LINEAR ALGEBRA

W W L CHEN

Chapter 5

INTRODUCTION TO VECTOR SPACES

5.1. Real Vector Spaces

Before we give any formal definition of a vector space, we shall consider a few concrete examples of such an abstract object. We first study two examples from the theory of vectors which we first discussed in Chapter 4.

EXAMPLE 5.1.1. Consider the set $\mathbb{R}^2$ of all vectors of the form $\mathbf{u} = (u_1, u_2)$, where $u_1, u_2 \in \mathbb{R}$. Consider vector addition and also multiplication of vectors by real numbers. It is easy to check that we have the following properties:

(1.1) For every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, we have $\mathbf{u} + \mathbf{v} \in \mathbb{R}^2$.
(1.2) For every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^2$, we have $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
(1.3) For every $\mathbf{u} \in \mathbb{R}^2$, we have $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$.
(1.4) For every $\mathbf{u} \in \mathbb{R}^2$, we have $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
(1.5) For every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
(2.1) For every $c \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^2$, we have $c\mathbf{u} \in \mathbb{R}^2$.
(2.2) For every $c \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, we have $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
(2.3) For every $a, b \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^2$, we have $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.
(2.4) For every $a, b \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^2$, we have $(ab)\mathbf{u} = a(b\mathbf{u})$.
(2.5) For every $\mathbf{u} \in \mathbb{R}^2$, we have $1\mathbf{u} = \mathbf{u}$.

EXAMPLE 5.1.2. Consider the set $\mathbb{R}^3$ of all vectors of the form $\mathbf{u} = (u_1, u_2, u_3)$, where $u_1, u_2, u_3 \in \mathbb{R}$. Consider vector addition and also multiplication of vectors by real numbers. It is easy to check that we have properties analogous to (1.1)–(1.5) and (2.1)–(2.5) in the previous example, with reference to $\mathbb{R}^2$ being replaced by $\mathbb{R}^3$.

We next turn to an example from the theory of matrices which we first discussed in Chapter 2.

Example 5.1.3. Consider the set $\mathcal{M}_{2,2}(\mathbb{R})$ of all 2×2 matrices with entries in $\mathbb{R}$. Consider matrix addition and also multiplication of matrices by real numbers. Denote by O the 2×2 null matrix. It is easy to check that we have the following properties:

(1.1) For every $P, Q \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $P + Q \in \mathcal{M}_{2,2}(\mathbb{R})$.
(1.2) For every $P, Q, R \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $P + (Q + R) = (P + Q) + R$.
(1.3) For every $P \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $P + O = O + P = P$.
(1.4) For every $P \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $P + (-P) = O$.
(1.5) For every $P, Q \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $P + Q = Q + P$.
(2.1) For every $c \in \mathbb{R}$ and $P \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $cP \in \mathcal{M}_{2,2}(\mathbb{R})$.
(2.2) For every $c \in \mathbb{R}$ and $P, Q \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $c(P + Q) = cP + cQ$.
(2.3) For every $a, b \in \mathbb{R}$ and $P \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $(a + b)P = aP + bP$.
(2.4) For every $a, b \in \mathbb{R}$ and $P \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $(ab)P = a(bP)$.
(2.5) For every $P \in \mathcal{M}_{2,2}(\mathbb{R})$, we have $1P = P$.

We also turn to an example from the theory of functions.

Example 5.1.4. Consider the set $\mathcal{A}$ of all functions of the form $f : \mathbb{R} \to \mathbb{R}$. For any two functions $f, g \in \mathcal{A}$, define the function $f + g : \mathbb{R} \to \mathbb{R}$ by writing $(f + g)(x) = f(x) + g(x)$ for every $x \in \mathbb{R}$. For every function $f \in \mathcal{A}$ and every number $c \in \mathbb{R}$, define the function $cf : \mathbb{R} \to \mathbb{R}$ by writing $(cf)(x) = cf(x)$ for every $x \in \mathbb{R}$. Denote by $\lambda : \mathbb{R} \to \mathbb{R}$ the function where $\lambda(x) = 0$ for every $x \in \mathbb{R}$. Then it is easy to check that we have the following properties:

(1.1) For every $f, g \in \mathcal{A}$, we have $f + g \in \mathcal{A}$.
(1.2) For every $f, g, h \in \mathcal{A}$, we have $f + (g + h) = (f + g) + h$.
(1.3) For every $f \in \mathcal{A}$, we have $f + \lambda = \lambda + f = f$.
(1.4) For every $f \in \mathcal{A}$, we have $f + (-f) = \lambda$.
(1.5) For every $f, g \in \mathcal{A}$, we have $f + g = g + f$.
(2.1) For every $c \in \mathbb{R}$ and $f \in \mathcal{A}$, we have $cf \in \mathcal{A}$.
(2.2) For every $c \in \mathbb{R}$ and $f, g \in \mathcal{A}$, we have $c(f + g) = cf + cg$.
(2.3) For every $a, b \in \mathbb{R}$ and $f \in \mathcal{A}$, we have $(a + b)f = af + bf$.
(2.4) For every $a, b \in \mathbb{R}$ and $f \in \mathcal{A}$, we have $(ab)f = a(bf)$.
(2.5) For every $f \in \mathcal{A}$, we have $1f = f$.

There are many more examples of sets where properties analogous to (1.1)–(1.5) and (2.1)–(2.5) in the four examples above hold. This apparent similarity leads us to consider an abstract object which will incorporate all these individual cases as examples. We say that these examples are all vector spaces over $\mathbb{R}$.

Definition. A vector space V over $\mathbb{R}$, or a real vector space V, is a set of objects, known as vectors, together with vector addition $+$ and multiplication of vectors by element of $\mathbb{R}$, and satisfying the following properties:

(VA1) For every $\mathbf{u}, \mathbf{v} \in V$, we have $\mathbf{u} + \mathbf{v} \in V$.
(VA2) For every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, we have $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
(VA3) There exists an element $\mathbf{0} \in V$ such that for every $\mathbf{u} \in V$, we have $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$.
(VA4) For every $\mathbf{u} \in V$, there exists $-\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
(VA5) For every $\mathbf{u}, \mathbf{v} \in V$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
(SM1) For every $c \in \mathbb{R}$ and $\mathbf{u} \in V$, we have $c\mathbf{u} \in V$.
(SM2) For every $c \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$, we have $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
(SM3) For every $a, b \in \mathbb{R}$ and $\mathbf{u} \in V$, we have $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.
(SM4) For every $a, b \in \mathbb{R}$ and $\mathbf{u} \in V$, we have $(ab)\mathbf{u} = a(b\mathbf{u})$.
(SM5) For every $\mathbf{u} \in V$, we have $1\mathbf{u} = \mathbf{u}$.

Remark. The elements $a, b, c \in \mathbb{R}$ discussed in (SM1)–(SM5) are known as scalars. Multiplication of vectors by elements of $\mathbb{R}$ is sometimes known as scalar multiplication.

EXAMPLE 5.1.5. Let $n \in \mathbb{N}$. Consider the set $\mathbb{R}^n$ of all vectors of the form $\mathbf{u} = (u_1, \ldots, u_n)$, where $u_1, \ldots, u_n \in \mathbb{R}$. For any two vectors $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ and any number $c \in \mathbb{R}$, write

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \ldots, u_n + v_n) \qquad \text{and} \qquad c\mathbf{u} = (cu_1, \ldots, cu_n).$$

To check (VA1), simply note that $u_1+v_1, \ldots, u_n+v_n \in \mathbb{R}$. To check (VA2), note that if $\mathbf{w} = (w_1, \ldots, w_n)$, then

$$\begin{aligned}\mathbf{u} + (\mathbf{v} + \mathbf{w}) &= (u_1, \ldots, u_n) + (v_1 + w_1, \ldots, v_n + w_n) = (u_1 + (v_1 + w_1), \ldots, u_n + (v_n + w_n)) \\ &= ((u_1 + v_1) + w_1, \ldots, (u_n + v_n) + w_n) = (u_1 + v_1, \ldots, u_n + v_n) + (w_1, \ldots, w_n) \\ &= (\mathbf{u} + \mathbf{v}) + \mathbf{w}.\end{aligned}$$

If we take $\mathbf{0}$ to be the zero vector $(0, \ldots, 0)$, then $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$, giving (VA3). Next, writing $-\mathbf{u} = (-u_1, \ldots, -u_n)$, we have $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$, giving (VA4). To check (VA5), note that

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \ldots, u_n + v_n) = (v_1 + u_1, \ldots, v_n + u_n) = \mathbf{v} + \mathbf{u}.$$

To check (SM1), simply note that $cu_1, \ldots, cu_n \in \mathbb{R}$. To check (SM2), note that

$$\begin{aligned}c(\mathbf{u} + \mathbf{v}) &= c(u_1 + v_1, \ldots, u_n + v_n) = (c(u_1 + v_1), \ldots, c(u_n + v_n)) \\ &= (cu_1 + cv_1, \ldots, cu_n + cv_n) = (cu_1, \ldots, cu_n) + (cv_1, \ldots, cv_n) = c\mathbf{u} + c\mathbf{v}.\end{aligned}$$

To check (SM3), note that

$$\begin{aligned}(a + b)\mathbf{u} &= ((a + b)u_1, \ldots, (a + b)u_n) = (au_1 + bu_1, \ldots, au_n + bu_n) \\ &= (au_1, \ldots, au_n) + (bu_1, \ldots, bu_n) = a\mathbf{u} + b\mathbf{u}.\end{aligned}$$

To check (SM4), note that

$$(ab)\mathbf{u} = ((ab)u_1, \ldots, (ab)u_n) = (a(bu_1), \ldots, a(bu_n)) = a(bu_1, \ldots, bu_n) = a(b\mathbf{u}).$$

Finally, to check (SM5), note that

$$1\mathbf{u} = (1u_1, \ldots, 1u_n) = (u_1, \ldots, u_n) = \mathbf{u}.$$

It follows that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$. This is known as the n-dimensional euclidean space.

EXAMPLE 5.1.6. Let $k \in \mathbb{N}$. Consider the set P_k of all polynomials of the form

$$p(x) = p_0 + p_1 x + \ldots + p_k x^k, \quad \text{where } p_0, p_1, \ldots, p_k \in \mathbb{R}.$$

In other words, P_k is the set of all polynomials of degree at most k and with coefficients in $\mathbb{R}$. For any two polynomials $p(x) = p_0 + p_1 x + \ldots + p_k x^k$ and $q(x) = q_0 + q_1 x + \ldots + q_k x^k$ in P_k and for any number $c \in \mathbb{R}$, write

$$p(x) + q(x) = (p_0 + q_0) + (p_1 + q_1)x + \ldots + (p_k + q_k)x^k \qquad \text{and} \qquad cp(x) = cp_0 + cp_1 x + \ldots + cp_k x^k.$$

To check (VA1), simply note that $p_0 + q_0, \ldots, p_k + q_k \in \mathbb{R}$. To check (VA2), note that if we write $r(x) = r_0 + r_1 x + \ldots + r_k x^k$, then we have

$$\begin{aligned}p(x) + (q(x) + r(x)) &= (p_0 + p_1 x + \ldots + p_k x^k) + ((q_0 + r_0) + (q_1 + r_1)x + \ldots + (q_k + r_k)x^k) \\ &= (p_0 + (q_0 + r_0)) + (p_1 + (q_1 + r_1))x + \ldots + (p_k + (q_k + r_k))x^k \\ &= ((p_0 + q_0) + r_0) + ((p_1 + q_1) + r_1)x + \ldots + ((p_k + q_k) + r_k)x^k \\ &= ((p_0 + q_0) + (p_1 + q_1)x + \ldots + (p_k + q_k)x^k) + (r_0 + r_1 x + \ldots + r_k x^k) \\ &= (p(x) + q(x)) + r(x).\end{aligned}$$

If we take $\mathbf{0}$ to be the zero polynomial $0+0x+\ldots+0x^k$, then $p(x)+\mathbf{0}=\mathbf{0}+p(x)=p(x)$, giving (VA3). Next, writing $-p(x)=-p_0-p_1x-\ldots-p_kx^k$, we have $p(x)+(-p(x))=\mathbf{0}$, giving (VA4). To check (VA5), note that

$$\begin{aligned}p(x)+q(x)&=(p_0+q_0)+(p_1+q_1)x+\ldots+(p_k+q_k)x^k\\&=(q_0+p_0)+(q_1+p_1)x+\ldots+(q_k+p_k)x^k=q(x)+p(x).\end{aligned}$$

To check (SM1), simply note that $cp_0,\ldots,cp_k\in\mathbb{R}$. To check (SM2), note that

$$\begin{aligned}c(p(x)+q(x))&=c((p_0+q_0)+(p_1+q_1)x+\ldots+(p_k+q_k)x^k)\\&=c(p_0+q_0)+c(p_1+q_1)x+\ldots+c(p_k+q_k)x^k\\&=(cp_0+cq_0)+(cp_1+cq_1)x+\ldots+(cp_k+cq_k)x^k\\&=(cp_0+cp_1x+\ldots+cp_kx^k)+(cq_0+cq_1x+\ldots+cq_kx^k)\\&=cp(x)+cq(x).\end{aligned}$$

To check (SM3), note that

$$\begin{aligned}(a+b)p(x)&=(a+b)p_0+(a+b)p_1x+\ldots+(a+b)p_kx^k\\&=(ap_0+bp_0)+(ap_1+bp_1)x+\ldots+(ap_k+bp_k)x^k\\&=(ap_0+ap_1x+\ldots+ap_kx^k)+(bp_0+bp_1x+\ldots+bp_kx^k)\\&=ap(x)+bp(x).\end{aligned}$$

To check (SM4), note that

$$\begin{aligned}(ab)p(x)&=(ab)p_0+(ab)p_1x+\ldots+(ab)p_kx^k=a(bp_0)+a(bp_1)x+\ldots+a(bp_k)x^k\\&=a(bp_0+bp_1x+\ldots+bp_kx^k)=a(bp(x)).\end{aligned}$$

Finally, to check (SM5), note that

$$1p(x)=1p_0+1p_1x+\ldots+1p_kx^k=p_0+p_1x+\ldots+p_kx^k=p(x).$$

It follows that P_k is a vector space over $\mathbb{R}$. Note also that the vectors are the polynomials.

There are a few simple properties of vector spaces that we can deduce easily from the definition.

PROPOSITION 5A. *Suppose that V is a vector space over $\mathbb{R}$, and that $\mathbf{u}\in V$ and $c\in\mathbb{R}$.*
(a) We have $0\mathbf{u}=\mathbf{0}$.
(b) We have $c\mathbf{0}=\mathbf{0}$.
(c) We have $(-1)\mathbf{u}=-\mathbf{u}$.
(d) If $c\mathbf{u}=\mathbf{0}$, then $c=0$ or $\mathbf{u}=\mathbf{0}$.

Proof. (a) By (SM1), we have $0\mathbf{u}\in V$. Hence

$$\begin{aligned}0\mathbf{u}+0\mathbf{u}&=(0+0)\mathbf{u}&&\text{(by (SM3))},\\&=0\mathbf{u}&&\text{(since }0\in\mathbb{R}).\end{aligned}$$

It follows that

$$\begin{aligned}0\mathbf{u}&=0\mathbf{u}+\mathbf{0}&&\text{(by (VA3))},\\&=0\mathbf{u}+(0\mathbf{u}+(-(0\mathbf{u})))&&\text{(by (VA4))},\\&=(0\mathbf{u}+0\mathbf{u})+(-(0\mathbf{u}))&&\text{(by (VA2))},\\&=0\mathbf{u}+(-(0\mathbf{u}))&&\text{(from above)},\\&=\mathbf{0}&&\text{(by (VA4))}.\end{aligned}$$

(b) By (SM1), we have $c\mathbf{0} \in V$. Hence

$$\begin{aligned} c\mathbf{0} + c\mathbf{0} &= c(\mathbf{0} + \mathbf{0}) && \text{(by (SM2))}, \\ &= c\mathbf{0} && \text{(by (VA3))}. \end{aligned}$$

It follows that

$$\begin{aligned} c\mathbf{0} &= c\mathbf{0} + \mathbf{0} && \text{(by (VA3))}, \\ &= c\mathbf{0} + (c\mathbf{0} + (-(c\mathbf{0}))) && \text{(by (VA4))}, \\ &= (c\mathbf{0} + c\mathbf{0}) + (-(c\mathbf{0})) && \text{(by (VA2))}, \\ &= c\mathbf{0} + (-(c\mathbf{0})) && \text{(from above)}, \\ &= \mathbf{0} && \text{(by (VA4))}. \end{aligned}$$

(c) We have

$$\begin{aligned} (-1)\mathbf{u} &= (-1)\mathbf{u} + \mathbf{0} && \text{(by (VA3))}, \\ &= (-1)\mathbf{u} + (\mathbf{u} + (-\mathbf{u})) && \text{(by (VA4))}, \\ &= ((-1)\mathbf{u} + \mathbf{u}) + (-\mathbf{u}) && \text{(by (VA2))}, \\ &= ((-1)\mathbf{u} + 1\mathbf{u}) + (-\mathbf{u}) && \text{(by (SM5))}, \\ &= ((-1) + 1)\mathbf{u} + (-\mathbf{u}) && \text{(by (SM3))}, \\ &= 0\mathbf{u} + (-\mathbf{u}) && \text{(since } 1 \in \mathbb{R}), \\ &= \mathbf{0} + (-\mathbf{u}) && \text{(from (a))}, \\ &= -\mathbf{u} && \text{(by (VA3))}. \end{aligned}$$

(d) Suppose that $c\mathbf{u} = \mathbf{0}$ and $c \neq 0$. Then $c^{-1} \in \mathbb{R}$ and

$$\begin{aligned} \mathbf{u} &= 1\mathbf{u} && \text{(by (SM5))}, \\ &= (c^{-1}c)\mathbf{u} && \text{(since } c \in \mathbb{R} \setminus \{0\}), \\ &= c^{-1}(c\mathbf{u}) && \text{(by (SM4))}, \\ &= c^{-1}\mathbf{0} && \text{(assumption)}, \\ &= \mathbf{0} && \text{(from (b))}, \end{aligned}$$

as required. ◯

5.2. Subspaces

Example 5.2.1. Consider the vector space $\mathbb{R}^2$ of all points (x, y), where $x, y \in \mathbb{R}$. Let L be a line through the origin $\mathbf{0} = (0, 0)$. Suppose that L is represented by the equation $\alpha x + \beta y = 0$; in other words,

$$L = \{(x, y) \in \mathbb{R}^2 : \alpha x + \beta y = 0\}.$$

Note first of all that $\mathbf{0} = (0, 0) \in L$, so that (VA3) and (VA4) clearly hold in L. Also (VA2) and (VA5) clearly hold in L. To check (VA1), note that if $(x, y), (u, v) \in L$, then $\alpha x + \beta y = 0$ and $\alpha u + \beta v = 0$, so that $\alpha(x + u) + \beta(y + v) = 0$, whence $(x, y) + (u, v) = (x + u, y + v) \in L$. Next, note that (SM2)–(SM5) clearly hold in L. To check (SM1), note that if $(x, y) \in L$, then $\alpha x + \beta y = 0$, so that $\alpha(cx) + \beta(cy) = 0$, whence $c(x, y) = (cx, cy) \in L$. It follows that L forms a vector space over $\mathbb{R}$. In fact, we have shown that every line in $\mathbb{R}^2$ through the origin is a vector space over $\mathbb{R}$.

DEFINITION. Suppose that V is a vector space over $\mathbb{R}$, and that W is a subset of V. Then we say that W is a subspace of V if W forms a vector space over $\mathbb{R}$ under the vector addition and scalar multiplication defined in V.

EXAMPLE 5.2.2. We have just shown in Example 5.2.1 that every line in $\mathbb{R}^2$ through the origin is a subspace of $\mathbb{R}^2$. On the other hand, if we work through the example again, then it is clear that we have really only checked conditions (VA1) and (SM1) for L, and that $\mathbf{0} = (0,0) \in L$.

PROPOSITION 5B. *Suppose that V is a vector space over $\mathbb{R}$, and that W is a non-empty subset of V. Then W is a subspace of V if the following conditions are satisfied:*
(SP1) For every $\mathbf{u},\mathbf{v} \in W$, we have $\mathbf{u}+\mathbf{v} \in W$.
(SP2) For every $c \in \mathbb{R}$ and $\mathbf{u} \in W$, we have $c\mathbf{u} \in W$.

PROOF. To show that W is a vector space over $\mathbb{R}$, it remains to check that W satisfies (VA2)–(VA5) and (SM2)–(SM5). To check (VA3) and (VA4) for W, it clearly suffices to check that $\mathbf{0} \in W$. Since W is non-empty, there exists $\mathbf{u} \in W$. Then it follows from (SP2) and Proposition 5A(a) that $\mathbf{0} = 0\mathbf{u} \in W$. The remaining conditions (VA2), (VA5) and (SM2)–(SM5) hold for all vectors in V, and hence also for all vectors in W. ◯

EXAMPLE 5.2.3. Consider the vector space $\mathbb{R}^3$ of all points (x,y,z), where $x,y,z \in \mathbb{R}$. Let P be a plane through the origin $\mathbf{0} = (0,0,0)$. Suppose that P is represented by the equation $\alpha x + \beta y + \gamma z = 0$; in other words,

$$P = \{(x,y,z) \in \mathbb{R}^2 : \alpha x + \beta y + \gamma z = 0\}.$$

To check (SP1), note that if $(x,y,z),(u,v,w) \in P$, then $\alpha x + \beta y + \gamma z = 0$ and $\alpha u + \beta v + \gamma w = 0$, so that $\alpha(x+u) + \beta(y+v) + \gamma(z+w) = 0$, whence $(x,y,z) + (u,v,w) = (x+u,y+v,z+w) \in P$. To check (SP2), note that if $(x,y,z) \in P$, then $\alpha x + \beta y + \gamma z = 0$, so that $\alpha(cx) + \beta(cy) + \gamma(cz) = 0$, whence $c(x,y,z) = (cx,cy,cz) \in P$. It follows that P is a subspace of $\mathbb{R}^3$. Next, let L be a line through the origin $\mathbf{0} = (0,0,0)$. Suppose that $(\alpha,\beta,\gamma) \in \mathbb{R}^3$ is a non-zero point on L. Then we can write

$$L = \{t(\alpha,\beta,\gamma) : t \in \mathbb{R}\}.$$

Suppose that $\mathbf{u} = t(\alpha,\beta,\gamma) \in L$ and $\mathbf{v} = s(\alpha,\beta,\gamma) \in L$, and that $c \in \mathbb{R}$. Then

$$\mathbf{u}+\mathbf{v} = t(\alpha,\beta,\gamma) + s(\alpha,\beta,\gamma) = (t+s)(\alpha,\beta,\gamma) \in L,$$

giving (SP1). Also, $c\mathbf{u} = c(t(\alpha,\beta,\gamma)) = (ct)(\alpha,\beta,\gamma) \in L$, giving (SP2). It follows that L is a subspace of $\mathbb{R}^3$. Finally, it is not difficult to see that both $\{\mathbf{0}\}$ and $\mathbb{R}^3$ are subspaces of $\mathbb{R}^3$.

EXAMPLE 5.2.4. Note that $\mathbb{R}^2$ is not a subspace of $\mathbb{R}^3$. First of all, $\mathbb{R}^2$ is not a subset of $\mathbb{R}^3$. Note also that vector addition and scalar multiplication are different in $\mathbb{R}^2$ and $\mathbb{R}^3$.

EXAMPLE 5.2.5. Suppose that A is an $m \times n$ matrix and $\mathbf{0}$ is the $m \times 1$ zero column matrix. Consider the system $A\mathbf{x} = \mathbf{0}$ of m homogeneous linear equations in the n unknowns $x_1, \ldots, x_n$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

is interpreted as an element of the vector space $\mathbb{R}^n$, with usual vector addition and scalar multiplication. Let S denote the set of all solutions of the system. Suppose that $\mathbf{x},\mathbf{y} \in S$ and $c \in \mathbb{R}$. Then

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0}+\mathbf{0} = \mathbf{0},$$

giving (SP1). Also, $A(c\mathbf{x}) = c(A\mathbf{x}) = c\mathbf{0} = \mathbf{0}$, giving (SP2). It follows that S is a subspace of $\mathbb{R}^n$. To summarize, the space of solutions of a system of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

EXAMPLE 5.2.6. As a special case of Example 5.2.5, note that if we take two non-parallel planes in $\mathbb{R}^3$ through the origin $\mathbf{0} = (0,0,0)$, then the intersection of these two planes is clearly a line through the origin. However, each plane is a homogeneous equation in the three unknowns $x,y,z \in \mathbb{R}$. It follows that the intersection of the two planes is the collection of all solutions $(x,y,z) \in \mathbb{R}^3$ of the system formed by the two homogeneous equations in the three unknowns x,y,z representing these two planes. We have already shown in Example 5.2.3 that the line representing all these solutions is a subspace of $\mathbb{R}^3$.

EXAMPLE 5.2.7. We showed in Example 5.1.3 that the set $\mathcal{M}_{2,2}(\mathbb{R})$ of all 2×2 matrices with entries in $\mathbb{R}$ forms a vector space over $\mathbb{R}$. Consider the subset

$$W = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} : a_{11}, a_{12}, a_{21} \in \mathbb{R} \right\}$$

of $\mathcal{M}_{2,2}(\mathbb{R})$. Since

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & 0 \end{pmatrix} \quad \text{and} \quad c\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} \\ ca_{21} & 0 \end{pmatrix},$$

it follows that (SP1) and (SP2) are satisfied. Hence W is a subspace of $\mathcal{M}_{2,2}(\mathbb{R})$.

EXAMPLE 5.2.8. We showed in Example 5.1.4 that the set $\mathcal{A}$ of all functions of the form $f : \mathbb{R} \to \mathbb{R}$ forms a vector space over $\mathbb{R}$. Let $\mathcal{C}_0$ denote the set of all functions of the form $f : \mathbb{R} \to \mathbb{R}$ which are continuous at $x = 2$, and let $\mathcal{C}_1$ denote the set of all functions of the form $f : \mathbb{R} \to \mathbb{R}$ which are differentiable at $x = 2$. Then it follows from the arithmetic of limits and the arithmetic of derivatives that $\mathcal{C}_0$ and $\mathcal{C}_1$ are both subspaces of $\mathcal{A}$. Furthermore, $\mathcal{C}_1$ is a subspace of $\mathcal{C}_0$ (why?). On the other hand, let $k \in \mathbb{N}$. Recall from Example 5.1.6 the vector space P_k of all polynomials of the form

$$p(x) = p_0 + p_1x + \ldots + p_kx^k, \quad \text{where } p_0, p_1, \ldots, p_k \in \mathbb{R}.$$

In other words, P_k is the set of all polynomials of degree at most k and with coefficients in $\mathbb{R}$. Clearly P_k is a subspace of $\mathcal{C}_1$.

5.3. Linear Combination

In this section and the next two, we shall study ways of describing the vectors in a vector space V. Our ultimate goal is to be able to determine a subset B of vectors in V and describe every element of V in terms of elements of B in a unique way. The first step in this direction is summarized below.

DEFINITION. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a vector space V over $\mathbb{R}$. By a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$, we mean an expression of the type

$$c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r,$$

where $c_1, \ldots, c_r \in \mathbb{R}$.

EXAMPLE 5.3.1. In $\mathbb{R}^2$, every vector (x,y) is a linear combination of the two vectors $\mathbf{i} = (1,0)$ and $\mathbf{j} = (0,1)$, for clearly $(x,y) = x\mathbf{i} + y\mathbf{j}$.

EXAMPLE 5.3.2. In $\mathbb{R}^3$, every vector (x,y,z) is a linear combination of the three vectors $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$ and $\mathbf{k} = (0,0,1)$, for clearly $(x,y,z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

Example 5.3.3. In $\mathbb{R}^4$, the vector $(1,4,-2,6)$ is a linear combination of the two vectors $(1,2,0,4)$ and $(1,1,1,3)$, for we have $(1,4,-2,6) = 3(1,2,0,4) - 2(1,1,1,3)$. On the other hand, the vector $(2,6,0,9)$ is not a linear combination of the two vectors $(1,2,0,4)$ and $(1,1,1,3)$, for

$$(2,6,0,9) = c_1(1,2,0,4) + c_2(1,1,1,3)$$

would lead to the system of four equations

$$\begin{aligned} c_1 + c_2 &= 2, \\ 2c_1 + c_2 &= 6, \\ c_2 &= 0, \\ 4c_1 + 3c_2 &= 9. \end{aligned}$$

It is easily checked that this system has no solutions.

Example 5.3.4. In the vector space $\mathcal{A}$ of all functions of the form $f : \mathbb{R} \to \mathbb{R}$ described in Example 5.1.4, the function $\cos 2x$ is a linear combination of the three functions $\cos^2 x$, $\cosh^2 x$ and $\sinh^2 x$. It is not too difficult to check that

$$\cos 2x = 2\cos^2 x + \sinh^2 x - \cosh^2 x,$$

noting that $\cos 2x = 2\cos^2 x - 1$ and $\cosh^2 x - \sinh^2 x = 1$.

We observe that in Example 5.3.1, every vector in $\mathbb{R}^2$ is a linear combination of the two vectors $\mathbf{i}$ and $\mathbf{j}$. Similarly, in Example 5.3.2, every vector in $\mathbb{R}^3$ is a linear combination of the three vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. On the other hand, we observe that in Example 5.3.3, not every vector in $\mathbb{R}^4$ is a linear combination of the two vectors $(1,2,0,4)$ and $(1,1,1,3)$.

Let us therefore investigate the collection of all vectors in a vector space that can be represented as linear combinations of a given set of vectors in V.

Definition. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a vector space V over $\mathbb{R}$. The set

$$\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = \{c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r : c_1, \ldots, c_r \in \mathbb{R}\}$$

is called the span of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$. We also say that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ span V if

$$\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = V;$$

in other words, if every vector in V can be expressed as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$.

Example 5.3.5. The two vectors $\mathbf{i} = (1,0)$ and $\mathbf{j} = (0,1)$ span $\mathbb{R}^2$.

Example 5.3.6. The three vectors $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$ and $\mathbf{k} = (0,0,1)$ span $\mathbb{R}^3$.

Example 5.3.7. The two vectors $(1,2,0,4)$ and $(1,1,1,3)$ do not span $\mathbb{R}^4$.

PROPOSITION 5C. *Suppose that* $\mathbf{v}_1, \ldots, \mathbf{v}_r$ *are vectors in a vector space* V *over* $\mathbb{R}$.
(a) Then $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ *is a subspace of* V.
(b) Suppose further that W *is a subspace of* V *and* $\mathbf{v}_1, \ldots, \mathbf{v}_r \in W$. *Then* $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subseteq W$.

Proof. (a) Suppose that $\mathbf{u}, \mathbf{w} \in \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ and $c \in \mathbb{R}$. There exist $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{R}$ such that

$$\mathbf{u} = a_1\mathbf{v}_1 + \ldots + a_r\mathbf{v}_r \qquad \text{and} \qquad \mathbf{w} = b_1\mathbf{v}_1 + \ldots + b_r\mathbf{v}_r.$$

Then

$$\begin{aligned}\mathbf{u}+\mathbf{w} &= (a_1\mathbf{v}_1+\ldots+a_r\mathbf{v}_r)+(b_1\mathbf{v}_1+\ldots+b_r\mathbf{v}_r)\\ &= (a_1+b_1)\mathbf{v}_1+\ldots+(a_r+b_r)\mathbf{v}_r \in \mathrm{span}\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}\end{aligned}$$

and

$$c\mathbf{u} = c(a_1\mathbf{v}_1+\ldots+a_r\mathbf{v}_r) = (ca_1)\mathbf{v}_1+\ldots+(ca_r)\mathbf{v}_r \in \mathrm{span}\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}.$$

It follows from Proposition 5B that $\mathrm{span}\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$ is a subspace of V.

(b) Suppose that $c_1,\ldots,c_r \in \mathbb{R}$ and $\mathbf{u} = c_1\mathbf{v}_1+\ldots+c_r\mathbf{v}_r \in \mathrm{span}\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$. If $\mathbf{v}_1,\ldots,\mathbf{v}_r \in W$, then it follows from (SM1) for W that $c_1\mathbf{v}_1,\ldots,c_r\mathbf{v}_r \in W$. It then follows from (VA1) for W that $\mathbf{u} = c_1\mathbf{v}_1+\ldots+c_r\mathbf{v}_r \in W$. ◯

EXAMPLE 5.3.8. In $\mathbb{R}^2$, any non-zero vector $\mathbf{v}$ spans the subspace $\{c\mathbf{v} : c \in \mathbb{R}\}$. This is clearly a line through the origin. Also, try to draw a picture to convince yourself that any two non-zero vectors that are not on the same line span $\mathbb{R}^2$.

EXAMPLE 5.3.9. In $\mathbb{R}^3$, try to draw pictures to convince yourself that any non-zero vector spans a subspace which is a line through the origin; any two non-zero vectors that are not on the same line span a subspace which is a plane through the origin; and any three non-zero vectors that do not lie on the same plane span $\mathbb{R}^3$.

5.4. Linear Independence

We first study two simple examples.

EXAMPLE 5.4.1. Consider the three vectors $\mathbf{v}_1 = (1,2,3)$, $\mathbf{v}_2 = (3,2,1)$ and $\mathbf{v}_3 = (3,3,3)$ in $\mathbb{R}^3$. Then

$$\begin{aligned}\mathrm{span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\} &= \{c_1(1,2,3)+c_2(3,2,1)+c_3(3,3,3) : c_1,c_2,c_3 \in \mathbb{R}\}\\ &= \{(c_1+3c_2+3c_3, 2c_1+2c_2+3c_3, 3c_1+c_2+3c_3) : c_1,c_2,c_3 \in \mathbb{R}\}.\end{aligned}$$

Write $(x,y,z) = (c_1+3c_2+3c_3, 2c_1+2c_2+3c_3, 3c_1+c_2+3c_3)$. Then it is not difficult to see that

$$\begin{pmatrix} x\\ y\\ z\end{pmatrix} = \begin{pmatrix} 1 & 3 & 3\\ 2 & 2 & 3\\ 3 & 1 & 3\end{pmatrix}\begin{pmatrix} c_1\\ c_2\\ c_3\end{pmatrix},$$

and so (do not worry if you cannot understand why we take this next step)

$$(\,1 \quad -2 \quad 1\,)\begin{pmatrix} x\\ y\\ z\end{pmatrix} = (\,1 \quad -2 \quad 1\,)\begin{pmatrix} 1 & 3 & 3\\ 2 & 2 & 3\\ 3 & 1 & 3\end{pmatrix}\begin{pmatrix} c_1\\ c_2\\ c_3\end{pmatrix} = (\,0 \quad 0 \quad 0\,)\begin{pmatrix} c_1\\ c_2\\ c_3\end{pmatrix} = (\,0\,),$$

so that $x-2y+z=0$. It follows that $\mathrm{span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ is a plane through the origin and not $\mathbb{R}^3$. Note, in fact, that $3\mathbf{v}_1+3\mathbf{v}_2-4\mathbf{v}_3 = \mathbf{0}$. Note also that

$$\det\begin{pmatrix} 1 & 3 & 3\\ 2 & 2 & 3\\ 3 & 1 & 3\end{pmatrix} = 0.$$

EXAMPLE 5.4.2. Consider the three vectors $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (5, 1, -3)$ and $\mathbf{v}_3 = (2, 7, 4)$ in $\mathbb{R}^3$. Then

$$\begin{aligned}\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} &= \{c_1(1, 1, 0) + c_2(5, 1, -3) + c_3(2, 7, 4) : c_1, c_2, c_3 \in \mathbb{R}\} \\ &= \{(c_1 + 5c_2 + 2c_3, c_1 + c_2 + 7c_3, -3c_2 + 4c_3) : c_1, c_2, c_3 \in \mathbb{R}\}.\end{aligned}$$

Write $(x, y, z) = (c_1 + 5c_2 + 2c_3, c_1 + c_2 + 7c_3, -3c_2 + 4c_3)$. Then it is not difficult to see that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

so that

$$\begin{aligned}\begin{pmatrix} -25 & 26 & -33 \\ 4 & -4 & 5 \\ 3 & -3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} -25 & 26 & -33 \\ 4 & -4 & 5 \\ 3 & -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.\end{aligned}$$

It follows that for every $(x, y, z) \in \mathbb{R}^3$, we can find $c_1, c_2, c_3 \in \mathbb{R}$ such that $(x, y, z) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Hence $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \mathbb{R}^3$. Note that

$$\det \begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix} \neq 0,$$

and that the only solution for

$$(0, 0, 0) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

is $c_1 = c_2 = c_3 = 0$.

DEFINITION. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a vector space V over $\mathbb{R}$.

(LD) We say that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly dependent if there exist $c_1, \ldots, c_r \in \mathbb{R}$, not all zero, such that $c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = \mathbf{0}$.

(LI) We say that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent if they are not linearly dependent; in other words, if the only solution of $c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = \mathbf{0}$ in $c_1, \ldots, c_r \in \mathbb{R}$ is given by $c_1 = \ldots = c_r = 0$.

EXAMPLE 5.4.3. Let us return to Example 5.4.1 and consider again the three vectors $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (3, 2, 1)$ and $\mathbf{v}_3 = (3, 3, 3)$ in $\mathbb{R}^3$. Consider the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$. This can be rewritten in matrix form as

$$\begin{pmatrix} 1 & 3 & 3 \\ 2 & 2 & 3 \\ 3 & 1 & 3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since

$$\det \begin{pmatrix} 1 & 3 & 3 \\ 2 & 2 & 3 \\ 3 & 1 & 3 \end{pmatrix} = 0,$$

the system has non-trivial solutions; for example, $(c_1, c_2, c_3) = (3, 3, -4)$, so that $3\mathbf{v}_1 + 3\mathbf{v}_2 - 4\mathbf{v}_3 = \mathbf{0}$. Hence $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent.

Example 5.4.4. Let us return to Example 5.4.2 and consider again the three vectors $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (5, 1, -3)$ and $\mathbf{v}_3 = (2, 7, 4)$ in $\mathbb{R}^3$. Consider the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$. This can be rewritten in matrix form as

$$\begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Since

$$\det \begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix} \neq 0,$$

the only solution is $c_1 = c_2 = c_3 = 0$. Hence $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.

Example 5.4.5. In the vector space $\mathcal{A}$ of all functions of the form $f : \mathbb{R} \to \mathbb{R}$ described in Example 5.1.4, the functions x, x^2 and $\sin x$ are linearly independent. To see this, note that for every $c_1, c_2, c_3 \in \mathbb{R}$, the linear combination $c_1x + c_2x^2 + c_3 \sin x$ is never identically zero unless $c_1 = c_2 = c_3 = 0$.

Example 5.4.6. In $\mathbb{R}^n$, the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$, where

$$\mathbf{e}_j = (\underbrace{0, \ldots, 0}_{j-1}, 1, \underbrace{0, \ldots, 0}_{n-j}) \quad \text{for every } j = 1, \ldots, n,$$

are linearly independent (why?).

We observe in Examples 5.4.3–5.4.4 that the determination of whether a collection of vectors in $\mathbb{R}^3$ are linearly dependent is based on whether a system of homogeneous linear equations has non-trivial solutions. The same idea can be used to prove the following result concerning $\mathbb{R}^n$.

PROPOSITION 5D. *Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in the vector space $\mathbb{R}^n$. If $r > n$, then $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly dependent.*

Proof. For every $j = 1, \ldots, r$, write

$$\mathbf{v}_j = (a_{1j}, \ldots, a_{nj}).$$

Then the equation $c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = \mathbf{0}$ can be rewritten in matrix form as

$$\begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nr} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

If $r > n$, then there are more variables than equations. It follows that there must be non-trivial solutions $c_1, \ldots, c_r \in \mathbb{R}$. Hence $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly dependent. ◯

Remarks. (1) Consider two vectors $\mathbf{v}_1 = (a_{11}, a_{21})$ and $\mathbf{v}_2 = (a_{12}, a_{22})$ in $\mathbb{R}^2$. To study linear independence, we consider the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$, which can be written in matrix form as

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent precisely when

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \neq 0.$$

This can be interpreted geometrically in the following way: The area of the parallelogram formed by the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is in fact equal to the absolute value of the determinant of the matrix formed with $\mathbf{v}_1$ and $\mathbf{v}_2$ as the columns; in other words,

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

It follows that the two vectors are linearly dependent precisely when the parallelogram has zero area; in other words, when the two vectors lie on the same line. On the other hand, if the parallelogram has positive area, then the two vectors are linearly independent.

(2) Consider three vectors $\mathbf{v}_1 = (a_{11}, a_{21}, a_{31})$, $\mathbf{v}_2 = (a_{12}, a_{22}, a_{32})$, and $\mathbf{v}_3 = (a_{13}, a_{23}, a_{33})$ in $\mathbb{R}^3$. To study linear independence, we consider the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$, which can be written in matrix form as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent precisely when

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \neq 0.$$

This can be interpreted geometrically in the following way: The volume of the parallelepiped formed by the three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ is in fact equal to the absolute value of the determinant of the matrix formed with $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ as the columns; in other words,

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

It follows that the three vectors are linearly dependent precisely when the parallelepiped has zero volume; in other words, when the three vectors lie on the same plane. On the other hand, if the parallelepiped has positive volume, then the three vectors are linearly independent.

(3) What is the geometric interpretation of two linearly independent vectors in $\mathbb{R}^3$? Well, note that if $\mathbf{v}_1$ and $\mathbf{v}_2$ are non-zero and linearly dependent, then there exist $c_1, c_2 \in \mathbb{R}$, not both zero, such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$. This forces the two vectors to be multiples of each other, so that they lie on the same line, whence the parallelogram they form has zero area. It follows that if two vectors in $\mathbb{R}^3$ form a parallelogram with positive area, then they are linearly independent.

5.5. Basis and Dimension

In this section, we complete the task of describing uniquely every element of a vector space V in terms of the elements of a suitable subset B. To motivate the ideas, we first consider an example.

EXAMPLE 5.5.1. Let us consider the three vectors $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (5, 1, -3)$ and $\mathbf{v}_3 = (2, 7, 4)$ in $\mathbb{R}^3$, as in Examples 5.4.2 and 5.4.4. We have already shown that $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \mathbb{R}^3$, and that the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent. Furthermore, we have shown that for every $\mathbf{u} = (x, y, z) \in \mathbb{R}^3$, we can write $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, where $c_1, c_2, c_3 \in \mathbb{R}$ are determined uniquely by

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} -25 & 26 & -33 \\ 4 & -4 & 5 \\ 3 & -3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

DEFINITION. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a vector space V over $\mathbb{R}$. We say that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a basis for V if the following two conditions are satisfied:
(B1) We have $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = V$.
(B2) The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent.

EXAMPLE 5.5.2. Consider two vectors $\mathbf{v}_1 = (a_{11}, a_{21})$ and $\mathbf{v}_2 = (a_{12}, a_{22})$ in $\mathbb{R}^2$. Suppose that

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \neq 0;$$

in other words, suppose that the parallelogram formed by the two vectors has non-zero area. Then it follows from Remark (1) in Section 5.4 that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. Furthermore, for every $\mathbf{u} = (x, y) \in \mathbb{R}^2$, there exist $c_1, c_2 \in \mathbb{R}$ such that $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Indeed, c_1 and c_2 are determined as the unique solution of the system

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \mathbb{R}^2$. It follows that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^2$.

EXAMPLE 5.5.3. Consider three vectors of the type $\mathbf{v}_1 = (a_{11}, a_{21}, a_{31})$, $\mathbf{v}_2 = (a_{12}, a_{22}, a_{32})$ and $\mathbf{v}_3 = (a_{13}, a_{23}, a_{33})$ in $\mathbb{R}^3$. Suppose that

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \neq 0;$$

in other words, suppose that the parallelepiped formed by the three vectors has non-zero volume. Then it follows from Remark (2) in Section 5.4 that $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent. Furthermore, for every $\mathbf{u} = (x, y, z) \in \mathbb{R}^3$, there exist $c_1, c_2, c_3 \in \mathbb{R}$ such that $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Indeed, c_1, c_2 and c_3 are determined as the unique solution of the system

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Hence $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \mathbb{R}^3$. It follows that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\mathbb{R}^3$.

EXAMPLE 5.5.4. In $\mathbb{R}^n$, the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$, where

$$\mathbf{e}_j = (\underbrace{0, \ldots, 0}_{j-1}, 1, \underbrace{0, \ldots, 0}_{n-j}) \quad \text{for every } j = 1, \ldots, n,$$

are linearly independent and span $\mathbb{R}^n$. Hence $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$. This is known as the standard basis for $\mathbb{R}^n$.

EXAMPLE 5.5.5. In the vector space $\mathcal{M}_{2,2}(\mathbb{R})$ of all 2×2 matrices with entries in $\mathbb{R}$ as discussed in Example 5.1.3, the set

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

is a basis.

EXAMPLE 5.5.6. In the vector space P_k of polynomials of degree at most k and with coefficients in $\mathbb{R}$ as discussed in Example 5.1.6, the set $\{1, x, x^2, \ldots, x^k\}$ is a basis.

PROPOSITION 5E. *Suppose that* $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ *is a basis for a vector space* V *over* $\mathbb{R}$*. Then every element* $\mathbf{u} \in V$ *can be expressed uniquely in the form*

$$\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r, \quad \textit{where } c_1, \ldots, c_r \in \mathbb{R}.$$

PROOF. Since $\mathbf{u} \in V = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, there exist $c_1, \ldots, c_r \in \mathbb{R}$ such that $\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r$. Suppose now that $b_1, \ldots, b_r \in \mathbb{R}$ such that

$$c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = b_1\mathbf{v}_1 + \ldots + b_r\mathbf{v}_r.$$

Then

$$(c_1 - b_1)\mathbf{v}_1 + \ldots + (c_r - b_r)\mathbf{v}_r = \mathbf{0}.$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent, it follows that $c_1 - b_1 = \ldots = c_r - b_r = 0$. Hence $c_1, \ldots, c_r$ are uniquely determined. ○

We have shown earlier that a vector space can have many bases. For example, any collection of three vectors not on the same plane is a basis for $\mathbb{R}^3$. In the following discussion, we attempt to find out some properties of bases. However, we shall restrict our discussion to the following simple case.

DEFINITION. A vector space V over $\mathbb{R}$ is said to be finite-dimensional if it has a basis containing only finitely many elements.

EXAMPLE 5.5.7. The vector spaces $\mathbb{R}^n$, $\mathcal{M}_{2,2}(\mathbb{R})$ and P_k that we have discussed earlier are all finite-dimensional.

Recall that in $\mathbb{R}^n$, the standard basis has exactly n elements. On the other hand, it follows from Proposition 5D that any basis for $\mathbb{R}^n$ cannot contain more than n elements. However, can a basis for $\mathbb{R}^n$ contain fewer than n elements?

We shall answer this question by showing that all bases for a given vector space have the same number of elements. As a first step, we establish the following generalization of Proposition 5D.

PROPOSITION 5F. *Suppose that* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *is a basis for a vector space* V *over* $\mathbb{R}$*. Suppose further that* $r > n$*, and that the vectors* $\mathbf{u}_1, \ldots, \mathbf{u}_r \in V$*. Then the vectors* $\mathbf{u}_1, \ldots, \mathbf{u}_r$ *are linearly dependent.*

PROOF. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for the vector space V, we can write

$$\begin{aligned}
\mathbf{u}_1 &= a_{11}\mathbf{v}_1 + \ldots + a_{n1}\mathbf{v}_n, \\
&\vdots \\
\mathbf{u}_r &= a_{1r}\mathbf{v}_1 + \ldots + a_{nr}\mathbf{v}_n,
\end{aligned}$$

where $a_{ij} \in \mathbb{R}$ for every $i = 1, \ldots, n$ and $j = 1, \ldots, r$. Let $c_1, \ldots, c_r \in \mathbb{R}$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, it follows that if

$$\begin{aligned}
c_1\mathbf{u}_1 + \ldots + c_r\mathbf{u}_r &= c_1(a_{11}\mathbf{v}_1 + \ldots + a_{n1}\mathbf{v}_n) + \ldots + c_r(a_{1r}\mathbf{v}_1 + \ldots + a_{nr}\mathbf{v}_n) \\
&= (a_{11}c_1 + \ldots + a_{1r}c_r)\mathbf{v}_1 + \ldots + (a_{n1}c_1 + \ldots + a_{nr}c_r)\mathbf{v}_n \\
&= \mathbf{0},
\end{aligned}$$

then $a_{11}c_1 + \ldots + a_{1r}c_r = \ldots = a_{n1}c_1 + \ldots + a_{nr}c_r = 0$; in other words, we have the homogeneous system

$$\begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nr} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

If $r > n$, then there are more variables than equations. It follows that there must be non-trivial solutions $c_1, \ldots, c_r \in \mathbb{R}$. Hence $\mathbf{u}_1, \ldots, \mathbf{u}_r$ are linearly dependent. ◯

PROPOSITION 5G. *Suppose that V is a finite-dimensional vector space V over $\mathbb{R}$. Then any two bases for V have the same number of elements.*

PROOF. Note simply that by Proposition 5F, the vectors in the "basis" with more elements must be linearly dependent, and so cannot be a basis. ◯

We are now in a position to make the following definition.

DEFINITION. Suppose that V is a finite-dimensional vector space over $\mathbb{R}$. Then we say that V is of dimension n if a basis for V contains exactly n elements.

EXAMPLE 5.5.8. The vector space $\mathbb{R}^n$ has dimension n.

EXAMPLE 5.5.9. The vector space $\mathcal{M}_{2,2}(\mathbb{R})$ of all 2×2 matrices with entries in $\mathbb{R}$, as discussed in Example 5.1.3, has dimension 4.

EXAMPLE 5.5.10. The vector space P_k of all polynomials of degree at most k and with coefficients in $\mathbb{R}$, as discussed in Example 5.1.6, has dimension $(k+1)$.

EXAMPLE 5.5.11. Recall Example 5.2.5, where we showed that the set of solutions to a system of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$. Consider now the homogeneous system

$$\begin{pmatrix} 1 & 3 & -5 & 1 & 5 \\ 1 & 4 & -7 & 3 & -2 \\ 1 & 5 & -9 & 5 & -9 \\ 0 & 3 & -6 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The solutions can be described in the form

$$\mathbf{x} = c_1 \begin{pmatrix} 1 \\ -2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 3 \\ 0 \\ -5 \\ -1 \end{pmatrix},$$

where $c_1, c_2 \in \mathbb{R}$ (the reader must check this). It can be checked that $(1, -2, -1, 0, 0)$ and $(1, 3, 0, -5, -1)$ are linearly independent and so form a basis for the space of solutions of the system. It follows that the space of solutions of the system has dimension 2.

Suppose that V is an n-dimensional vector space over $\mathbb{R}$. Then any basis for V consists of exactly n linearly independent vectors in V. Suppose now that we have a set of n linearly independent vectors in V. Will this form a basis for V?

We have already answered this question in the affirmative in the cases when the vector space is $\mathbb{R}^2$ or $\mathbb{R}^3$. To seek an answer to the general case, we first establish the following result.

PROPOSITION 5H. *Suppose that V is a finite-dimensional vector space over $\mathbb{R}$. Then any finite set of linearly independent vectors in V can be expanded, if necessary, to a basis for V.*

PROOF. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a finite set of linearly independent vectors in V. If S spans V, then the proof is complete. If S does not span V, then there exists $\mathbf{v}_{k+1} \in V$ that is not a linear combination of the elements of S. The set $T = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is a finite set of linearly independent vectors in V; for otherwise, there exist $c_1, \ldots, c_k, c_{k+1}$, not all zero, such that

$$c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} = \mathbf{0}.$$

If $c_{k+1} = 0$, then $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$, contradicting the assumption that S is a finite set of linearly independent vectors in V. If $c_{k+1} \neq 0$, then

$$\mathbf{v}_{k+1} = -\frac{c_1}{c_{k+1}}\mathbf{v}_1 - \ldots - \frac{c_k}{c_{k+1}}\mathbf{v}_k,$$

contradicting the assumption that $\mathbf{v}_{k+1}$ is not a linear combination of the elements of S. We now study the finite set T of linearly independent vectors in V. If T spans V, then the proof is complete. If T does not span V, then we repeat the argument. Note that the number of vectors in a linearly independent expansion of S cannot exceed the dimension of V, in view of Proposition 5F. So eventually some linearly independent expansion of S will span V. ◯

PROPOSITION 5J. *Suppose that V is an n-dimensional vector space over $\mathbb{R}$. Then any set of n linearly independent vectors in V is a basis for V.*

PROOF. Let S be a set of n linearly independent vectors in V. By Proposition 5H, S can be expanded, if necessary, to a basis for V. By Proposition 5F, any expansion of S will result in a linearly dependent set of vectors in V. It follows that S is already a basis for V. ◯

EXAMPLE 5.5.12. Consider the three vectors $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (3, 2, 1)$ and $\mathbf{v}_3 = (3, 3, 3)$ in $\mathbb{R}^3$, as in Examples 5.4.1 and 5.4.3. We showed that these three vectors are linearly dependent, and span the plane $x - 2y + z = 0$. Note that

$$\mathbf{v}_3 = \frac{3}{4}\mathbf{v}_1 + \frac{3}{4}\mathbf{v}_2,$$

and that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. Consider now the vector $\mathbf{v}_4 = (0, 0, 1)$. Note that $\mathbf{v}_4$ does not lie on the plane $x - 2y + z = 0$, so that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ form a linearly independent set. It follows that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is a basis for $\mathbb{R}^3$.

Problems for Chapter 5

1. Determine whether each of the following subsets of $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$:
   a) $\{(x, y, z) \in \mathbb{R}^3 : x = 0\}$
   b) $\{(x, y, z) \in \mathbb{R}^3 : x + y = 0\}$
   c) $\{(x, y, z) \in \mathbb{R}^3 : xz = 0\}$
   d) $\{(x, y, z) \in \mathbb{R}^3 : y \geq 0\}$
   e) $\{(x, y, z) \in \mathbb{R}^3 : x = y = z\}$

2. For each of the following collections of vectors, determine whether the first vector is a linear combination of the remaining ones:
   a) $(1, 2, 3)$; $(1, 0, 1)$, $(2, 1, 0)$ in $\mathbb{R}^3$
   b) $x^3 + 2x^2 + 3x + 1$; x^3, $x^2 + 3x$, $x^2 + 1$ in P_4
   c) $(1, 3, 5, 7)$; $(1, 0, 1, 0)$, $(0, 1, 0, 1)$, $(0, 0, 1, 1)$ in $\mathbb{R}^4$

3. For each of the following collections of vectors, determine whether the vectors are linearly independent:
   a) $(1, 2, 3)$, $(1, 0, 1)$, $(2, 1, 0)$ in $\mathbb{R}^3$
   b) $(1, 2)$, $(3, 5)$, $(-1, 3)$ in $\mathbb{R}^2$
   c) $(2, 5, -3, 6)$, $(1, 0, 0, 1)$, $(4, 0, 9, 6)$ in $\mathbb{R}^4$
   d) $x^2 + 1$, $x + 1$, $x^2 + x$ in P_3

4. Find the volume of the parallelepiped in $\mathbb{R}^3$ formed by the vectors $(1, 2, 3)$, $(1, 0, 1)$ and $(3, 0, 2)$.

5. Let S be the set of all functions y that satisfy the differential equation

$$2\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - 3\frac{\mathrm{d}y}{\mathrm{d}x} + y = 0.$$

   Show that S is a subspace of the vector space $\mathcal{A}$ described in Example 5.1.4.

6. For each of the sets in Problem 1 which is a subspace of $\mathbb{R}^3$, find a basis for the subspace, and then extend it to a basis for $\mathbb{R}^3$.