\mathscr{A}$ since not all coefficients are zero (the v_j has -1 as its coefficient). Hence $\mathscr{A}$ is linearly dependent. $\square$

Example 3.13 *Show that the vectors*

$$\mathbf{e}_1 = (1,0,0), \qquad \mathbf{e}_2 = (0,1,0), \qquad \mathbf{e}_3 = (0,0,1)$$

in $\mathbb{R}^3$ *are linearly independent.*

Solution: Consider the equation

$$\alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 = \mathbf{0}.$$

We need to show that the only solution for α_1, α_2 and α_3 is the zero solution. The left-hand side of the equation equals

$$\alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 = \alpha_1(1,0,0) + \alpha_2(0,1,0) + \alpha_3(0,0,1) = (\alpha_1, \alpha_2, \alpha_3).$$

Hence we solve

$$(\alpha_1, \alpha_2, \alpha_3) = (0,0,0),$$

which forces

$$\alpha_1 = \alpha_2 = \alpha_3 = 0.$$

Hence the set of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is linearly independent.

Example 3.14 *Let V be the vector space $M_{2\times 2}(F)$ of 2×2 matrices over the field F. Show that the following are LI sets of vectors in the space:*

- $A := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$;
- $B := \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\}$

Solution: For A, we solve

$$\alpha_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

that is,

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

From this, it is immediate that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$, so the set is LI.
For B, we solve

$$\alpha_1 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

that is,

$$\begin{pmatrix} \alpha_1 + \alpha_2 & \alpha_2 + \alpha_3 \\ \alpha_1 + 2\alpha_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This forces $\alpha_1 = -\alpha_2$, $\alpha_2 = -\alpha_3$ and $\alpha_1 = -2\alpha_2$. Hence we must have $\alpha_2 = 0$ and thus $\alpha_1 = \alpha_3 = 0$. So the set is LI.

Next, we give a method for determining the linear independence (or otherwise) of n vectors in $\mathbb{R}^n$.

Example 3.15 *Determine whether the set*

$$\mathscr{A} = \{(1,2,1),(1,1,2),(1,0,1)\}$$

is a linearly independent set in $\mathbb{R}^3$.

Solution: We consider the equation

$$\alpha_1(1,2,1) + \alpha_2(1,1,2) + \alpha_3(1,0,1) = (0,0,0)$$

where $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. This is equivalent to the system of equations

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 0 \\ 2\alpha_1 + \alpha_2 + 0 &= 0 \\ \alpha_1 + 2\alpha_2 + \alpha_3 &= 0 \end{aligned}$$

which we may rewrite as

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.1}$$

(Notice that the matrix is formed by writing the vectors in $\mathscr{A}$ as columns.)

Clearly $\alpha_1 = \alpha_2 = \alpha_3 = 0$ is a solution to this equation. We are interested in whether this is the *only* solution. We know that Equation (3.1) has a unique solution if and only if the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

has non-zero determinant, i.e. is invertible. We calculate

$$\begin{aligned}\det M &= \det\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix} - \det\begin{pmatrix}2 & 0\\ 1 & 1\end{pmatrix} + \det\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}\\ &= 1-2+(4-1)\\ &= 2 \neq 0.\end{aligned}$$

Hence M is invertible, so Equation (3.1) has a unique solution, namely $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore $\mathscr{A}$ is linearly independent.

Theorem 3.16 *The non-zero rows in an $m \times n$ echelon matrix form a linearly independent set of vectors.*

Proof: Let $E \in M_{m\times n}(F)$ be an echelon matrix and consider its rows as vectors in F^n. If E is the all-zero matrix, its set of non-zero rows is empty, and so trivially linearly independent. Now suppose E possesses at least one non-zero row; label the non-zero rows from top downwards as $R_1, R_2, \ldots, R_l$. Suppose we have the following equation (the vectors are row matrices, i.e. elements of F^n, and the λ_i are scalars):

$$\lambda_1 R_1 + \lambda_2 R_2 + \cdots + \lambda_l R_l = \mathbf{0}$$

Consider the position in R_1 containing its leading entry; this position contains the entry zero for each of $R_2, \ldots, R_l$, so

$$\lambda_1 \times \text{ (leading entry of } R_1\text{) } = 0,$$

i.e. $\lambda_1 = 0$. Hence

$$\lambda_2 R_2 + \cdots + \lambda_l R_l = \mathbf{0}$$

and we may repeat the argument, this time for the leading entry of R_2, to see that $\lambda_2 = 0$. Proceeding thus, we eventually obtain that all the scalars $\lambda_1, \lambda_2, \ldots, \lambda_l$ equal zero, and so the row vectors of E form an LI set. □

An illustration is the set of rows of the echelon matrix in Example 1.10, given by $\{(1,1,1,3),(0,2,4,-1),(0,0,0,\frac{5}{2})\}$; these are visibly LI due to the placement of the zeros.

Proposition 3.17 *Let $S = \{v_1, \ldots, v_i, \ldots, v_j, \ldots, v_s\}$ be a set of vectors in vector space V. Suppose we alter S by:*

- *re-ordering the vectors: let $S_1 := \{v_1, \ldots, v_j, \ldots, v_i, \ldots, v_s\}$;*
- *modifying a vector by a non–zero scalar: let $S_2 = \{v_1, \ldots, \beta v_i, \ldots, v_j, \ldots, v_s\}$ ($\beta \neq 0$);*
- *modifying a vector by adding to it a scalar multiple of another vector: let $S_3 = \{v_1, \ldots, v_i + \gamma v_j, \ldots, v_j, \ldots, v_s\}$ ($\gamma \neq 0$).*

Then S, S_1, S_2, S_3 are either all LI or all LD.

Proof: Exercise. □

An application of this is that, given a set of vectors which we wish to test for linear independence, we can form a matrix using the vectors as rows, then apply elementary row operations and row-reduce to (reduced) echelon form. If we obtain any all-zero rows, we know the set of rows is linearly dependent, and hence that the original set of vectors was linearly dependent as well. If we have no all-zero rows in our echelon form, our previous result tells us that the whole set is linearly independent.

Crucially, by our earlier result, the new set of rows spans the same space as the originals! Thus, in the linearly dependent case, we can find a linearly independent set of rows spanning the same space as the original rows (namely the non-zero rows of the echelon form).

Example 3.18 *Is the set of vectors* $\{(2,0,-2,4),(1,3,2,-3),(5,6,1,0)\}$ *a linearly independent set in* $\mathbb{R}^4$*? If not, find an LI set which spans the same space as the original set.*

Solution: From this set, we form the 3×4 matrix

$$\begin{pmatrix} 2 & 0 & -2 & 4 \\ 1 & 3 & 2 & -3 \\ 5 & 6 & 1 & 0 \end{pmatrix}$$

Applying row operations reduces this to the following echelon form:

$$\begin{pmatrix} 1 & 3 & 2 & -3 \\ 0 & 1 & 1 & -\frac{5}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

It is clear that the three vectors forming the rows of the echelon matrix form an LD set, since one is the all-zero vector. Hence, by the above theorem which proves that applying row operations to a set does not change whether it is LI or LD, we see that the original set of vectors must have been LD. In fact it can be shown that they are related by:

$$3(2,0,-2,4) + 4(1,3,2-3) = 2(5,6,1,0).$$

And the space spanned by the original three vectors is spanned by the two vectors $\{(1,3,2,-3),(0,1,1,-\frac{5}{3})\}$, clearly LI.

Suppose we have n vectors in $\mathbb{R}^n$, and let M be the matrix formed by using them as rows. In the case where the vectors are LI, row-reduction yields the identity, whose rows are clearly LI (this is the case where $\det M \neq 0$). In the case where the vectors are LD, row-reduction yields an echelon matrix with at least one all-zero row (this is the case where $\det M = 0$).

Bases

We now put together the two concepts introduced so far in this chapter.

Definition 3.19 Let V be a vector space over a field F. A subset $\mathscr{B}$ of V is called a *basis* if

(i) $\mathscr{B}$ is a spanning set for V, and

(ii) $\mathscr{B}$ is linearly independent.

Throughout this course, we shall assume that any vector space we work with has a finite basis (that is, a basis containing only a finite number of vectors). Linear algebra can be done with infinite bases, but we shall avoid such complications here.

Theorem 3.20 *Let $\mathscr{B}$ be a basis for a vector space V. Then every vector in V can be expressed in* precisely one way *as a linear combination of the vectors in the basis $\mathscr{B}$.*

PROOF: Suppose that $\mathscr{B} = \{v_1, v_2, \dots, v_k\}$. Since $\mathscr{B}$ is, by definition, a spanning set for V, every vector in V is a linear combination of the vectors in $\mathscr{B}$. We need to use the fact that $\mathscr{B}$ is linearly independent to show that every such linear combination is unique.

Let $v \in V$ and suppose that we have two linear combination expressions for v in terms of $\mathscr{B}$, say

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k. \tag{3.2}$$

Rearranging the terms we obtain

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_k - \beta_k)v_k = \mathbf{0}.$$

However, $\mathscr{B}$ is linearly independent, so the only possible way that a linear combination of the vectors in $\mathscr{B}$ can equal the zero vector $\mathbf{0}$ is if all coefficients involved are zero. Hence

$$\alpha_1 - \beta_1 = \alpha_2 - \beta_2 = \cdots = \alpha_k - \beta_k = 0;$$

that is,

$$\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2, \quad \dots, \quad \alpha_k = \beta_k.$$

Hence the coefficients occurring in Equation (3.2) are the same and we conclude that every linear combination expression for v in terms of the basis $\mathscr{B}$ is indeed unique. □

Example 3.21 *The set of unit vectors $\{\mathbf{e}_1 = (1,0,0), \mathbf{e}_2 = (0,1,0), \mathbf{e}_3 = (0,0,1)\}$ forms a basis for $\mathbb{R}^3$. Indeed, we saw in Example 3.3 that this set spans $\mathbb{R}^3$, while Example 3.13 tells us that it is linearly independent.*

We also call this set the standard basis *for $\mathbb{R}^3$.*

Example 3.22 *Analogously, the standard basis for $\mathbb{R}^n$ ($n \in \mathbb{N}$) is given by*

$$\{\mathbf{e}_1 = (1,0,\dots,0), \mathbf{e}_2 = (0,1,0,\dots,0), \dots, \mathbf{e}_n = (0,\dots,0,1)\}.$$

Example 3.23 *The set of matrices*

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

is a basis for the vector space $M_{2\times 2}(F)$ of all 2×2 matrices over the field F.

Solution: We have already seen that $\mathscr{B}$ is a spanning set for the vector space $M_{2\times 2}(F)$ (Example 3.4) and that it is a linearly independent set (Example 3.14), and hence $\mathscr{B}$ is a basis for $M_{2\times 2}(F)$.

Example 3.24 *The complex numbers $\mathbb{C}$ form a vector space over $\mathbb{R}$. The set $\{1, i\}$ is a basis for $\mathbb{C}$.*

Solution: We have already shown that this is a spanning set. It is easy to see that it is also LI, since clearly $a + bi = 0$ $(a, b \in \mathbb{R})$ if and only if $a = b = 0$.

Example 3.25 *Show that $\{1, x, x^2\}$ forms a basis for the space $\mathcal{P}_2$ of all polynomials of degree at most 2 with real coefficients.*

Solution: Let $\mathscr{B} = \{1, x, x^2\}$. First note that if $f(x)$ is any polynomial in $\mathcal{P}_2$, then it has the form

$$f(x) = a_0 + a_1 x + a_2 x^2,$$

i.e. a linear combination of the polynomials 1, x and x^2, and any such combination is in $\mathcal{P}_2$. Hence $\mathscr{B}$ is a spanning set for $\mathcal{P}_2$.

To determine linear independence, note that $a_0 + a_1 x + a_2 x^2 = 0$ for all x if and only if $a_0 = a_1 = a_2 = 0$. Hence $\mathscr{B}$ is a linearly independent set.

Hence $\mathscr{B}$ is a basis for $\mathcal{P}_2$.

Example 3.26 *Find a basis for the subspace W of $\mathbb{R}^3$ given by*

$$W = \{(x, y, z) : 2x + 4y - 3z = 0\}.$$

Solution: An element $v = (x, y, z)$ lies in W precisely if it satisfies $2x + 4y - 3z = 0$, i.e. $z = \frac{2}{3}x + \frac{4}{3}y$. In other words, W is the set of elements of the form

$$(x, y, \frac{2}{3}x + \frac{4}{3}y)$$

$(x, y \in \mathbb{R})$. With a view to obtaining a spanning set, we can express this as

$$(x, y, \frac{2}{3}x + \frac{4}{3}y) = x(1, 0, \frac{2}{3}) + y(0, 1, \frac{4}{3}).$$

So the two vectors $\{(1, 0, \frac{2}{3}), (0, 1, \frac{4}{3})\}$ span W. Are they LI? Clearly neither is a multiple of the other, and so they do form an LI set. Hence these two vectors form a basis for W.

The next example shows that a basis for a vector space V need not be unique. (Indeed, very few vector spaces have a unique basis.)

Example 3.27 *Consider the following three polynomials*

$$p_1(x) = 1 + 2x + x^2, \qquad p_2(x) = 1 + x + 2x^2, \qquad p_3(x) = 1 + x^2.$$

Show that $\mathscr{C} = \{p_1(x), p_2(x), p_3(x)\}$ is a basis for $\mathcal{P}_2$.

Solution: We seek to show that every polynomial can be expressed in the form

$$\alpha p_1(x) + \beta p_2(x) + \gamma p_3(x)$$

(for this shows that $\mathscr{C}$ spans $\mathcal{P}_2$) and, moreover, that the coefficients α, β and γ are uniquely determined. The latter will ensure linear independence, since when we solve

$$\alpha p_1(x) + \beta p_2(x) + \gamma p_3(x) = 0 = 0 \cdot p_1(x) + 0 \cdot p_2(x) + 0 \cdot p_3(x)$$

the uniqueness forces $\alpha = \beta = \gamma = 0$, as required.

So let $f(x) = a + bx + cx^2 \in \mathcal{P}_2$ and solve

$$\alpha p_1(x) + \beta p_2(x) + \gamma p_3(x) = f(x); \tag{3.3}$$

that is,

$$\alpha(1 + 2x + x^2) + \beta(1 + x + 2x^2) + \gamma(1 + x^2) = a + bx + cx^2.$$

Equating coefficients gives the system of three linear equations:

$$\begin{aligned} \alpha + \beta + \gamma &= a \\ 2\alpha + \beta &= b \\ \alpha + 2\beta + \gamma &= c; \end{aligned}$$

that is,

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

We know that this has a unique solution if and only if the matrix

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

is invertible (i.e., has non-zero determinant). We calculate that

$$\det A = 1 - 2 + (4 - 1) = 2 \neq 0,$$

so there is indeed a unique solution to the Equation (3.3).

We conclude that $\mathscr{C}$ is a linearly independent spanning set, and is a basis for $\mathcal{P}_2$.

We consider the basis of the row-space of a matrix.

Theorem 3.28 *The non-zero rows in an echelon matrix E form a basis for the row-space of E, and the row-space of every matrix row-equivalent to E.*

PROOF: The case when $E = \mathbf{0}$ is trivial (the set of non-zero rows is the empty set). For $E \neq \mathbf{0}$, the non-zero rows of E span the row-space of E, and Theorem 3.16 tells us that this set is LI. Hence this set is a basis for the row-space of E and hence, by Theorem 3.10, for the row-space of every matrix row-equivalent to E. $\square$

For example, in Example 3.18, the matrix $A = \begin{pmatrix} 2 & 0 & -2 & 4 \\ 1 & 3 & 2 & -3 \\ 5 & 6 & 1 & 0 \end{pmatrix}$ is shown to be equivalent to the echelon matrix $E = \begin{pmatrix} 1 & 3 & 2 & -3 \\ 0 & 1 & 1 & -\frac{5}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$, from which we can deduce that the row-space of A is spanned by the two vectors $(1, 3, 2, -3), (0, 1, 1, -\frac{5}{3})$.

It is clear that, for a given matrix, a basis for its row-space is not unique, as the choice of echelon form is not unique.

The key fact about bases for vector spaces is that they always exist. In fact, given any spanning set for a vector space, a basis can be produced from it by omitting the correct choice of vectors; and any linearly independent set can be extended to a basis for the vector space.

Theorem 3.29 *Let V be a vector space and let $\mathscr{A} = \{v_1, v_2, \ldots, v_k\}$ be a set of vectors in V.*

(i) If $\mathscr{A}$ spans V, then there is some subset of $\mathscr{A}$ which is a basis for V.

(ii) If $\mathscr{A}$ is linearly independent, then we can adjoin some number of vectors to $\mathscr{A}$ to produce a set $\mathscr{B}$ which is a basis for V and contains $\mathscr{A}$.

PROOF: (Proof sketch; more details in MT3501) For (i), if $\mathscr{A}$ is not linearly independent, then some vector v_i in $\mathscr{A}$ is a linear combination of the other vectors. If we omit v_i then we show that the resulting set also spans V. Repeating this process produces a set which is linearly independent and spans V; that is, a basis for V. For (ii), if $\mathscr{A}$ does not span V, then there is some vector which is not a linear combination of the vectors in $\mathscr{A}$. Adjoining this vector will still give a linearly independent set. Repeating this process produces a basis when V is finite-dimensional. $\square$

Dimension

Definition 3.30 Let V be a vector space over a field F. 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 $\dim V$.

For the definition of dimension to make sense, we need to know that the dimension is uniquely determined by the vector space; i.e., that it is not possible for a vector space to have bases of different sizes. We state this as a theorem, though we defer the proof to MT3501.

Theorem 3.31 *Every basis for a finite-dimensional vector space contains the same number of vectors.*

Example 3.32 *The dimension of $\mathbb{R}^3$ as a vector space over $\mathbb{R}$ is 3. Indeed, in Example 3.21 we observed that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a basis for $\mathbb{R}^3$.*

Example 3.33 *Example 3.23 tells us that $M_{2\times 2}(F)$ is a vector space of dimension 4 over the field F.*

Example 3.34 *Example 3.25 tells us that $\mathcal{P}_2$ is a vector space of dimension* 3 *over* $\mathbb{R}$.

The following describes an important convention:

Example 3.35 *Let V be the vector space consisting of just the zero vector. What is its dimension and basis?*

Solution: We adopt the convention that it has dimension 0. (This is not something you should attempt to understand intuitively; it is a decision that has been made as it makes the algebra work.) What do we take as its basis? We may want to take the zero vector itself, since this is a spanning set. But this will not work, as it is not linearly independent, and would give dimension 1. The convention is to take the empty set, which gives a dimension of 0 and is vacuously LI. It is not clear, from our definition of spanning set, that this is a valid spanning set for V. However, it turns out that you can make a more general definition of spanning set, which makes this ok, though you do not need to worry about this here. (For those who are interested, the definition is that $U = Span(A)$ if U is the smallest subspace containing A).

Example 3.36 *Consider the differential equation*

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - y = 0.$$

Solution: The general solution to this equation (see for example MT1002) is

$$y = A\mathrm{e}^x + B\mathrm{e}^{-x}$$

where A and B are any constants. We can verify that the set of all functions which are solutions to the differential equation forms a vector space over $\mathbb{R}$. Observe that every solution can be written uniquely as a linear combination of the two functions $f_1(x) = \mathrm{e}^x$ and $f_2(x) = \mathrm{e}^{-x}$, and every such linear combination will be a solution. Hence these two functions form a basis for the vector space of solutions and we conclude that the solution space has dimension 2.

This is a general phenomenon of solutions to linear differential equations. A linear differential equation of degree n will typically have a solution space of dimension n.

We can now state an extremely useful result.

Theorem 3.37 *Let V be a finite-dimensional vector space over a field F. Let $A = \{v_1, \dots, v_k\}$ be a finite set of vectors in V. Then, if A satisfies any* two *of the following three conditions, it is a basis of V:*

(i) it is a spanning set for V;

(ii) it is linearly independent;

(iii) it contains precisely $\dim V$ *vectors.*

Proof: If A satisfies (i) and (ii), it is a basis by definition.

Now suppose A has properties (i) and (iii). By (i), A spans V, and by Theorem 3.29, A therefore contains a basis of V. But any basis of V must contain $\dim V$ vectors — so A itself must already be a basis for V.

Suppose A has properties (ii) and (iii). By (ii), A is a linearly independent set of vectors in V. If it is not already a basis of V, then by Theorem 3.29, it can be extended to form a basis for V by adding some number of vectors. However, A comprises $\dim V$ vectors, so must already be a basis for V. □

Thus, for example, any set of n linearly independent vectors in $\mathbb{R}^n$ is a basis for $\mathbb{R}^n$.

Example 3.38 *Consider the following set of vectors in* $\mathbb{R}^4$*:*

$$A = \{(1,2,3,4),(0,1,0,0),(0,0,5,6),(0,0,0,1)\}$$

Prove that A is a basis for the vector space $\mathbb{R}^4$ over $\mathbb{R}$.

Solution: Since $\dim \mathbb{R}^4$ is 4 (recall the standard basis $\{e_1, e_2, e_3, e_4\}$), A contains precisely $\dim \mathbb{R}^4$ vectors. Hence by Theorem 3.37, it suffices to verify either that A is a spanning set, or that A is linearly independent. We choose to verify linear independence. The four vectors in A are linearly independent if

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 0 & 5 & 0 \\ 4 & 0 & 6 & 1 \end{pmatrix}$$

has non-zero determinant. Here, the determinant is $5 \neq 0$, and so the four vectors in A are linearly independent. Thus A is a basis for $\mathbb{R}^4$.

We end with a particularly important instance of dimension.

Definition 3.39 Let $A \in M_{n\times n}(F)$.

- The *row-rank* of A is the dimension of its row-space.
- The *column-rank* of A is the dimension of its column-space.

Theorem 3.40 *If the matrix A is row-equivalent to the echelon matrix E, then the row-rank of A is equal to the number of non-zero rows in E.*

Example 3.41 *Find the row-rank of the matrix*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 4 & 7 \end{pmatrix}.$$

Solution: We first apply row operations to reduce A into echelon form:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 4 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & -2 & -2 \end{pmatrix} \qquad r_2 \mapsto r_2 - 2r_1,\ r_3 \mapsto r_3 - 3r_1$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad\qquad r_3 \mapsto r_3 + 2r_2.$$

This matrix in echelon form has three non-zero rows. We conclude that A has row-rank 3.

We have seen in Chapter 1 that row and column operations are not "equivalent". We may ask whether there is any relationship between the row rank and the column rank of a matrix. The (perhaps surprising) answer is that they are the same.

Theorem 3.42 *Let $A \in M_{m\times n}(F)$. Then the row rank of A equals the column rank of A.*

PROOF: (Sketch) The proof proceeds by taking A and converting it, by row operations (which do not change its row or column rank) and then by column operations (which also do not change its row or column rank), into a matrix of the form $\begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix}$ By construction this has row and column rank both equal to the same number, namely p. $\square$

Thus we may talk simply of *the rank* of a matrix.

Example 3.43 *Find the column-rank of the matrix*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 4 & 7 \end{pmatrix}.$$

Solution: Column operations convert A to

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

so the column-rank is also 3, as expected.

Corollary 3.44 *Consider the square matrix $A \in M_{n\times n}(F)$. Then A is invertible if and only if $rank(A) = n$.*