CHAPTER 3

VECTOR SPACES

We return now to elaborate on the geometric discussion of solutions of systems of linear equations initiated in Chapter 1. Because every solution of a homogeneous system of linear equations is given as a linear combination of vectors, we should view the sets of solutions geometrically as generalizations of lines, planes, and hyperplanes. Intuitively, lines and planes differ in that it takes only one free variable (parameter) to describe points on a line (so a line is "one-dimensional"), but two to describe points on a plane (so a plane is "two-dimensional"). One of the goals of this chapter is to make algebraically precise the geometric notion of *dimension*, so that we may assign a dimension to every subspace of $\mathbb{R}^n$. Finally, at the end of this chapter, we shall see that these ideas extend far beyond the realm of $\mathbb{R}^n$ to the notion of an "abstract" vector space.

1 Subspaces of $\mathbb{R}^n$

In Chapter 1 we learned to write the general solution of a system of linear equations in standard form; one consequence of this procedure is that it enables us to express the solution set of a *homogeneous* system as the span of a particular set of vectors. The alert reader will realize she learned one way of reversing this process in Chapter 1, and we will learn others shortly. However, we should stop to understand that the span of a set of vectors in $\mathbb{R}^n$ and the set of solutions of a homogeneous system of linear equations share some salient properties.

Definition. A set $V \subset \mathbb{R}^n$ (a *subset* of $\mathbb{R}^n$) is called a *subspace* of $\mathbb{R}^n$ if it satisfies all the following properties:

1. $\mathbf{0} \in V$ (the zero vector belongs to V).
2. Whenever $\mathbf{v} \in V$ and $c \in \mathbb{R}$, we have $c\mathbf{v} \in V$ (V is closed under scalar multiplication).
3. Whenever $\mathbf{v}, \mathbf{w} \in V$, we have $\mathbf{v} + \mathbf{w} \in V$ (V is closed under addition).

EXAMPLE 1

Let's begin with some familiar examples.

(a) The *trivial subspace* consisting of just the zero vector $\mathbf{0} \in \mathbb{R}^n$ is a subspace, since $c\mathbf{0} = \mathbf{0}$ for any scalar c and $\mathbf{0} + \mathbf{0} = \mathbf{0}$.

(b) $\mathbb{R}^n$ itself is a subspace of $\mathbb{R}^n$.

(c) Any line ℓ through the origin in $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$: If the direction vector of ℓ is $\mathbf{u} \in \mathbb{R}^n$, this means that

$$\ell = \{t\mathbf{u} : t \in \mathbb{R}\}.$$

To prove that ℓ is a subspace, we must check that the three criteria hold:

1. Setting $t = 0$, we see that $\mathbf{0} \in \ell$.
2. If $\mathbf{v} \in \ell$ and $c \in \mathbb{R}$, then $\mathbf{v} = t\mathbf{u}$ for some $t \in \mathbb{R}$, and so $c\mathbf{v} = c(t\mathbf{u}) = (ct)\mathbf{u}$, which is again a scalar multiple of $\mathbf{u}$ and hence an element of ℓ.
3. If $\mathbf{v}, \mathbf{w} \in \ell$, this means that $\mathbf{v} = s\mathbf{u}$ and $\mathbf{w} = t\mathbf{u}$ for some scalars s and t. Then $\mathbf{v} + \mathbf{w} = s\mathbf{u} + t\mathbf{u} = (s + t)\mathbf{u}$, so $\mathbf{v} + \mathbf{w} \in \ell$, as needed.

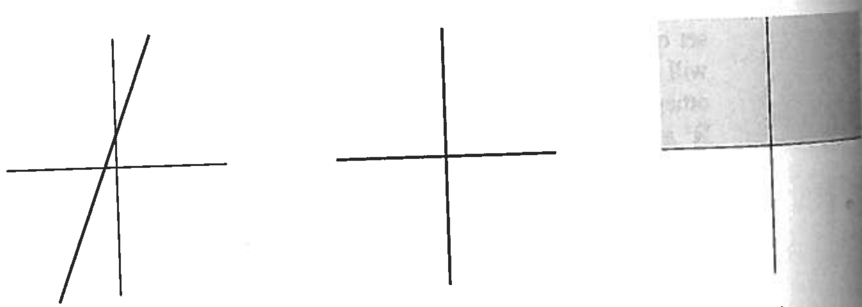
(d) Similarly, any plane through the origin in $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$. We leave this to the reader to check, but it is a special case of Proposition 1.2 below.

(e) Let $\mathbf{a} \in \mathbb{R}^n$ be a nonzero vector, and consider the hyperplane passing through the origin defined by $V = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a} \cdot \mathbf{x} = 0\}$. Recall that $\mathbf{a}$ is the normal vector of the hyperplane. We claim that V is a subspace. As expected, we check the three criteria:

1. Since $\mathbf{a} \cdot \mathbf{0} = 0$, we conclude that $\mathbf{0} \in V$.
2. Suppose $\mathbf{v} \in V$ and $c \in \mathbb{R}$. Then $\mathbf{a} \cdot (c\mathbf{v}) = c(\mathbf{a} \cdot \mathbf{v}) = c0 = 0$, and so $c\mathbf{v} \in V$ as well.
3. Suppose $\mathbf{v}, \mathbf{w} \in V$. Then $\mathbf{a} \cdot (\mathbf{v} + \mathbf{w}) = (\mathbf{a} \cdot \mathbf{v}) + (\mathbf{a} \cdot \mathbf{w}) = 0 + 0 = 0$, and therefore $\mathbf{v} + \mathbf{w} \in V$, as we needed to show. ▲

EXAMPLE 2

Let's consider next a few subsets of $\mathbb{R}^2$ that are *not* subspaces, as pictured in Figure 1.1.



(a) (b) (c)

Not subspaces of $\mathbb{R}^2$

FIGURE 1.1

As we commented on p. 93, to show that a multi-part definition *fails*, we only need to find *one* of the criteria that does not hold.

(a) $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 2x_1 + 1\}$ is not a subspace. All three criteria fail, but it suffices to point out $\mathbf{0} \notin S$.

(b) $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 = 0\}$ is not a subspace. Each of the vectors $\mathbf{v} = (1, 0)$ and $\mathbf{w} = (0, 1)$ lies in S, and yet their sum $\mathbf{v} + \mathbf{w} = (1, 1)$ does not.

(c) $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0\}$ is not a subspace. The vector $\mathbf{v} = (0, 1)$ lies in S, and yet any negative scalar multiple of it, e.g., $(-2)\mathbf{v} = (0, -2)$, does not. ▲

We now return to our motivating discussion. First, we consider the solution set of a homogeneous linear system.

Proposition 1.1. *Let A be an $m \times n$ matrix, and consider the set of solutions of the homogeneous system of linear equations $A\mathbf{x} = \mathbf{0}$; that is, let*

$$V = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

Then V is a subspace of $\mathbb{R}^n$.

Proof. The proof is essentially the same as Example 1(e) if we think of the equation $A\mathbf{x} = \mathbf{0}$ as being the collection of equations $A_1 \cdot \mathbf{x} = A_2 \cdot \mathbf{x} = \cdots = A_m \cdot \mathbf{x} = 0$. But we would rather phrase the argument in terms of the linearity properties of matrix multiplication, discussed in Section 1 of Chapter 2.

As usual, we need only check that the three defining criteria all hold.

1. To check that $\mathbf{0} \in V$, we recall that $A\mathbf{0} = \mathbf{0}$, as a consequence of either of our ways of thinking of matrix multiplication.
2. If $\mathbf{v} \in V$ and $c \in \mathbb{R}$, then we must show that $c\mathbf{v} \in V$. Well, $A(c\mathbf{v}) = c(A\mathbf{v}) = c\mathbf{0} = \mathbf{0}$.
3. If $\mathbf{v}, \mathbf{w} \in V$, then we must show that $\mathbf{v} + \mathbf{w} \in V$. Since $A\mathbf{v} = A\mathbf{w} = \mathbf{0}$, we have $A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, as required.

Thus, V is indeed a subspace of $\mathbb{R}^n$. □

Next, let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be vectors in $\mathbb{R}^n$. In Chapter 1 we defined $\mathrm{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ to be the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_k$; that is,

$$\mathrm{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \text{ for some scalars } c_1, \dots, c_k\}.$$

Generalizing what we observed in Examples 1(c) and (d), we have the following proposition.

Proposition 1.2. *Let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$. Then $V = \mathrm{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ is a subspace of $\mathbb{R}^n$.*

Proof. We check that all three criteria hold.

1. To see that $\mathbf{0} \in V$, we merely take $c_1 = c_2 = \cdots = c_k = 0$. Then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k = \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$.
2. Suppose $\mathbf{v} \in V$ and $c \in \mathbb{R}$. By definition, there are scalars $c_1, \dots, c_k$ so that $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. Thus,
$$c\mathbf{v} = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 + \cdots + (cc_k)\mathbf{v}_k,$$
which is again a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$, so $c\mathbf{v} \in V$, as desired.
3. Suppose $\mathbf{v}, \mathbf{w} \in V$. This means there are scalars $c_1, \dots, c_k$ and $d_1, \dots, d_k$ so that[1]
$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \quad \text{and} \quad \mathbf{w} = d_1\mathbf{v}_1 + \cdots + d_k\mathbf{v}_k;$$

[1] This might be a good time to review the content of the box following Exercise 1.1.22.

adding, we obtain

$$\begin{aligned}\mathbf{v}+\mathbf{w} &= (c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k)+(d_1\mathbf{v}_1+\cdots+d_k\mathbf{v}_k)\\ &= (c_1+d_1)\mathbf{v}_1+\cdots+(c_k+d_k)\mathbf{v}_k,\end{aligned}$$

which is again a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$ and hence an element of V.

This completes the verification that V is a subspace of $\mathbb{R}^n$. □

Remark. Let $V \subset \mathbb{R}^n$ be a subspace and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$. Then of course the subspace $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ is a subset of V. We say that $\mathbf{v}_1, \dots, \mathbf{v}_k$ *span* V if $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = V$. (The point here is that *every* vector in V must be a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$.)

EXAMPLE 3

The plane

$$\mathcal{P}_1 = \left\{ s\begin{bmatrix}1\\-1\\2\end{bmatrix} + t\begin{bmatrix}2\\0\\1\end{bmatrix} : s, t \in \mathbb{R} \right\}$$

is the span of the vectors

$$\mathbf{v}_1 = \begin{bmatrix}1\\-1\\2\end{bmatrix} \quad\text{and}\quad \mathbf{v}_2 = \begin{bmatrix}2\\0\\1\end{bmatrix}$$

and is therefore a subspace of $\mathbb{R}^3$. On the other hand, the plane

$$\mathcal{P}_2 = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix} + s\begin{bmatrix}1\\-1\\2\end{bmatrix} + t\begin{bmatrix}2\\0\\1\end{bmatrix} : s, t \in \mathbb{R} \right\}$$

is not a subspace. This is most easily verified by checking that $\mathbf{0} \notin \mathcal{P}_2$. Well, $\mathbf{0} \in \mathcal{P}_2$ precisely when we can find values of s and t such that

$$\begin{bmatrix}0\\0\\0\end{bmatrix} = \begin{bmatrix}1\\0\\0\end{bmatrix} + s\begin{bmatrix}1\\-1\\2\end{bmatrix} + t\begin{bmatrix}2\\0\\1\end{bmatrix}.$$

This amounts to the system of equations

$$\begin{aligned} s + 2t &= -1\\ -s &= 0\\ 2s + t &= 0,\end{aligned}$$

which we easily see is inconsistent.

A word of warning here: We might have expressed $\mathcal{P}_1$ in the form

$$\left\{ \begin{bmatrix}1\\1\\-1\end{bmatrix} + s\begin{bmatrix}1\\-1\\2\end{bmatrix} + t\begin{bmatrix}2\\0\\1\end{bmatrix} : s, t \in \mathbb{R} \right\};$$

the presence of the "shifting" term may not prevent the plane from passing through the origin. ▲

EXAMPLE 4

Let

$$\mathcal{P}_1 = \text{Span}\left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right) \quad \text{and} \quad \mathcal{P}_2 = \text{Span}\left(\begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\1\\2\end{bmatrix}\right).$$

We wish to find all the vectors contained in *both* $\mathcal{P}_1$ and $\mathcal{P}_2$, i.e., the intersection $\mathcal{P}_1 \cap \mathcal{P}_2$.

A vector $\mathbf{x}$ lies in both $\mathcal{P}_1$ and $\mathcal{P}_2$ if and only if we can write $\mathbf{x}$ in both the forms

$$\mathbf{x} = a\begin{bmatrix}1\\0\\0\end{bmatrix} + b\begin{bmatrix}0\\1\\1\end{bmatrix} \quad \text{and} \quad \mathbf{x} = c\begin{bmatrix}-1\\1\\0\end{bmatrix} + d\begin{bmatrix}2\\1\\2\end{bmatrix}$$

for some scalars a, b, c, and d. Setting the two expressions for $\mathbf{x}$ equal to one another and moving all the vectors to one side, we obtain the system of equations

$$-a\begin{bmatrix}1\\0\\0\end{bmatrix} - b\begin{bmatrix}0\\1\\1\end{bmatrix} + c\begin{bmatrix}-1\\1\\0\end{bmatrix} + d\begin{bmatrix}2\\1\\2\end{bmatrix} = \mathbf{0}.$$

In other words, we want to find all solutions of the system

$$\begin{bmatrix}1 & 0 & -1 & 2\\0 & 1 & 1 & 1\\0 & 1 & 0 & 2\end{bmatrix}\begin{bmatrix}-a\\-b\\c\\d\end{bmatrix} = \mathbf{0},$$

and so we reduce the matrix

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

to reduced echelon form

$$R = \begin{bmatrix}1 & 0 & 0 & 1\\0 & 1 & 0 & 2\\0 & 0 & 1 & -1\end{bmatrix}$$

and find that every solution of $A\mathbf{y} = \mathbf{0}$ is a scalar multiple of the vector

$$\begin{bmatrix}-a\\-b\\c\\d\end{bmatrix} = \begin{bmatrix}-1\\-2\\1\\1\end{bmatrix}.$$

This means that

$$\mathbf{x} = 1\begin{bmatrix}1\\0\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\\1\end{bmatrix} = 1\begin{bmatrix}-1\\1\\0\end{bmatrix} + 1\begin{bmatrix}2\\1\\2\end{bmatrix} = \begin{bmatrix}1\\2\\2\end{bmatrix}$$

spans the intersection of $\mathcal{P}_1$ and $\mathcal{P}_2$. We expected such a result on geometric grounds, since the intersection of two distinct planes through the origin in $\mathbb{R}^3$ should be a line. ▲

We ask the reader to show in Exercise 6 that, more generally, the intersection of subspaces is again always a subspace. We now investigate some other ways to concoct new subspaces from old.

EXAMPLE 5

Let U and V be subspaces of $\mathbb{R}^n$. We define their *sum* to be

$$U + V = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{u} + \mathbf{v} \text{ for some } \mathbf{u} \in U \text{ and } \mathbf{v} \in V\}.$$

That is, $U + V$ consists of all vectors that can be obtained by adding *some* vector in U to *some* vector in V, as shown in Figure 1.2. Be careful to note that, unless one of U or V is

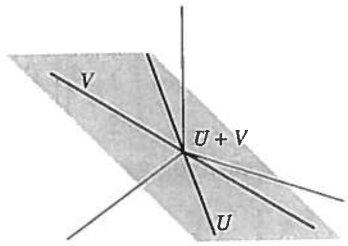


FIGURE 1.2

contained in the other, $U + V$ is much larger than $U \cup V$. We check that if U and V are subspaces, then $U + V$ is again a subspace:

1. Since $\mathbf{0} \in U$ and $\mathbf{0} \in V$, we have $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U + V$.
2. Suppose $\mathbf{x} \in U + V$ and $c \in \mathbb{R}$. We are to show that $c\mathbf{x} \in U + V$. By definition, $\mathbf{x}$ can be written in the form
$$\mathbf{x} = \mathbf{u} + \mathbf{v} \qquad \text{for some } \mathbf{u} \in U \quad \text{and} \quad \mathbf{v} \in V.$$
Then we have
$$c\mathbf{x} = c(\mathbf{u} + \mathbf{v}) = (c\mathbf{u}) + (c\mathbf{v}) \in U + V,$$
noting that $c\mathbf{u} \in U$ and $c\mathbf{v} \in V$ since each of U and V is closed under scalar multiplication.
3. Suppose $\mathbf{x}, \mathbf{y} \in U + V$. Then
$$\mathbf{x} = \mathbf{u} + \mathbf{v} \quad \text{and} \quad \mathbf{y} = \mathbf{u}' + \mathbf{v}' \qquad \text{for some } \mathbf{u}, \mathbf{u}' \in U \quad \text{and} \quad \mathbf{v}, \mathbf{v}' \in V.$$
Therefore, we have
$$\mathbf{x} + \mathbf{y} = (\mathbf{u} + \mathbf{v}) + (\mathbf{u}' + \mathbf{v}') = (\mathbf{u} + \mathbf{u}') + (\mathbf{v} + \mathbf{v}') \in U + V,$$
noting that $\mathbf{u} + \mathbf{u}' \in U$ and $\mathbf{v} + \mathbf{v}' \in V$ since U and V are both closed under addition.

Thus, as required, $U + V$ is a subspace. Indeed, it is the smallest subspace containing both U and V. (See Exercise 7.) ▲

Given an $m \times n$ matrix A, we can think of the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ as the set of all vectors that are orthogonal to each of the row vectors $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$, and, hence, by Exercise 1.2.11, are orthogonal to every vector in $V = \text{Span}\,(\mathbf{A}_1, \dots, \mathbf{A}_m)$. This leads us to a very important and natural notion.

Definition. Given a subspace $V \subset \mathbb{R}^n$, define

$$V^{\perp} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{v} = 0 \ \text{ for every } \mathbf{v} \in V\}.$$

$V^{\perp}$ (read "V perp") is called the *orthogonal complement* of V.[2] (See Figure 1.3.)

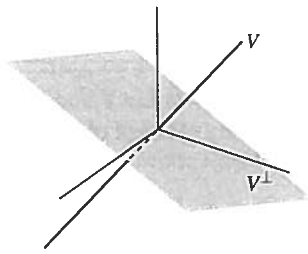


FIGURE 1.3

Proposition 1.3. *$V^{\perp}$ is a subspace of $\mathbb{R}^n$.*

Proof. We check the requisite three properties.

1. $\mathbf{0} \in V^{\perp}$ because $\mathbf{0} \cdot \mathbf{v} = 0$ for every $\mathbf{v} \in V$.
2. Suppose $\mathbf{x} \in V^{\perp}$ and $c \in \mathbb{R}$. We must check that $c\mathbf{x} \in V^{\perp}$. We calculate
$$(c\mathbf{x}) \cdot \mathbf{v} = c(\mathbf{x} \cdot \mathbf{v}) = 0$$
for all $\mathbf{v} \in V$, as required.
3. Suppose $\mathbf{x}, \mathbf{y} \in V^{\perp}$; we must check that $\mathbf{x} + \mathbf{y} \in V^{\perp}$. Well,
$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{v} = (\mathbf{x} \cdot \mathbf{v}) + (\mathbf{y} \cdot \mathbf{v}) = 0 + 0 = 0$$
for all $\mathbf{v} \in V$, as needed. □

EXAMPLE 6

Let $V = \text{Span}\,\big((1, 2, 1)\big) \subset \mathbb{R}^3$. Then $V^{\perp}$ is by definition the plane $W = \{\mathbf{x} : x_1 + 2x_2 + x_3 = 0\}$. And what is $W^{\perp}$? Clearly, any multiple of $(1, 2, 1)$ must be orthogonal to every vector in W; but is $\text{Span}\,\big((1, 2, 1)\big)$ *all* of $W^{\perp}$? Common sense suggests that the answer is yes, but let's be sure.

We know that the vectors

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

span W (why?), so we can find $W^{\perp}$ by solving the equations

$$(-2, 1, 0) \cdot \mathbf{x} = (-1, 0, 1) \cdot \mathbf{x} = 0.$$

[2]In fact, both this definition and Proposition 1.3 work just fine for any subset $V \subset \mathbb{R}^n$.

By finding the reduced echelon form of the coefficient matrix

$$A = \begin{bmatrix} -2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix},$$

we see that, indeed, every vector in $W^\perp$ is a multiple of $(1, 2, 1)$, as we suspected. ▲

It is extremely important to observe that if $\mathbf{c} \in V^\perp$, then all the elements of V satisfy the linear equation $\mathbf{c} \cdot \mathbf{x} = 0$. Thus, there is an intimate relation between elements of $V^\perp$ and Cartesian equations defining the subspace V. We will explore and exploit this relation more fully in the next few sections.

It will be useful for us to make the following definition.

Definition. Let V and W be subspaces of $\mathbb{R}^n$. We say V and W are *orthogonal subspaces* if every element of V is orthogonal to every element of W, i.e., if

$$\mathbf{v} \cdot \mathbf{w} = 0 \quad \text{for every } \mathbf{v} \in V \text{ and every } \mathbf{w} \in W.$$

Remark. If $V = W^\perp$ or $W = V^\perp$, then clearly V and W are orthogonal subspaces. On the other hand, if V and W are orthogonal subspaces of $\mathbb{R}^n$, then certainly $W \subset V^\perp$ and $V \subset W^\perp$. (See Exercise 12.) Of course, W need not be equal to $V^\perp$: Consider, for example, V to be the x_1-axis and W to be the x_2-axis in $\mathbb{R}^3$. Then $V^\perp$ is the x_2x_3-plane, which contains W and more. It is natural, however, to ask the following question: If $W = V^\perp$, must $V = W^\perp$? We will return to this shortly.

Exercises 3.1

*1. Which of the following are subspaces? Justify your answer in each case.

a. $\{\mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 = 1\}$

b. $\left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \begin{bmatrix} a \\ b \\ a+b \end{bmatrix} \text{ for some } a, b \in \mathbb{R}\right\}$

c. $\{\mathbf{x} \in \mathbb{R}^3 : x_1 + 2x_2 < 0\}$

d. $\{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$

e. $\{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 0\}$

f. $\{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = -1\}$

g. $\left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = s\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \text{ for some } s, t \in \mathbb{R}\right\}$

h. $\left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \text{ for some } s, t \in \mathbb{R}\right\}$

i. $\left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix} + s\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \text{ for some } s, t \in \mathbb{R}\right\}$

*2. Decide whether each of the following collections of vectors spans $\mathbb{R}^3$.
   a. $(1, 1, 1), (1, 2, 2)$
   b. $(1, 1, 1), (1, 2, 2), (1, 3, 3)$
   c. $(1, 0, 1), (1, -1, 1), (3, 5, 3), (2, 3, 2)$
   d. $(1, 0, -1), (2, 1, 1), (0, 1, 5)$

*3. Criticize the following argument: For any vector $\mathbf{v}$, we have $0\mathbf{v} = \mathbf{0}$. So the first criterion for subspaces is, in fact, a consequence of the second criterion and could therefore be omitted.

4. Let A be an $n \times n$ matrix. Verify that
$$V = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = 3\mathbf{x}\}$$
is a subspace of $\mathbb{R}^n$.

5. Let A and B be $m \times n$ matrices. Show that
$$V = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = B\mathbf{x}\}$$
is a subspace of $\mathbb{R}^n$.

6. a. Let U and V be subspaces of $\mathbb{R}^n$. Define the *intersection* of U and V to be
$$U \cap V = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in U \text{ and } \mathbf{x} \in V\}.$$
Show that $U \cap V$ is a subspace of $\mathbb{R}^n$. Give two examples.
   b. Is $U \cup V = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in U \text{ or } \mathbf{x} \in V\}$ always a subspace of $\mathbb{R}^n$? Give a proof or counterexample.

7. Prove that if U and V are subspaces of $\mathbb{R}^n$ and W is a subspace of $\mathbb{R}^n$ containing all the vectors of U and all the vectors of V (that is, $U \subset W$ and $V \subset W$), then $U + V \subset W$. This means that $U + V$ is the smallest subspace containing both U and V.

8. Let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ and let $\mathbf{v} \in \mathbb{R}^n$. Prove that
$$\operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}) \quad \text{if and only if} \quad \mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

9. Determine the intersection of the subspaces $\mathcal{P}_1$ and $\mathcal{P}_2$ in each case:
   *a. $\mathcal{P}_1 = \operatorname{Span}\big((1, 0, 1), (2, 1, 2)\big)$, $\mathcal{P}_2 = \operatorname{Span}\big((1, -1, 0), (1, 3, 2)\big)$
   b. $\mathcal{P}_1 = \operatorname{Span}\big((1, 2, 2), (0, 1, 1)\big)$, $\mathcal{P}_2 = \operatorname{Span}\big((2, 1, 1), (1, 0, 0)\big)$
   c. $\mathcal{P}_1 = \operatorname{Span}\big((1, 0, -1), (1, 2, 3)\big)$, $\mathcal{P}_2 = \{\mathbf{x} : x_1 - x_2 + x_3 = 0\}$
   *d. $\mathcal{P}_1 = \operatorname{Span}\big((1, 1, 0, 1), (0, 1, 1, 0)\big)$, $\mathcal{P}_2 = \operatorname{Span}\big((0, 0, 1, 1), (1, 1, 0, 0)\big)$
   e. $\mathcal{P}_1 = \operatorname{Span}\big((1, 0, 1, 2), (0, 1, 0, -1)\big)$, $\mathcal{P}_2 = \operatorname{Span}\big((1, 1, 2, 1), (1, 1, 0, 1)\big)$

♯*10. Let $V \subset \mathbb{R}^n$ be a subspace. Show that $V \cap V^\perp = \{\mathbf{0}\}$.

11. Suppose V and W are orthogonal subspaces of $\mathbb{R}^n$, i.e., $\mathbf{v} \cdot \mathbf{w} = 0$ for every $\mathbf{v} \in V$ and every $\mathbf{w} \in W$. Prove that $V \cap W = \{\mathbf{0}\}$.

♯*12. Suppose V and W are orthogonal subspaces of $\mathbb{R}^n$, i.e., $\mathbf{v} \cdot \mathbf{w} = 0$ for every $\mathbf{v} \in V$ and every $\mathbf{w} \in W$. Prove that $V \subset W^\perp$.

♯13. Let $V \subset \mathbb{R}^n$ be a subspace. Show that $V \subset (V^\perp)^\perp$. Do you think more is true?

♯14. Let V and W be subspaces of $\mathbb{R}^n$ with the property that $V \subset W$. Prove that $W^\perp \subset V^\perp$.

15. Let A be an $m \times n$ matrix. Let $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$ be subspaces.
   a. Show that $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \in W\}$ is a subspace of $\mathbb{R}^n$.
   b. Show that $\{\mathbf{y} \in \mathbb{R}^m : \mathbf{y} = A\mathbf{x} \text{ for some } \mathbf{x} \in V\}$ is a subspace of $\mathbb{R}^m$.

16. Suppose A is a symmetric $n \times n$ matrix. Let $V \subset \mathbb{R}^n$ be a subspace with the property that $A\mathbf{x} \in V$ for every $\mathbf{x} \in V$. Show that $A\mathbf{y} \in V^\perp$ for all $\mathbf{y} \in V^\perp$.

17. Use Exercises 13 and 14 to prove that for any subspace $V \subset \mathbb{R}^n$, we have $V^\perp = \big((V^\perp)^\perp\big)^\perp$.

18. Suppose U and V are subspaces of $\mathbb{R}^n$. Prove that $(U + V)^\perp = U^\perp \cap V^\perp$.

2 The Four Fundamental Subspaces

As we have seen, two of the most important constructions we've studied in linear algebra—the span of a collection of vectors and the set of solutions of a homogeneous linear system of equations—lead to subspaces. Let's use these notions to define four important subspaces associated to an $m \times n$ matrix.

The first two are already quite familiar to us from our work in Chapter 1, and we have seen in Section 1 of this chapter that they are in fact subspaces. Here we will give them their official names.

Definition (Nullspace). Let A be an $m \times n$ matrix. The *nullspace* of A is the set of solutions of the homogeneous system $A\mathbf{x} = \mathbf{0}$:

$$\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

Definition (Column Space). Let A be an $m \times n$ matrix with column vectors $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$. We define the *column space* of A to be the subspace of $\mathbb{R}^m$ spanned by the column vectors:

$$\mathbf{C}(A) = \mathrm{Span}\,(\mathbf{a}_1, \dots, \mathbf{a}_n) \subset \mathbb{R}^m.$$

Of course, the nullspace, $\mathbf{N}(A)$, is just the set of solutions of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$ that we first encountered in Section 4 of Chapter 1. What is less obvious is that we encountered the column space, $\mathbf{C}(A)$, in Section 5 of Chapter 1, as we now see.

Proposition 2.1. *Let A be an $m \times n$ matrix. Let $\mathbf{b} \in \mathbb{R}^m$. Then $\mathbf{b} \in \mathbf{C}(A)$ if and only if $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$. That is,*

$$\mathbf{C}(A) = \{\mathbf{b} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{b} \text{ is consistent}\}.$$

Proof. By definition, $\mathbf{C}(A) = \mathrm{Span}\,(\mathbf{a}_1, \dots, \mathbf{a}_n)$, and so $\mathbf{b} \in \mathbf{C}(A)$ if and only if $\mathbf{b}$ is a linear combination of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$; i.e., $\mathbf{b} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ for some scalars $x_1, \dots, x_n$. Recalling our crucial observation $(*)$ on p. 53, we conclude that $\mathbf{b} \in \mathbf{C}(A)$ if and only if $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$. The final reformulation is straightforward so long as we remember that the system $A\mathbf{x} = \mathbf{b}$ is consistent provided it has a solution. $\square$

Remark. If, as in Section 1 of Chapter 2, we think of A as giving a function $\mu_A\colon \mathbb{R}^n \to \mathbb{R}^m$, then $\mathbf{C}(A) \subset \mathbb{R}^m$ is the set of all the values of the function μ_A, i.e., the *image* of μ_A. It is important to keep track of where each subspace "lives" as you continue through this chapter: The nullspace $\mathbf{N}(A)$ consists of $\mathbf{x}$'s (inputs of μ_A) and is a subspace of $\mathbb{R}^n$; the column space $\mathbf{C}(A)$ consists of $\mathbf{b}$'s (outputs of the function μ_A) and is a subspace of $\mathbb{R}^m$.

A theme we explored in Chapter 1 was that lines and planes can be described either parametrically or by Cartesian equations. This idea should work for general subspaces of $\mathbb{R}^n$. We give a *parametric* description of a subspace V when we describe V as the span of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$. Putting these vectors as the columns of a matrix A amounts to writing $V = \mathbf{C}(A)$. Similarly, giving *Cartesian equations* for V, once we translate them into matrix form, is giving $V = \mathbf{N}(A)$ for the appropriate matrix A.[3] Much of Sections 4 and 5 of Chapter 1 was devoted to going from one description to the other: In our present language, by finding the general solution of $A\mathbf{x} = \mathbf{0}$, we obtain a parametric description of $\mathbf{N}(A)$ and thus obtain vectors that span that subspace. On the other hand, finding the constraint equations for $A\mathbf{x} = \mathbf{b}$ to be consistent provides a set of Cartesian equations for $\mathbf{C}(A)$.

EXAMPLE 1

Let

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

Of course, we bring A to its reduced echelon form

$$R = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and read off the general solution of $A\mathbf{x} = \mathbf{0}$:

$$\begin{aligned} x_1 &= -x_3 - x_4 \\ x_2 &= \phantom{-x_3 - {}} x_4 \\ x_3 &= x_3 \\ x_4 &= \phantom{-x_3 - {}} x_4, \end{aligned}$$

that is,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_3 - x_4 \\ x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

From this we see that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

span $\mathbf{N}(A)$.

On other hand, we know that the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ span $\mathbf{C}(A)$. To find Cartesian equations for $\mathbf{C}(A)$, we find the constraint equations for $A\mathbf{x} = \mathbf{b}$ to be consistent

[3]The astute reader may be worried that we have not yet shown that *every* subspace can be described in either manner. We will address this matter in Section 4.

by reducing the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & -1 & 1 & 2 & b_1 \\ 1 & 0 & 1 & 1 & b_2 \\ 1 & 2 & 1 & -1 & b_3 \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} 1 & -1 & 1 & 2 & b_1 \\ 0 & 1 & 0 & -1 & b_2 - b_1 \\ 0 & 0 & 0 & 0 & 2b_1 - 3b_2 + b_3 \end{array}\right],$$

from which we see that $2b_1 - 3b_2 + b_3 = 0$ gives a Cartesian description of $\mathbf{C}(A)$. Of course, we might want to replace b's with x's and just write

$$\mathbf{C}(A) = \{\mathbf{x} \in \mathbb{R}^3 : 2x_1 - 3x_2 + x_3 = 0\}.$$

We can summarize these results by defining new matrices

$$X = \begin{bmatrix} -1 & -1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 2 & -3 & 1 \end{bmatrix},$$

and then we have $\mathbf{N}(A) = \mathbf{C}(X)$ and $\mathbf{C}(A) = \mathbf{N}(Y)$. One final remark: Note that the coefficients of the constraint equation(s), i.e., the row(s) of Y, give vectors orthogonal to $\mathbf{C}(A)$, just as the rows of A are orthogonal to $\mathbf{N}(A)$ (and hence to the columns of X). ▲

We now move on to discuss the last two of the four subspaces associated to the matrix A. In the interest of fair play, since we've already dedicated a subspace to the columns of A, it is natural to make the following definition.

Definition (Row Space). Let A be an $m \times n$ matrix with row vectors $\mathbf{A}_1, \dots, \mathbf{A}_m \in \mathbb{R}^n$. We define the *row space* of A to be the subspace of $\mathbb{R}^n$ spanned by the row vectors $\mathbf{A}_1, \dots, \mathbf{A}_m$:

$$\mathbf{R}(A) = \operatorname{Span}(\mathbf{A}_1, \dots, \mathbf{A}_m) \subset \mathbb{R}^n.$$

It is important to remember that, as vectors in $\mathbb{R}^n$, the $\mathbf{A}_i$ are still represented by column vectors with n entries. But we continue our practice of writing vectors in parentheses when it is typographically more convenient.

Noting that $\mathbf{R}(A) = \mathbf{C}(A^\mathsf{T})$, it is natural then to complete the quartet as follows:

Definition (Left Nullspace). We define the *left nullspace* of the $m \times n$ matrix A to be

$$\mathbf{N}(A^\mathsf{T}) = \{\mathbf{x} \in \mathbb{R}^m : A^\mathsf{T}\mathbf{x} = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x}^\mathsf{T} A = \mathbf{0}^\mathsf{T}\}.$$

(The latter description accounts for the terminology.)

Just as elements of the nullspace of A give us the linear combinations of the *column* vectors of A that result in the zero vector, elements of the left nullspace give us the linear combinations of the *row* vectors of A that result in zero.

Once again, we pause to remark on the "locations" of the subspaces. $\mathbf{N}(A)$ and $\mathbf{R}(A)$ are "neighbors," both being subspaces of $\mathbb{R}^n$ (the domain of the linear map μ_A). $\mathbf{C}(A)$ and $\mathbf{N}(A^\mathsf{T})$ are "neighbors" in $\mathbb{R}^m$, the range of μ_A and the domain of μ_{A^T}. We will soon have a more complete picture of the situation.

In the discussion leading up to Proposition 1.3 we observed that vectors in the nullspace of A are orthogonal to all the row vectors of A—that is, that $\mathbf{N}(A)$ and $\mathbf{R}(A)$ are orthogonal subspaces. In fact, the orthogonality relations among our "neighboring" subspaces will provide a lot of information about linear maps. We begin with the following proposition.

Proposition 2.2. *Let A be an $m \times n$ matrix. Then* $\mathbf{N}(A) = \mathbf{R}(A)^{\perp}$.

Proof. If $\mathbf{x} \in \mathbf{N}(A)$, then $\mathbf{x}$ is orthogonal to each row vector $\mathbf{A}_1, \dots, \mathbf{A}_m$ of A. By Exercise 1.2.11, $\mathbf{x}$ is orthogonal to every vector in $\mathbf{R}(A)$ and is therefore an element of $\mathbf{R}(A)^{\perp}$. Thus, $\mathbf{N}(A)$ is a subset of $\mathbf{R}(A)^{\perp}$, and so we need only show that $\mathbf{R}(A)^{\perp}$ is a subset of $\mathbf{N}(A)$. (Recall the box on p. 12.) If $\mathbf{x} \in \mathbf{R}(A)^{\perp}$, this means that $\mathbf{x}$ is orthogonal to every vector in $\mathbf{R}(A)$, so, in particular, $\mathbf{x}$ is orthogonal to each of the row vectors $\mathbf{A}_1, \dots, \mathbf{A}_m$. But this means that $A\mathbf{x} = \mathbf{0}$, so $\mathbf{x} \in \mathbf{N}(A)$, as required. □

Since $\mathbf{C}(A) = \mathbf{R}(A^{\mathsf{T}})$, when we substitute A^{T} for A the following result is an immediate consequence of Proposition 2.2.

Proposition 2.3. *Let A be an $m \times n$ matrix. Then* $\mathbf{N}(A^{\mathsf{T}}) = \mathbf{C}(A)^{\perp}$.

Proposition 2.3 has a very pleasant interpretation in terms of the constraint equations for $A\mathbf{x} = \mathbf{b}$ to be consistent—the Cartesian equations for $\mathbf{C}(A)$. As we commented in Section 1, the coefficients of such a Cartesian equation give a vector orthogonal to $\mathbf{C}(A)$, i.e., an element of $\mathbf{C}(A)^{\perp} = \mathbf{N}(A^{\mathsf{T}})$. Thus, a constraint equation gives a linear combination of the rows that results in the zero vector. But, of course, this is where constraint equations come from in the first place. Conversely, any such relation among the row vectors of A gives an element of $\mathbf{N}(A^{\mathsf{T}}) = \mathbf{C}(A)^{\perp}$, and hence the coefficients of a constraint equation that $\mathbf{b}$ must satisfy in order for $A\mathbf{x} = \mathbf{b}$ to be consistent.

EXAMPLE 2

Let

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

We find the constraint equations for $A\mathbf{x} = \mathbf{b}$ to be consistent by row reducing the augmented matrix:

$$\left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 1 & 1 & b_2 \\ 0 & 1 & b_3 \\ 1 & 2 & b_4 \end{array}\right] \rightsquigarrow \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & -1 & b_2 - b_1 \\ 0 & 1 & b_3 \\ 0 & 0 & b_4 - b_1 \end{array}\right] \rightsquigarrow \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & 1 & b_1 - b_2 \\ 0 & 0 & -b_1 + b_2 + b_3 \\ 0 & 0 & -b_1 + b_4 \end{array}\right].$$

The constraint equations are

$$\begin{aligned} -b_1 + b_2 + b_3 &= 0 \\ -b_1 + b_4 &= 0. \end{aligned}$$

Note that the vectors

$$\mathbf{c}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{c}_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

are in $\mathbf{N}(A^{\mathsf{T}})$ and correspond to linear combinations of the rows yielding $\mathbf{0}$. ▲

Proposition 2.3 tells us that $\mathbf{N}(A^\top) = \mathbf{C}(A)^\perp$, and so $\mathbf{N}(A^\top)$ and $\mathbf{C}(A)$ are orthogonal subspaces. It is natural, then, to ask whether $\mathbf{N}(A^\top)^\perp = \mathbf{C}(A)$, as well.

Proposition 2.4. *Let A be an $m \times n$ matrix. Then $\mathbf{C}(A) = \mathbf{N}(A^\top)^\perp$.*

Proof. Since $\mathbf{C}(A)$ and $\mathbf{N}(A^\top)$ are orthogonal subspaces, we infer from Exercise 3.1.12 that $\mathbf{C}(A) \subset \mathbf{N}(A^\top)^\perp$. On the other hand, from Section 5 of Chapter 1 we know that there is a system of constraint equations

$$\mathbf{c}_1 \cdot \mathbf{b} = \cdots = \mathbf{c}_k \cdot \mathbf{b} = 0$$

that give necessary and sufficient conditions for $\mathbf{b} \in \mathbb{R}^m$ to belong to $\mathbf{C}(A)$. Setting $V = \operatorname{Span}(\mathbf{c}_1, \dots, \mathbf{c}_k) \subset \mathbb{R}^m$, this means that $\mathbf{C}(A) = V^\perp$. Since each such vector $\mathbf{c}_j$ is an element of $\mathbf{C}(A)^\perp = \mathbf{N}(A^\top)$, we conclude that $V \subset \mathbf{N}(A^\top)$. It follows from Exercise 3.1.14 that $\mathbf{N}(A^\top)^\perp \subset V^\perp = \mathbf{C}(A)$. Combining the two inclusions, we have $\mathbf{C}(A) = \mathbf{N}(A^\top)^\perp$, as required. $\square$

Now that we have proved Proposition 2.4, we can complete the circle of ideas. We have the following result, summarizing the geometric relations of the pairs of the four fundamental subspaces.

Theorem 2.5. *Let A be an $m \times n$ matrix. Then*

1. $\mathbf{R}(A)^\perp = \mathbf{N}(A)$
2. $\mathbf{N}(A)^\perp = \mathbf{R}(A)$
3. $\mathbf{C}(A)^\perp = \mathbf{N}(A^\top)$
4. $\mathbf{N}(A^\top)^\perp = \mathbf{C}(A)$

Proof. All but the second are the contents of Propositions 2.2, 2.3, and 2.4. The second follows from Proposition 2.4 by substituting $A^\top$ for A. $\square$

Figure 2.1 is a schematic diagram giving a visual representation of these results.

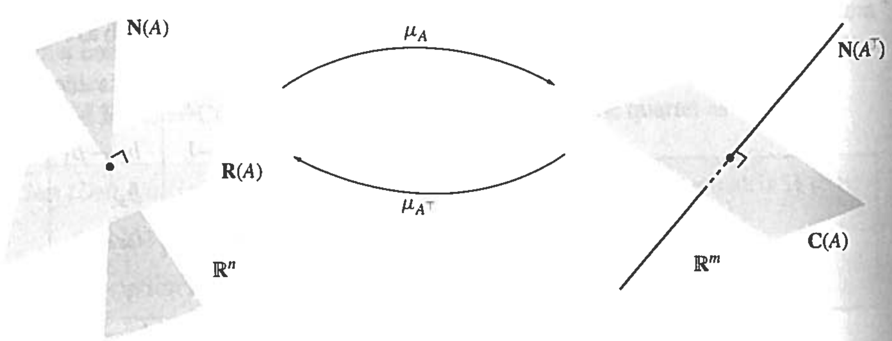


FIGURE 2.1

Remark. Combining these pairs of results, we conclude that for any of the four fundamental subspaces $V = \mathbf{R}(A), \mathbf{N}(A), \mathbf{C}(A)$, and $\mathbf{N}(A^\top)$, it is the case that $(V^\perp)^\perp = V$. If we knew that every subspace of $\mathbb{R}^n$ could be so written, we would have the result in general; this will come soon.

EXAMPLE 3

Let's look for matrices whose row spaces are the plane in $\mathbb{R}^3$ spanned by $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\\-1\end{bmatrix}$ and satisfy the extra conditions given below. Note, first of all, that these must be $m \times 3$ matrices for some positive integers m.

(a) Suppose we want such a matrix A with $\begin{bmatrix}1\\2\\-1\end{bmatrix}$ in its nullspace. Remember that $\mathbf{N}(A) = \mathbf{R}(A)^{\perp}$. We cannot succeed: Although $\begin{bmatrix}1\\2\\-1\end{bmatrix}$ is orthogonal to $\begin{bmatrix}1\\-1\\-1\end{bmatrix}$, it is not orthogonal to $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and hence not orthogonal to every vector in the row space.

(b) Suppose we want such a matrix with its column space equal to $\mathbb{R}^2$. Now we win: We need a 2×3 matrix, and we just try

$$A = \begin{bmatrix}1 & 1 & 1\\1 & -1 & -1\end{bmatrix}.$$

Then $\mathbf{C}(A) = \text{Span}\left(\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right) = \mathbb{R}^2$, as required.

(c) Suppose we want such a matrix A whose column space is spanned by $\begin{bmatrix}1\\-1\end{bmatrix}$. This seems impossible, but here's an argument to that effect. If we had such a matrix A, note that $\mathbf{C}(A)^{\perp} = \mathbf{N}(A^{\mathsf{T}})$ is spanned by $\begin{bmatrix}1\\1\end{bmatrix}$, and so we would have to have $\mathbf{A}_1 + \mathbf{A}_2 = \mathbf{0}$. This means that the row space of A is a line.

(d) Following this reasoning, let's look for a matrix A whose column space is spanned by $\begin{bmatrix}1\\0\\1\end{bmatrix}$ and $\begin{bmatrix}0\\1\\1\end{bmatrix}$. We note that A now must be a 3×3 matrix. As before, note that $\begin{bmatrix}1\\1\\-1\end{bmatrix} \in \mathbf{C}(A)^{\perp} = \mathbf{N}(A^{\mathsf{T}})$, and so the third row of A must be the sum of the first two rows. So now we just try

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

Perhaps it's not obvious that A really works, but if we add the first and second columns, we get $\begin{bmatrix}2\\0\\2\end{bmatrix}$, and if we subtract them we get $\begin{bmatrix}0\\2\\2\end{bmatrix}$, so $\mathbf{C}(A)$ contains both the desired vectors and hence their span. We leave it to the reader to check that $\mathbf{C}(A)$ is not larger than this span. ▲

Exercises 3.2

♯*1. Show that if B is obtained from A by performing one or more row operations, then $\mathbf{R}(B) = \mathbf{R}(A)$.

2. What vectors $\mathbf{b}$ are in the column space of A in each case? (Give constraint equations.) Check that the coefficients of the constraint equations give linear combinations of the rows of A summing to $\mathbf{0}$.

*a. $A = \begin{bmatrix}3&-1\\6&-2\\-9&3\end{bmatrix}$ *b. $A = \begin{bmatrix}1&1&1\\-1&1&2\\1&3&-5\end{bmatrix}$ c. $A = \begin{bmatrix}1&-2&1\\0&1&3\\-1&3&0\\1&0&-1\end{bmatrix}$

3. Given each matrix A, find matrices X and Y so that $\mathbf{C}(A) = \mathbf{N}(X)$ and $\mathbf{N}(A) = \mathbf{C}(Y)$.

*a. $A = \begin{bmatrix}3&-1\\6&-2\\-9&3\end{bmatrix}$ b. $A = \begin{bmatrix}1&1&0\\2&1&1\\1&-1&2\end{bmatrix}$ c. $A = \begin{bmatrix}1&1&1\\1&2&0\\1&1&1\\1&0&2\end{bmatrix}$

*4. Let $A = \begin{bmatrix}1&2&1&1\\-1&0&3&4\\2&2&-2&-3\end{bmatrix}$.

a. Give constraint equations for $\mathbf{C}(A)$.

b. Find vectors spanning $\mathbf{N}(A^{\mathsf{T}})$.

*5. Let

$$L = \begin{bmatrix}1&0&0\\2&1&0\\1&-1&1\end{bmatrix} \quad\text{and}\quad U = \begin{bmatrix}1&2&1&1\\0&0&2&-2\\0&0&0&0\end{bmatrix}.$$

If $A = LU$, give vectors that span $\mathbf{R}(A)$, $\mathbf{C}(A)$, and $\mathbf{N}(A)$.

6. a. Construct a matrix whose column space contains $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}0\\1\\1\end{bmatrix}$ and whose nullspace contains $\begin{bmatrix}1\\0\\1\end{bmatrix}$ and $\begin{bmatrix}0\\1\\0\end{bmatrix}$, or explain why none can exist.

*b. Construct a matrix whose column space contains $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}0\\1\\1\end{bmatrix}$ and whose nullspace contains $\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$, or explain why none can exist.

7. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 0 & -1 & -1 \end{bmatrix}$.
   a. Give $\mathbf{C}(A)$ and $\mathbf{C}(B)$. Are they lines, planes, or all of $\mathbb{R}^3$?
   b. Describe $\mathbf{C}(A+B)$ and $\mathbf{C}(A)+\mathbf{C}(B)$. Compare your answers.

8. *a. Construct a 3×3 matrix A with $\mathbf{C}(A) \subset \mathbf{N}(A)$.
   b. Construct a 3×3 matrix A with $\mathbf{N}(A) \subset \mathbf{C}(A)$.
   c. Do you think there can be a 3×3 matrix A with $\mathbf{N}(A) = \mathbf{C}(A)$? Why or why not?
   d. Construct a 4×4 matrix A with $\mathbf{C}(A) = \mathbf{N}(A)$.

*9. Let A be an $m \times n$ matrix and recall that we have the associated function $\mu_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $\mu_A(\mathbf{x}) = A\mathbf{x}$. Show that μ_A is a one-to-one function if and only if $\mathbf{N}(A) = \{\mathbf{0}\}$.

♯10. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Prove that
   a. $\mathbf{N}(B) \subset \mathbf{N}(AB)$.
   b. $\mathbf{C}(AB) \subset \mathbf{C}(A)$. (*Hint:* Use Proposition 2.1.)
   c. $\mathbf{N}(B) = \mathbf{N}(AB)$ when A is $n \times n$ and nonsingular. (*Hint:* See the box on p. 12.)
   d. $\mathbf{C}(AB) = \mathbf{C}(A)$ when B is $n \times n$ and nonsingular.

♯11. Let A be an $m \times n$ matrix. Prove that $\mathbf{N}(A^{\mathsf{T}}A) = \mathbf{N}(A)$. (*Hint:* Use Exercise 10 and Exercise 2.5.15.)

12. Suppose A and B are $m \times n$ matrices. Prove that $\mathbf{C}(A)$ and $\mathbf{C}(B)$ are orthogonal subspaces of $\mathbb{R}^m$ if and only if $A^{\mathsf{T}}B = O$.

13. Suppose A is an $n \times n$ matrix with the property that $A^2 = A$.
   a. Prove that $\mathbf{C}(A) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = A\mathbf{x}\}$.
   b. Prove that $\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{u} - A\mathbf{u} \text{ for some } \mathbf{u} \in \mathbb{R}^n\}$.
   c. Prove that $\mathbf{C}(A) \cap \mathbf{N}(A) = \{\mathbf{0}\}$.
   d. Prove that $\mathbf{C}(A) + \mathbf{N}(A) = \mathbb{R}^n$.

3 Linear Independence and Basis

In view of our discussion in the preceding section, it is natural to ask the following question:

Given vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$, is $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$?

Of course, we recognize that this is a question of whether there *exist* scalars $c_1, \dots, c_k$ such that $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. As we are well aware, this is, in turn, a question of whether a certain (inhomogeneous) system of linear equations has a solution. As we saw in Chapter 1, one is often interested in the allied question: Is that solution *unique*?

EXAMPLE 1

Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

We ask first of all whether $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. This is a familiar question when we recast it in matrix notation: Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Is the system $A\mathbf{x} = \mathbf{b}$ consistent? Immediately we write down the appropriate augmented matrix and reduce to echelon form:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{array}\right],$$

so the system is obviously inconsistent. The answer is: No, $\mathbf{v}$ is not in $\operatorname{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.

What about

$$\mathbf{w} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}?$$

As the reader can easily check, $\mathbf{w} = 3\mathbf{v}_1 - \mathbf{v}_3$, so $\mathbf{w} \in \operatorname{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. What's more, $\mathbf{w} = 2\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3$, as well. So, obviously, there is no unique expression for $\mathbf{w}$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. But we can conclude more: Setting the two expressions for $\mathbf{w}$ equal, we obtain

$$3\mathbf{v}_1 - \mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3, \quad \text{i.e.,} \quad \mathbf{v}_1 + \mathbf{v}_2 - 2\mathbf{v}_3 = \mathbf{0}.$$

That is, there is a nontrivial relation among the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, and this is why we have different ways of expressing $\mathbf{w}$ as a linear combination of the three of them. Indeed, because $\mathbf{v}_1 = -\mathbf{v}_2 + 2\mathbf{v}_3$, we can see easily that any linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ is a linear combination of just $\mathbf{v}_2$ and $\mathbf{v}_3$:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = c_1(-\mathbf{v}_2 + 2\mathbf{v}_3) + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = (c_2 - c_1)\mathbf{v}_2 + (c_3 + 2c_1)\mathbf{v}_3.$$

The vector $\mathbf{v}_1$ was redundant, since

$$\operatorname{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \operatorname{Span}(\mathbf{v}_2, \mathbf{v}_3).$$

We might surmise that the vector $\mathbf{w}$ can now be written *uniquely* as a linear combination of $\mathbf{v}_2$ and $\mathbf{v}_3$. This is easy to check with an augmented matrix:

$$[A' \mid \mathbf{w}] = \left[\begin{array}{cc|c} 1 & 1 & 2 \\ -1 & 0 & 3 \\ 0 & 1 & 5 \end{array}\right] \rightsquigarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{array}\right];$$

from the fact that the matrix A' has rank 2, we infer that the system of equations has a unique solution. ▲

In the language of functions, the question of uniqueness is the question of whether the function $\mu_A\colon \mathbb{R}^3 \to \mathbb{R}^3$ is *one-to-one*. Remember that we say f is a *one-to-one* function if

$$\text{whenever } a \neq b, \text{ it must be the case that } f(a) \neq f(b).$$

Given some function $y = f(x)$, we might ask if, for a certain value r, we can solve the equation $f(x) = r$. When r is in the image of the function, there is at least one solution. Is the solution unique? If f is a one-to-one function, there can be *at most* one solution of the equation $f(x) = r$.

Next we show that the question of uniqueness we raised earlier can be reduced to one basic question, which will be crucial to all our future work.

Proposition 3.1. *Let* $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$. *If the zero vector has a* unique *expression as a linear combination of* $\mathbf{v}_1, \dots, \mathbf{v}_k$, *that is, if*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0} \implies c_1 = c_2 = \dots = c_k = 0,$$

then every *vector* $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ *has a* unique *expression as a linear combination of* $\mathbf{v}_1, \dots, \mathbf{v}_k$.

Proof. By considering the matrix A whose column vectors are $\mathbf{v}_1, \dots, \mathbf{v}_k$, we can deduce this immediately from Proposition 5.4 of Chapter 1. However, we prefer to give a coordinate-free proof that is typical of many of the arguments we shall be encountering for a while.

Suppose that for some $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ there are two expressions

$$\begin{aligned} \mathbf{v} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k \quad \text{and} \\ \mathbf{v} &= d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \dots + d_k\mathbf{v}_k. \end{aligned}$$

Then, subtracting, we obtain

$$\mathbf{0} = (c_1 - d_1)\mathbf{v}_1 + \dots + (c_k - d_k)\mathbf{v}_k.$$

Since the only way to express the zero vector as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$ is with every coefficient equal to 0, we conclude that $c_1 - d_1 = c_2 - d_2 = \dots = c_k - d_k = 0$, which means, of course, that $c_1 = d_1$, $c_2 = d_2, \dots, c_k = d_k$. That is, $\mathbf{v}$ has a unique expression as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$. □

This discussion leads us to make the following definition.

Definition. The (indexed) set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is called *linearly independent* if

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0} \implies c_1 = c_2 = \dots = c_k = 0,$$

that is, if the *only* way of expressing the zero vector as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$ is the *trivial* linear combination $0\mathbf{v}_1 + \dots + 0\mathbf{v}_k$.

The set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is called *linearly dependent* if it is not linearly independent—i.e., if there is some expression

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0}, \quad \text{where not all the } c_i\text{'s are } 0.$$

The language is problematic here. Many mathematicians—including at least one of the authors of this text—often say things like "the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent." But linear independence (or dependence) is a property of the whole *collection* of vectors, not of the individual vectors. What's worse, we really should refer to an *ordered list* of vectors rather than to a set of vectors. For example, any list in which some vector, $\mathbf{v}$, appears twice is obviously giving a linearly dependent collection, but the set $\{\mathbf{v}, \mathbf{v}\}$ is indistinguishable from the set $\{\mathbf{v}\}$. There seems to be no ideal route out of this morass! Having said all this, we warn the gentle reader that we may occasionally say, "the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly (in)dependent" where it would be too clumsy to be more pedantic. Just stay alert!!

EXAMPLE 2

We wish to decide whether the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^4$$

form a linearly independent set.

Here is a piece of advice: It is virtually always the case that when you are presented with a set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ that you are to prove linearly independent, you should write,

"Suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$. I must show that $c_1 = \cdots = c_k = 0$."

You then use whatever hypotheses you're given to arrive at that conclusion.

The definition of linear independence is a particularly subtle one, largely because of the syntax. Suppose we know that $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent. As a result, we know that *if* it should happen that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, *then* it must be that $c_1 = c_2 = \cdots = c_k = 0$. But we may never blithely assert that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$.

Suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$, i.e.,

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} = \mathbf{0}.$$

Can we conclude that $c_1 = c_2 = c_3 = 0$? We recognize this as a homogeneous system of linear equations:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \mathbf{0}.$$

By now we are old hands at solving such systems. We find that the echelon form of the coefficient matrix is

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and so our system of equations in fact has infinitely many solutions. For example, we can take $c_1 = 1$, $c_2 = -1$, and $c_3 = 1$. The vectors therefore form a linearly dependent set. ▲

EXAMPLE 3

Suppose $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. We show next that if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, then so is $\{\mathbf{u}+\mathbf{v}, \mathbf{v}+\mathbf{w}, \mathbf{u}+\mathbf{w}\}$. Suppose

$$c_1(\mathbf{u}+\mathbf{v}) + c_2(\mathbf{v}+\mathbf{w}) + c_3(\mathbf{u}+\mathbf{w}) = \mathbf{0}.$$

We must show that $c_1 = c_2 = c_3 = 0$. We use the distributive property to rewrite our equation as

$$(c_1 + c_3)\mathbf{u} + (c_1 + c_2)\mathbf{v} + (c_2 + c_3)\mathbf{w} = \mathbf{0}.$$

Since $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, we may infer that the coefficients of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ must each be equal to 0. Thus,

$$\begin{aligned} c_1 \qquad + c_3 &= 0 \\ c_1 + c_2 \qquad &= 0 \\ c_2 + c_3 &= 0, \end{aligned}$$

and we leave it to the reader to check that the only solution of this system of equations is, in fact, $c_1 = c_2 = c_3 = 0$, as desired. ▲

EXAMPLE 4

Any time one has a list of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in which one of the vectors is the zero vector, say $\mathbf{v}_1 = \mathbf{0}$, then the set of vectors must be linearly dependent, because the equation

$$1\mathbf{v}_1 = \mathbf{0}$$

is a nontrivial linear combination of the vectors yielding the zero vector. ▲

EXAMPLE 5

How can two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ give rise to a linearly dependent set? By definition, this means that there is a linear combination

$$a\mathbf{u} + b\mathbf{v} = \mathbf{0},$$

where, to start, *either* $a \neq 0$ *or* $b \neq 0$. But if, say, $a = 0$, then the equation reduces to $b\mathbf{v} = \mathbf{0}$; since $b \neq 0$, we must have $\mathbf{v} = \mathbf{0}$, which contradicts the hypothesis that the vectors are nonzero. Thus, in this case, we must have *both* a *and* $b \neq 0$. We may write $\mathbf{u} = -\frac{b}{a}\mathbf{v}$, so $\mathbf{u}$ is a scalar multiple of $\mathbf{v}$. Hence two nonzero linearly dependent vectors are parallel (and vice versa).

How can a collection of three nonzero vectors be linearly dependent? As before, there must be a linear combination

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0},$$

where (at least) one of a, b, and c is nonzero. Say $a \neq 0$. This means that we can solve

$$\mathbf{u} = -\frac{1}{a}(b\mathbf{v} + c\mathbf{w}) = \left(-\frac{b}{a}\right)\mathbf{v} + \left(-\frac{c}{a}\right)\mathbf{w},$$

so $\mathbf{u} \in \mathrm{Span}(\mathbf{v}, \mathbf{w})$. In particular, $\mathrm{Span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ is either a line (if all three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are parallel) or a plane (when $\mathbf{v}$ and $\mathbf{w}$ are nonparallel). We leave it to the reader to think about what must happen when $a = 0$. ▲

The appropriate generalization of the last example is the following useful criterion, depicted in Figure 3.1.

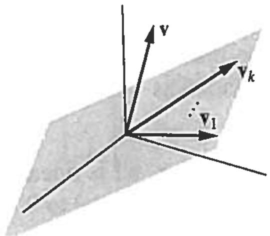


FIGURE 3.1

Proposition 3.2. *Suppose $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ form a linearly independent set, and suppose $\mathbf{v} \in \mathbb{R}^n$. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}\}$ is linearly independent if and only if $\mathbf{v} \notin \mathrm{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.*

The *contrapositive* of the statement

"if P, then Q"

is

"if Q is false, then P is false."

One of the fundamental points of logic underlying all of mathematics is that these statements are equivalent: One is true precisely when the other is. (This is quite reasonable. For instance, if Q must be true whenever P is true and we know that Q is false, then P must be false as well, for if not, Q would have had to be true.)

It probably is a bit more convincing to consider a couple of examples:

- If we believe the statement "Whenever it is raining, the ground is wet" (or "*if* it is raining, *then* the ground is wet"), we should equally well grant that "If the ground is dry, then it is not raining."
- If we believe the statement "If $x = 2$, then $x^2 = 4$," then we should believe that "if $x^2 \neq 4$, then $x \neq 2$."

It is important not to confuse the contrapositive of a statement with the *converse* of the statement. The converse of the statement "if P, then Q" is

"if Q, then P."

Note that even if we believe our two earlier statements, we do *not* believe their converses:

- "If the ground is wet, then it is raining"—it may have stopped raining a while ago, or someone may have washed a car earlier.
- "If $x^2 = 4$, then $x = 2$"—even though this is a common error, it is an error nevertheless: x might be -2.

Proof. We will prove the contrapositive: Still supposing that $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ form a linearly independent set,

$$\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}\} \text{ is linearly dependent if and only if } \mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

Suppose that $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. Then $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ for some scalars $c_1, \dots, c_k$, so

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + (-1)\mathbf{v} = \mathbf{0},$$

from which we conclude that $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}\}$ is linearly dependent (since at least one of the coefficients is nonzero).

Now suppose $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}\}$ is linearly dependent. This means that there are scalars $c_1, \dots, c_k$, and c, not all 0, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + c\mathbf{v} = \mathbf{0}.$$

Note that we cannot have $c = 0$: For if c were 0, we'd have $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, and linear independence of $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ implies $c_1 = \cdots = c_k = 0$, which contradicts our assumption that $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}\}$ is linearly dependent. Therefore $c \neq 0$, and so

$$\mathbf{v} = -\frac{1}{c}(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = \left(-\frac{c_1}{c}\right)\mathbf{v}_1 + \left(-\frac{c_2}{c}\right)\mathbf{v}_2 + \cdots + \left(-\frac{c_k}{c}\right)\mathbf{v}_k,$$

which tells us that $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, as required. □

We now understand that when we have a set of linearly independent vectors, no proper subset will yield the same span. In other words, we will have an "efficient" set of spanning vectors (that is, there is no redundancy in the vectors we've chosen; no proper subset will do). This motivates the following definition.

Definition. Let $V \subset \mathbb{R}^n$ be a subspace. The set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is called a *basis* for V if

(i) $\mathbf{v}_1, \dots, \mathbf{v}_k$ span V, that is, $V = \operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, and

(ii) $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent.

We comment that the plural of *basis* is *bases*.[4]

EXAMPLE 6

Let $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, $\dots$, $\mathbf{e}_n = (0, \dots, 0, 1) \in \mathbb{R}^n$. Then $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$, called the *standard basis*. To check this, we must establish that properties (i) and (ii) above hold for $V = \mathbb{R}^n$. The first is obvious: If $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, then $\mathbf{x} =$

[4]Pronounced *bāsēz*, to rhyme with Macy's.

$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$. The second is not much harder. Suppose $c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n = \mathbf{0}$. This means that $(c_1, c_2, \dots, c_n) = (0, 0, \dots, 0)$, and so $c_1 = c_2 = \cdots = c_n = 0$. ▲

EXAMPLE 7

Consider the plane given by $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 - x_2 + 2x_3 = 0\} \subset \mathbb{R}^3$. Our algorithms of Chapter 1 tell us that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

span V. Since these vectors are not parallel, it follows from Example 5 that they must be linearly independent.

For the practice, however, we give a direct argument. Suppose

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = \mathbf{0}.$$

Writing out the entries explicitly, we obtain

$$\begin{bmatrix} c_1 - 2c_2 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

from which we conclude that $c_1 = c_2 = 0$, as required. (For future reference, we note that this information came from the free variable "slots.") Therefore, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent and gives a basis for V, as required. ▲

The following observation may prove useful.

Corollary 3.3. *Let $V \subset \mathbb{R}^n$ be a subspace, and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a basis for V if and only if every vector of V can be written uniquely as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$.*

Proof. This is immediate from Proposition 3.1. □

This result is so important that we introduce a bit of terminology.

Definition. When we write $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$, we refer to $c_1, \dots, c_k$ as the *coordinates* of $\mathbf{v}$ with respect to the (ordered) basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

EXAMPLE 8

Consider the three vectors

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

Let's take a general vector $\mathbf{b} \in \mathbb{R}^3$ and ask first of all whether it has a unique expression as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Forming the augmented matrix and row reducing, we find

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 2 & 1 & 0 & b_2 \\ 1 & 2 & 2 & b_3 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2b_1 - b_3 \\ 0 & 1 & 0 & -4b_1 + b_2 + 2b_3 \\ 0 & 0 & 1 & 3b_1 - b_2 - b_3 \end{array}\right].$$

It follows from Corollary 3.3 that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\mathbb{R}^3$, because an arbitrary vector $\mathbf{b} \in \mathbb{R}^3$ can be written in the form

$$\mathbf{b} = \underbrace{(2b_1 - b_3)}_{c_1}\mathbf{v}_1 + \underbrace{(-4b_1 + b_2 + 2b_3)}_{c_2}\mathbf{v}_2 + \underbrace{(3b_1 - b_2 - b_3)}_{c_3}\mathbf{v}_3.$$

And, what's more,

$$\begin{aligned} c_1 &= 2b_1 - b_3, \\ c_2 &= -4b_1 + b_2 + 2b_3, \quad \text{and} \\ c_3 &= 3b_1 - b_2 - b_3 \end{aligned}$$

give the coordinates of $\mathbf{b}$ with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. ▲

Our experience in Example 8 leads us to make the following general observation:

Proposition 3.4. *Let A be an $n \times n$ matrix. Then A is nonsingular if and only if its column vectors form a basis for $\mathbb{R}^n$.*

Proof. As usual, let's denote the column vectors of A by $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Using Corollary 3.3, we are to prove that A is nonsingular if and only if every vector in $\mathbb{R}^n$ can be written uniquely as a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. But this is exactly what Proposition 5.5 of Chapter 1 tells us. □

Somewhat more generally (see Exercise 12), we have the following result.

EXAMPLE 9

Suppose A is a nonsingular $n \times n$ matrix and $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$. Then we wish to show that $\{A\mathbf{v}_1, \dots, A\mathbf{v}_n\}$ is likewise a basis for $\mathbb{R}^n$.

First, we show that $\{A\mathbf{v}_1, \dots, A\mathbf{v}_n\}$ is linearly independent. Following our ritual, we start by supposing that

$$c_1(A\mathbf{v}_1) + c_2(A\mathbf{v}_2) + \cdots + c_n(A\mathbf{v}_n) = \mathbf{0},$$

and we wish to show that $c_1 = \cdots = c_n = 0$. By linearity properties we have

$$\begin{aligned} \mathbf{0} = c_1 A\mathbf{v}_1 + c_2 A\mathbf{v}_2 + \cdots + c_n A\mathbf{v}_n &= A(c_1\mathbf{v}_1) + A(c_2\mathbf{v}_2) + \cdots + A(c_n\mathbf{v}_n) \\ &= A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n). \end{aligned}$$

Since A is nonsingular, the only solution of $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$, and so we must have $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$. From the linear independence of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ we now conclude that $c_1 = c_2 = \cdots = c_n = 0$, as required.

Now, why do these vectors span $\mathbb{R}^n$? (The result follows from Exercise 1.5.13, but we give the argument here.) Given $\mathbf{b} \in \mathbb{R}^n$, we know from Proposition 5.5 of Chapter 1 that there is a unique $\mathbf{x} \in \mathbb{R}^n$ with $A\mathbf{x} = \mathbf{b}$. Since $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis for $\mathbb{R}^n$, we can write $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ for some scalars $c_1, \dots, c_n$. Then, again by linearity properties, we have

$$\begin{aligned}\mathbf{b} = A\mathbf{x} &= A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = A(c_1\mathbf{v}_1) + A(c_2\mathbf{v}_2) + \cdots + A(c_n\mathbf{v}_n)\\ &= c_1(A\mathbf{v}_1) + c_2(A\mathbf{v}_2) + \cdots + c_n(A\mathbf{v}_n),\end{aligned}$$

as required. ▲

Given a subspace $V \subset \mathbb{R}^n$, how do we know there is some basis for it? This is a consequence of Proposition 3.2 as well.

Theorem 3.5. *Any subspace $V \subset \mathbb{R}^n$ other than the trivial subspace has a basis.*

Proof. Because $V \neq \{\mathbf{0}\}$, we can choose a nonzero vector $\mathbf{v}_1 \in V$. If $\mathbf{v}_1$ spans V, then we know $\{\mathbf{v}_1\}$ will constitute a basis for V. If not, choose $\mathbf{v}_2 \notin \text{Span}(\mathbf{v}_1)$. From Proposition 3.2 we infer that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent. If $\mathbf{v}_1, \mathbf{v}_2$ span V, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ will be a basis for V. If not, choose $\mathbf{v}_3 \notin \text{Span}(\mathbf{v}_1, \mathbf{v}_2)$. Once again, we know that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ will be linearly independent and hence will form a basis for V if the three vectors span V. We continue in this fashion, and we are guaranteed that the process will terminate in at most n steps: Once we have $n + 1$ vectors in $\mathbb{R}^n$, they must form a linearly dependent set, because an $n \times (n + 1)$ matrix has rank at most n (see Exercise 15). □

From this fact it follows that every subspace $V \subset \mathbb{R}^n$ can be expressed as the row space (or column space) of a matrix. This settles the issue raised in the footnote on p. 137. As an application, we can now follow through on the substance of the remark on p. 140.

Proposition 3.6. *Let $V \subset \mathbb{R}^n$ be a subspace. Then $(V^\perp)^\perp = V$.*

Proof. Choose a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ for V, and consider the $k \times n$ matrix A whose rows are $\mathbf{v}_1, \dots, \mathbf{v}_k$. By construction, $V = \mathbf{R}(A)$. By Theorem 2.5, $V^\perp = \mathbf{R}(A)^\perp = \mathbf{N}(A)$, and $\mathbf{N}(A)^\perp = \mathbf{R}(A)$, so $(V^\perp)^\perp = V$. □

We conclude this section with the problem of determining bases for each of the four fundamental subspaces of a matrix.

EXAMPLE 10

Let

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ 1 & 2 & 1 & 1 & 6 \\ 0 & 1 & 1 & 1 & 3 \\ 2 & 2 & 0 & 1 & 7 \end{bmatrix}.$$

Gaussian elimination gives us the reduced echelon form R:

$$R = \begin{bmatrix} 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ -1 & -1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ 1 & 2 & 1 & 1 & 6 \\ 0 & 1 & 1 & 1 & 3 \\ 2 & 2 & 0 & 1 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

From this information, we wish to find bases for $\mathbf{R}(A)$, $\mathbf{N}(A)$, $\mathbf{C}(A)$, and $\mathbf{N}(A^\top)$.

Since any row of R is a linear combination of rows of A and vice versa, it is easy to see that $\mathbf{R}(A) = \mathbf{R}(R)$ (see Exercise 3.2.1), so we concentrate on the rows of R. We may as well use only the nonzero rows of R; now we need only check that they form a linearly independent set. We keep an eye on the pivot "slots": Suppose

$$c_1 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \mathbf{0}.$$

This means that

$$\begin{bmatrix} c_1 \\ c_2 \\ -c_1 + c_2 \\ c_3 \\ c_1 + 2c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

and so $c_1 = c_2 = c_3 = 0$, as promised.

From the reduced echelon form R, we read off the vectors that span $\mathbf{N}(A)$: The general solution of $A\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} x_3 - x_5 \\ -x_3 - 2x_5 \\ x_3 \\ -x_5 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -1 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix},$$

so

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

span $\mathbf{N}(A)$. On the other hand, these vectors are linearly independent, because if we take a linear combination

$$x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -1 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix} = \mathbf{0},$$

we infer (from the free variable slots) that $x_3 = x_5 = 0$.

Obviously, $\mathbf{C}(A)$ is spanned by the five column vectors of A. But these vectors cannot be linearly independent—that's what vectors in the nullspace of A tell us. From our vectors spanning $\mathbf{N}(A)$, we know that

$$(*) \qquad \mathbf{a}_1 - \mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0} \quad \text{and} \quad -\mathbf{a}_1 - 2\mathbf{a}_2 - \mathbf{a}_4 + \mathbf{a}_5 = \mathbf{0}.$$

These equations tell us that $\mathbf{a}_3$ and $\mathbf{a}_5$ can be written as linear combinations of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_4$. If we can check that $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4\}$ is linearly independent, we'll be finished. So we form a matrix A' with these columns (easier: cross out the third and fifth columns of A), and reduce it to echelon form (easier yet: cross out the third and fifth columns of R). Well, we have

$$A' = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = R',$$

and so only the trivial linear combination of the columns of A' will yield the zero vector. In conclusion, the vectors

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{a}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

give a basis for $\mathbf{C}(A)$.

Remark. The puzzled reader may wonder why, looking at the equations ($*$), we chose to use the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_4$ and discard the vectors $\mathbf{a}_3$ and $\mathbf{a}_5$. These are the columns in which pivots appear in the echelon form; the subsequent reasoning establishes their linear independence. There might in any specific case be other viable choices for vectors to discard, but then the proof that the remaining vectors form a linearly independent set may be less straightforward.

What about the left nullspace? The only row of 0's in R arises as the linear combination

$$-\mathbf{A}_1 - \mathbf{A}_2 + \mathbf{A}_3 + \mathbf{A}_4 = \mathbf{0}$$

of the rows of A, so we expect the vector

$$\mathbf{v} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

to give a basis for $\mathbf{N}(A^{\mathsf{T}})$. As a check, we note it is orthogonal to the basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_4$ for $\mathbf{C}(A)$. Could there be any vectors in $\mathbf{C}(A)^{\perp}$ besides multiples of $\mathbf{v}$? ▲

What is lurking in the background here is a notion of dimension, and we turn to this important topic in the next section.

Exercises 3.3

1. Let $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (2, 4, 5)$, and $\mathbf{v}_3 = (2, 4, 6) \in \mathbb{R}^3$. Is each of the following statements correct or incorrect? Explain.
   a. The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.
   b. Each of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ can be written as a linear combination of the others.

*2. Decide whether each of the following sets of vectors is linearly independent.
   a. $\{(1,4),(2,9)\} \subset \mathbb{R}^2$
   b. $\{(1,4,0),(2,9,0)\} \subset \mathbb{R}^3$
   c. $\{(1,4,0),(2,9,0),(3,-2,0)\} \subset \mathbb{R}^3$
   d. $\{(1,1,1),(2,3,3),(0,1,2)\} \subset \mathbb{R}^3$
   e. $\{(1,1,1,3),(1,1,3,1),(1,3,1,1),(3,1,1,1)\} \subset \mathbb{R}^4$
   f. $\{(1,1,1,-3),(1,1,-3,1),(1,-3,1,1),(-3,1,1,1)\} \subset \mathbb{R}^4$

*3. Decide whether the following sets of vectors give a basis for the indicated space.
   a. $\{(1,2,1),(2,4,5),(1,2,3)\}$; $\mathbb{R}^3$
   b. $\{(1,0,1),(1,2,4),(2,2,5),(2,2,-1)\}$; $\mathbb{R}^3$
   c. $\{(1,0,2,3),(0,1,1,1),(1,1,4,4)\}$; $\mathbb{R}^4$
   d. $\{(1,0,2,3),(0,1,1,1),(1,1,4,4),(2,-2,1,2)\}$; $\mathbb{R}^4$

4. In each case, check that $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$ and give the coordinates of the given vector $\mathbf{b} \in \mathbb{R}^n$ with respect to that basis.

   a. $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}; \mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

   *b. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}; \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

   c. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \mathbf{b} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

   *d. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 4 \end{bmatrix}; \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

5. Following Example 10, for each of the following matrices A, give a basis for each of the subspaces $\mathbf{R}(A)$, $\mathbf{C}(A)$, $\mathbf{N}(A)$, and $\mathbf{N}(A^{\mathsf{T}})$.

   *a. $A = \begin{bmatrix} 3 & -1 \\ 6 & -2 \\ -9 & 3 \end{bmatrix}$

   *c. $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

   b. $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$

   d. $A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -1 & -1 \end{bmatrix}$

*6. Give a basis for the orthogonal complement of the subspace $W \subset \mathbb{R}^4$ spanned by $(1,1,1,2)$ and $(1,-1,5,2)$.

7. Let $V \subset \mathbb{R}^5$ be spanned by $(1,0,1,1,1)$ and $(0,1,-1,0,2)$. By finding the left nullspace of an appropriate matrix, give a homogeneous system of equations having V as its solution set. Explain how you are using Proposition 3.6.

8. Suppose $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $\{\mathbf{v}, \mathbf{w}\}$ is linearly independent. Prove that $\{\mathbf{v} - \mathbf{w},\ 2\mathbf{v} + \mathbf{w}\}$ is linearly independent as well.

9. Suppose $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ form a linearly independent set. Prove that $\mathbf{u} + \mathbf{v}$, $\mathbf{v} + 2\mathbf{w}$, and $-\mathbf{u} + \mathbf{v} + \mathbf{w}$ likewise form a linearly independent set.

♯**10.** Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are nonzero vectors with the property that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ whenever $i \neq j$. Prove that $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent. (*Hint:* "Suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$." Start by showing $c_1 = 0$.)

♯**11.** Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are nonzero, mutually orthogonal vectors in $\mathbb{R}^n$.

a. Prove that they form a basis for $\mathbb{R}^n$. (Use Exercise 10.)

b. Given any $\mathbf{x} \in \mathbb{R}^n$, give an explicit formula for the coordinates of $\mathbf{x}$ with respect to the basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

c. Deduce from your answer to part *b* that $\mathbf{x} = \sum_{i=1}^{n} \mathrm{proj}_{\mathbf{v}_i}\mathbf{x}$.

12. Give an alternative proof of Example 9 by applying Proposition 3.4 and Exercise 2.1.10.

***13.** Prove that if $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly dependent, then every vector $\mathbf{v} \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ infinitely many ways.

♯***14.** Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ form a linearly independent set. Show that for any $1 \le \ell < k$, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is linearly independent as well.

♯**15.** Suppose $k > n$. Prove that any k vectors in $\mathbb{R}^n$ must form a linearly dependent set. (So what can you conclude if you have k linearly independent vectors in $\mathbb{R}^n$?)

16. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ form a linearly dependent set. Prove that for some j between 1 and k we have $\mathbf{v}_j \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_k)$. That is, one of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ can be written as a linear combination of the remaining vectors.

17. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ form a linearly dependent set. Prove that either $\mathbf{v}_1 = \mathbf{0}$ or $\mathbf{v}_i \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$ for some $i = 2, 3, \ldots, k$. (*Hint:* There is a relation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ with at least one $c_j \neq 0$. Consider the largest such j.)

18. Let A be an $m \times n$ matrix and suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$. Prove that if $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_k\}$ is linearly independent, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ must be linearly independent.

19. Let A be an $n \times n$ matrix. Prove that if A is nonsingular and $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent, then $\{A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_k\}$ is likewise linearly independent. Give an example to show that the result is false if A is singular.

20. Suppose U and V are subspaces of $\mathbb{R}^n$. Prove that $(U \cap V)^\perp = U^\perp + V^\perp$. (*Hint:* Use Exercise 3.1.18 and Proposition 3.6.)

♯**21.** Let A be an $m \times n$ matrix of rank n. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent. Prove that $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_k\} \subset \mathbb{R}^m$ is likewise linearly independent. (**N.B.**: If you did not explicitly make use of the assumption that $\mathrm{rank}(A) = n$, your proof cannot be correct. Why?)

22. Let A be an $n \times n$ matrix and suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^n$ are nonzero vectors that satisfy

$$\begin{aligned} A\mathbf{v}_1 &= \mathbf{v}_1 \\ A\mathbf{v}_2 &= 2\mathbf{v}_2 \\ A\mathbf{v}_3 &= 3\mathbf{v}_3. \end{aligned}$$

Prove that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent. (*Hint:* Start by showing that $\{\mathbf{v}_1, \mathbf{v}_2\}$ must be linearly independent.)

***23.** Suppose U and V are subspaces of $\mathbb{R}^n$ with $U \cap V = \{\mathbf{0}\}$. If $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a basis for U and $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is a basis for V, prove that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is a basis for $U + V$.