Chapter 7

Linear Dependence

All vector spaces are understood to be over a fixed field, $\mathbb{F}$.

Definition 7.1. The vector, $\mathbf{x}$, is a *linear combination of the vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ if and only if there are $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ with

$$\mathbf{x} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \sum_{i=1}^{n} \lambda_i\mathbf{v}_i.$$

The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are *linearly independent (over* $\mathbb{F}$*)* if and only if the equation

$$\sum_{i=1}^{n} \lambda_i\mathbf{v}_i = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \mathbf{0}_V \quad (\lambda_1, \ldots, \lambda_n \in \mathbb{F})$$

has only the trivial solution $\lambda_1 = \cdots \lambda_n = 0$. Otherwise $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are said to be *linearly dependent*.

Example 7.2. Let $V = \mathbb{R}^2$, with its usual vector space structure over $\mathbb{R}$.
Then $(3, 2)$ is a linear combination of $(1, 1)$ and $(5, 4)$, because

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

The vectors $(1, 1), (5, 4)$ and $(3, 2)$ are linearly dependent, because

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

The vectors $(1, 1)$ and $(3, 2)$ are linearly independent, because

$$\lambda(1, 1) + \mu(3, 2) = (0, 0)$$

if and only if

$$\begin{array}{ccccc} \lambda & + & 3\mu & = & 0 \\ \lambda & + & 2\mu & = & 0 \end{array}$$

which, clearly, is the case if and only if $\lambda = \mu = 0$.

Lemma 7.3. *The vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *are linearly dependent if and only if at least one of them can be expressed as a linear combination of the others.*

Proof. Since $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly dependent, there are $\lambda_1, \dots, \lambda_n \in \mathbb{F}$, not all 0, with

$$\sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \mathbf{0}_V.$$

Suppose that $\lambda_i \neq 0$. Then

$$\mathbf{v}_i = \sum_{j \neq i} \mu_j \mathbf{v}_j,$$

where $\mu_j := \frac{-\lambda_j}{\lambda_i} \in \mathbb{F}, \ (j = 1, \dots, n, \ j \neq i)$. □

Theorem 7.4. *If each* $\mathbf{w}_j \quad (j = 1, \dots, m)$ *is a linear combination of* $\mathbf{v}_1, \dots, \mathbf{v}_n$*, then any linear combination of* $\mathbf{w}_1, \dots, \mathbf{w}_m$ *is a linear combination of* $\mathbf{v}_1, \dots, \mathbf{v}_n$*.*

Proof. Suppose that $\mathbf{w}_j = a_{j1}\mathbf{v}_1 + \cdots a_{jn}\mathbf{v}_n \quad (j = 1, \dots, m)$ and that $\mathbf{u} = \lambda_1\mathbf{w}_1 + \cdots + \lambda_m\mathbf{w}_m$. Then

$$\begin{aligned}
\mathbf{u} &= \lambda_1\mathbf{w}_1 + \cdots + \lambda_m\mathbf{w}_m \\
&= \lambda_1(a_{11}\mathbf{v}_1 + \cdots a_{1n}\mathbf{v}_n) + \lambda_2(a_{21}\mathbf{v}_1 + \cdots a_{2n}\mathbf{v}_n) + \ldots + \lambda_m(a_{m1}\mathbf{v}_1 + \cdots a_{mn}\mathbf{v}_n) \\
&= (\lambda_1 a_{11} + \cdots + \lambda_m a_{m1})\mathbf{v}_1 + \ldots + (\lambda_1 a_{1n} + \cdots + \lambda_m a_{mn})\mathbf{v}_n \\
&= \mu_1\mathbf{v}_1 + \cdots + \mu_n\mathbf{v}_n
\end{aligned}$$

where $\mu_i = \lambda_1 a_{1i} + \cdots + \lambda_m a_{mi} \quad (i = 1, \dots, n)$. □

Recall that $\langle S \rangle$, the vector subspace of V generated by the subset, S, of V, is the smallest vector subspace of V containing S. It is the intersection of all those vector subspaces of V which have S as a subset.. We can now provide an alternative and intrinsic description of it.

Corollary 7.5. *Given* $S \subseteq V$*,* $\langle S \rangle$ *comprises all linear combinations of elements of* S*.*

Proof. Let $\mathcal{L}(S)$ denote the set of all linear combinations of elements of S.

$$\mathcal{L}(S) := \big\{ \sum_{j=1}^{n} \lambda_j \mathbf{v}_j \mid \lambda_j \in \mathbb{F}, \mathbf{v}_j \in S \text{ for all } j \leq n, n \in \mathbb{N} \big\}$$

Plainly, $\mathcal{L}(S)$ is a vector subspace of V and $S \subseteq \mathcal{L}(S)$.

Hence, $\langle S \rangle \leq \mathcal{L}(S)$.

For the reverse inclusion, note that every vector space is closed under forming linear combinations of its elements.Thus, if $U \leq V$ and $S \subseteq U$, then $\mathcal{L}(S) \leq U$.

Taking $U = \langle S \rangle$ completes the proof. □

We may thus reformulate the condition for the subset U of the vector space V to be a vector subspace as:

$U \subseteq V$ is a vector subspace of V if and only if U is closed under linear combinations.

Lemma 7.6. *Any non-zero vector is linearly independent.*

Proof. If $\mathbf{v} \neq \mathbf{0}_V$, then by Theorem 3.18 (e), $\lambda\mathbf{v} = \mathbf{0}_V$ if and only if $\lambda = 0$. □

Theorem 7.7. *Suppose that* $\mathbf{v}_1, \dots, \mathbf{v}_n$ *are linearly independent, but* $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}$ *are linearly dependent. Then* $\mathbf{w}$ *is a linear combination of* $\mathbf{v}_1, \dots, \mathbf{v}_n$*.*

Proof. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}$ are linearly dependent, there are $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{F}$ such that

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n + \lambda_{n+1}\mathbf{w} = \mathbf{0}_V,$$

with not all $\lambda_j = 0$.
If $\lambda_{n+1} = 0$, then our equation reduces to $\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \mathbf{0}_V$.
Since$\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, this has only the trivial solution $\lambda_1 = \cdots = \lambda_n = 0$.
Thus we may assume that $\lambda_{n+1} \neq 0$. Then

$$\mathbf{w} = \mu_1\mathbf{v}_1 + \cdots + \mu_n\mathbf{v}_n,$$

where $\mu_j = -\frac{\lambda_j}{\lambda_{n+1}}$ for $j = 1, \ldots, n$. □

Example 7.8. We saw in Example 7.2 that $(1, 1)$ and $(3, 2)$ are linearly independent, but $(1, 1), (3, 2)$ and $(5, 4)$ are linearly dependent in $\mathbb{R}^2$. Plainly

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

Theorem 7.9. *The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent if and only if any vector which can be written as a linear combination of them can be written in this form in precisely one way.*

Proof.

$$\mathbf{x} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \mu_1\mathbf{v}_1 + \cdots + \mu_n\mathbf{v}_n$$

if and only if

$$(\lambda_1 - \mu_1)\mathbf{v}_1 + \cdots + (\lambda_n - \mu_n)\mathbf{v}_n = \mathbf{0}_V,$$

the expression of each vector $\mathbf{x}$ as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is unique if and only if

$$\gamma_1\mathbf{v}_1 + \cdots + \gamma_n\mathbf{v}_n = \mathbf{0}_V$$

has only the trivial solution $\gamma_1 = \cdots = \gamma_n = 0$. □

Linear independence and the property of generating a vector space are closely linked to significant properties of linear transformations, as the next theorem shows.

Theorem 7.10. *Let $T : V \to W$ be a linear transformation. Then*

(i) T is injective if and only if $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ are linearly independent in W whenever $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent in V.

(ii) T is surjective if and only if $T(S)$ generates W whenever S generates V.

Proof. (i) Suppose that T is injective and that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent in V. Then

$$\begin{aligned}
\sum_{j=1}^{n} \lambda_j T(\mathbf{v}_j) = \mathbf{0}_W \quad & \text{if and only if} \quad T(\sum_{j=1}^{n} \lambda_j \mathbf{v}_j) = \mathbf{0}_W, && \text{as } T \text{ is a linear transformation.} \\
& \text{if and only if} \quad \sum_{j=1}^{n} \lambda_j \mathbf{v}_j \mathbf{v}_n \in \ker(T) && \\
& \text{if and only if} \quad \sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \mathbf{0}_V, && \text{as } T \text{ is injective} \\
& \text{if and only if} \quad \text{each } \lambda_j = 0, && \text{as } \mathbf{v}_1, \ldots \mathbf{v}_n \text{ are linearly independent,}
\end{aligned}$$

showing that $T(\mathbf{v}_1), \ldots T(\mathbf{v}_n)$ are linearly independent in W.

Suppose, now, that $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ are linearly independent in W whenever $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent in V.

Take $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}_V \in V$.

Then $\mathbf{v}$ is linearly independent in V.

By hypothesis, $T(\mathbf{v})$ is then linearly independent in W.

Thus $T(\mathbf{v}) \neq \mathbf{0}_W$, whence $\mathbf{v} \notin \ker(T)$.

Hence $\ker(T) = \{\mathbf{0}_V\}$, so that T is injective.

(ii) Suppose that T is surjective and that S generates V.

Take $\mathbf{w} \in W$.

Since T is surjective, there is a $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$.

Since S generates V, there are $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ with

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n.$$

Since T is a linear transformation,

$$\mathbf{w} = T(\mathbf{v}) = T(\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n) = \lambda_1 T(\mathbf{v}_1) + \cdots + \lambda_n T(\mathbf{v}_n),$$

showing that $T(S)$ generate W

Conversely, suppose that $T(S)$ generate W whenever S generates V. Since clearly V generates V, $T(V)$ must generate W,

But $T(V) = \operatorname{im}(T)$ is a vector subspace of W. Hence $W = T(V)$. □

Sets of vectors that are both linearly independent and generate a given vector space lie at the heart of working with vector spaces, since their behaviour completely determines the behaviour of the entire vector space. Such a set of vectors comprises a *basis* for the vector space in question.

Definition 7.11. Let V be a vector space over the field $\mathbb{F}$. The vectors $\{\mathbf{e}_\lambda \mid \lambda \in \Lambda\}$ form a *basis for* V if and only if they are linearly independent and generate V.

Example 7.12. By the definition of the standard real vector space structure on $\mathbb{R}^2$, $\{(1,0),(0,1)\}$ is a basis for $\mathbb{R}^2$. This is the *standard basis* for $\mathbb{R}^2$.

It is left as an exercise for the reader to verify that $\{(0,-1)(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})\}$ is also a basis for $\mathbb{R}^2$.

Example 7.13. The set of all solutions of the real differential equation

$$\frac{d^2y}{dx^2} = -y$$

forms a real vector space, $V = \{f : \mathbb{R} \to \mathbb{R} \mid \frac{d^2f}{dx^2} + f = 0\}$. By the theory of linear differential equations with constant coefficients (cf. MATH102), each $f \in V$ can be expressed uniquely as

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto A\cos x + B\sin x$$

Thus, the functions

$$\begin{aligned} \cos &: \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos x \\ \sin &: \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \sin x \end{aligned}$$

form a basis for V.

It is left as an exercise for the reader to verify that the functions

$$\mathbf{e}_1 : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos(x + \tfrac{\pi}{4})$$
$$\mathbf{e}_2 : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos(x + \tfrac{\pi}{2})$$

also form a basis for V.

The significance of bases is illustrated by the facts that any two bases for a given vector space must have the same number of elements and that two vector spaces over a given field are isomorphic if and only if a basis for one can be found with the same number of elements as a basis for the other. (This is proved in the next chapter for the case of finitely generated vector spaces.)

It is therefore crucial to know when a vector space admits a basis. The answer is provided by the next theorem.

Theorem 7.14. *Every finitely generated vector space admits a basis.*

Proof. Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a finite generating set for V. We may assume, without loss of generality, that $\mathbf{v}_j \neq \mathbf{0}_V \quad (j = 1, \dots n)$.

We construct inductively a basis from S by omitting successively those elements of S which are linearly dependent on their predecessors in the ordering induced by their subscripts.

By Lemma 7.6, $\mathbf{v}_1$ is linearly independent. So put $\mathbf{e}_1 := \mathbf{v}_1$. Clearly $\langle \mathbf{e}_1 \rangle = \langle \mathbf{v}_1 \rangle$.

Now suppose that for $j \geq 1$ we have chosen $\mathbf{e}_1, \dots \mathbf{e}_j$ from S in such a manner that

(i) $\mathbf{e}_1, \dots \mathbf{e}_j$ are linearly independent and

(ii) $\langle \mathbf{e}_1, \dots \mathbf{e}_j \rangle = \langle \mathbf{v}_1, \dots \mathbf{v}_{n_j} \rangle$.

If $\langle \mathbf{e}_1, \dots \mathbf{e}_j \rangle = V$, we are finished.

Otherwise, let n_{j+1} be the least integer such that $\mathbf{e}_1, \dots \mathbf{e}_j, \mathbf{v}_{n_{j+1}}$ are linearly independent.

Put $\mathbf{e}_{j+1} := \mathbf{v}_{n_{j+1}}$.

Then $\mathbf{e}_1, \dots \mathbf{e}_{j+1}$ are obviously linearly independent and $\langle \mathbf{e}_1, \dots \mathbf{e}_{j+1} \rangle = \langle \mathbf{v}_1, \dots \mathbf{v}_{n_{j+1}} \rangle$.

Since S is finite, this procedure must terminate after at most n steps. $\square$

Observation 7.15. The last theorem is still true without the restriction to finite generation, but, of course, our proof would not suffice then. The more general statement requires the *Axiom of Choice* or some equivalent of it. This would take us into the realm of formal set theory, which is not within the scope of this course.

Set theory lies at the basis of most of mathematics and it was through axiomatic set theory that the theory of recursive functions, the theory of computability and the theory of Turing machines arose. Thus, in addition to its centrality in the development of modern mathematics, set theory is the historical, conceptual and theoretical parent of modern computing and modern computers.

7.1 Exercises

Exercise 7.1. Let $T : V \to W$ be a linear transformation of $\mathbb{F}$ vector spaces.

Show that if $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly dependent, then so are $T(\mathbf{v}_1), \dots, T(\mathbf{v}_k)$.

Find an example with $\mathbf{v}_1, \dots, \mathbf{v}_k$ linearly independent and $T(\mathbf{v}_1), \dots, T(\mathbf{v}_k)$ linearly dependent.

Exercise 7.2. Find a basis for the vector subspace of $\mathbb{R}^4$ generated by

$$\{(1,1,0,0), (0,0,0,0), (1,0,0,1), (0,1,0,1), (1,-1,0,3)\}$$

Exercise 7.3. Let $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for the vector space V over the field $\mathbb{F}$.

Show that V is isomorphic with $\mathcal{F}(\mathcal{B})$, the vector space of all functions $f \colon \mathcal{B} \to \mathbb{F}$.

Exercise 7.4. Show that $\{(0, -1), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})\}$ is a basis for $\mathbb{R}^2$ with its standard vector space structure.

Exercise 7.5. Show that the functions

$$\begin{aligned} \mathbf{e}_1 &\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos(x + \tfrac{\pi}{4}) \\ \mathbf{e}_2 &\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto \cos(x + \tfrac{\pi}{2}) \end{aligned}$$

form a basis for $V = \{f \colon \mathbb{R} \to \mathbb{R} \mid \frac{d^2 f}{dx^2} + f = 0\}$.

Exercise 7.6. Let $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ be a basis for the vector space V over the field $\mathbb{F}$.

Prove that for every vector space, W, over $\mathbb{F}$ and every function, $f : \mathcal{B} \to W$, there is a unique linear transformation $T : V \to W$ with $T(\mathbf{e}_j) = \varphi(\mathbf{e}_j)$ for all $j \in \{1, \ldots, m\}$.

This is the *universal property* of a basis. It is customary to express it in the form of the following *commutative diagram.*

$$\begin{tikzcd}
V \arrow[r, dashed, "\exists!\, T"] & W \\
\mathcal{B} \arrow[u, "i_{\mathcal{B}}^{V}"] \arrow[ur, "f"'] &
\end{tikzcd}$$