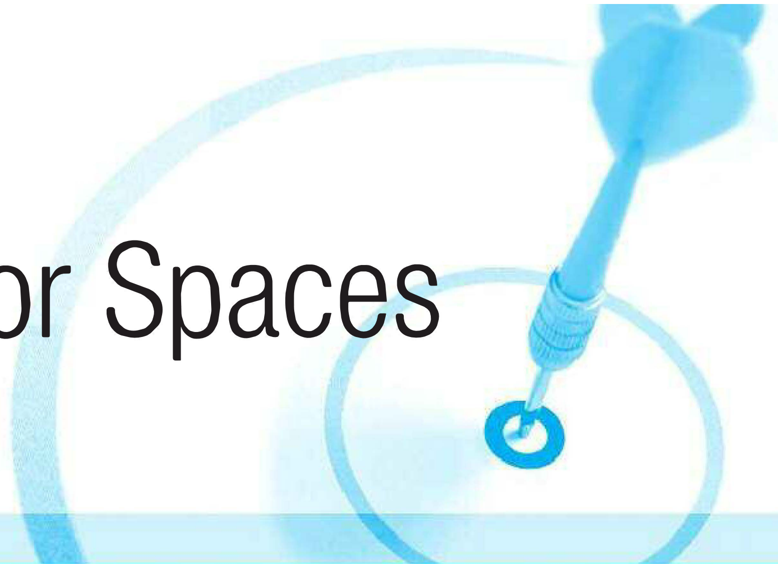
6 Vector Spaces



6.0 Introduction: Fibonacci in (Vector) Space

Algebra is generous; she often gives more than is asked of her.
—Jean le Rond d'Alembert (1717–1783)
In Carl B. Boyer
A History of Mathematics
Wiley, 1968, p. 481

The Fibonacci sequence was introduced in Section 4.6. It is the sequence

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

of nonnegative integers with the property that after the first two terms, each term is the sum of the two terms preceding it. Thus $0 + 1 = 1$, $1 + 1 = 2$, $1 + 2 = 3$, $2 + 3 = 5$, and so on.

If we denote the terms of the Fibonacci sequence by $f_0, f_1, f_2, \ldots$, then the entire sequence is completely determined by specifying that

$$f_0 = 0, f_1 = 1 \quad \text{and} \quad f_n = f_{n-1} + f_{n-2} \quad \text{for } n \geq 2$$

By analogy with vector notation, let's write a sequence $x_0, x_1, x_2, x_3, \ldots$ as

$$\mathbf{x} = [x_0, x_1, x_2, x_3, \ldots)$$

The Fibonacci sequence then becomes

$$\mathbf{f} = [f_0, f_1, f_2, f_3, \ldots) = [0, 1, 1, 2, \ldots)$$

We now generalize this notion.

Definition A ***Fibonacci-type sequence*** is *any* sequence $\mathbf{x} = [x_0, x_1, x_2, x_3, \ldots)$ such that x_0 and x_1 are real numbers and $x_n = x_{n-1} + x_{n-2}$ for $n \geq 2$.

For example, $[1, \sqrt{2}, 1 + \sqrt{2}, 1 + 2\sqrt{2}, 2 + 3\sqrt{2}, \ldots)$ is a Fibonacci-type sequence.

Problem 1 Write down the first five terms of three more Fibonacci-type sequences.

By analogy with vectors again, let's define the *sum* of two sequences $\mathbf{x} = [x_0, x_1, x_2, \ldots)$ and $\mathbf{y} = [y_0, y_1, y_2, \ldots)$ to be the sequence

$$\mathbf{x} + \mathbf{y} = [x_0 + y_0, x_1 + y_1, x_2 + y_2, \ldots)$$

If c is a scalar, we can likewise define the scalar multiple of a sequence by

$$c\mathbf{x} = [cx_0, cx_1, cx_2, \ldots)$$

Problem 2 (a) Using your examples from Problem 1 or other examples, compute the sums of various pairs of Fibonacci-type sequences. Do the resulting sequences appear to be Fibonacci-type?

(b) Compute various scalar multiples of your Fibonacci-type sequences from Problem 1. Do the resulting sequences appear to be Fibonacci-type?

Problem 3 (a) Prove that if $\mathbf{x}$ and $\mathbf{y}$ are Fibonacci-type sequences, then so is $\mathbf{x} + \mathbf{y}$.

(b) Prove that if $\mathbf{x}$ is a Fibonacci-type sequence and c is a scalar, then $c\mathbf{x}$ is also a Fibonacci-type sequence.

Let's denote the set of all Fibonacci-type sequences by *Fib*. Problem 3 shows that, like $\mathbb{R}^n$, *Fib* is closed under addition and scalar multiplication. The next exercises show that *Fib* has much more in common with $\mathbb{R}^n$.

Problem 4 Review the algebraic properties of vectors in Theorem 1.1. Does *Fib* satisfy all of these properties? What Fibonacci-type sequence plays the role of $\mathbf{0}$? For a Fibonacci-type sequence $\mathbf{x}$, what is $-\mathbf{x}$? Is $-\mathbf{x}$ also a Fibonacci-type sequence?

Problem 5 In $\mathbb{R}^n$, we have the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. The Fibonacci sequence $\mathbf{f} = [0, 1, 1, 2, \ldots)$ can be thought of as the analogue of $\mathbf{e}_2$ because its first two terms are 0 and 1. What sequence $\mathbf{e}$ in *Fib* plays the role of $\mathbf{e}_1$?

What about $\mathbf{e}_3, \mathbf{e}_4, \ldots$? Do these vectors have analogues in *Fib*?

Problem 6 Let $\mathbf{x} = [x_0, x_1, x_2, \ldots)$ be a Fibonacci-type sequence. Show that $\mathbf{x}$ is a linear combination of $\mathbf{e}$ and $\mathbf{f}$.

Problem 7 Show that $\mathbf{e}$ and $\mathbf{f}$ are linearly independent. (That is, show that if $c\mathbf{e} + d\mathbf{f} = \mathbf{0}$, then $c = d = 0$.)

Problem 8 Given your answers to Problems 6 and 7, what would be a sensible value to assign to the "dimension" of *Fib*? Why?

Problem 9 Are there any geometric sequences in *Fib*? That is, if

$$[1, r, r^2, r^3, \ldots)$$

is a Fibonacci-type sequence, what are the possible values of r?

Problem 10 Find a "basis" for *Fib* consisting of geometric Fibonacci-type sequences.

Problem 11 Using your answer to Problem 10, give an alternative derivation of *Binet's formula* [formula (5) in Section 4.6]:

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

for the terms of the Fibonacci sequence $\mathbf{f} = [f_0, f_1, f_2, \ldots)$. [*Hint:* Express $\mathbf{f}$ in terms of the basis from Problem 10.]

The Lucas sequence is named after Edouard Lucas (see page 336).

The ***Lucas sequence*** is the Fibonacci-type sequence

$$\mathbf{l} = [l_0, l_1, l_2, l_3, \ldots) = [2, 1, 3, 4, \ldots)$$

Problem 12 Use the basis from Problem 10 to find an analogue of Binet's formula for the nth term l_n of the Lucas sequence.

Problem 13 Prove that the Fibonacci and Lucas sequences are related by the identity

$$f_{n-1} + f_{n+1} = l_n \quad \text{for } n \geq 1$$

[*Hint:* The Fibonacci-type sequences $\mathbf{f}^- = [1, 1, 2, 3, \ldots)$ and $\mathbf{f}^+ = [1, 0, 1, 1, \ldots)$ form a basis for *Fib*. (Why?)]

In this Introduction, we have seen that the collection *Fib* of all Fibonacci-type sequences behaves in many respects like $\mathbb{R}^2$, even though the "vectors" are actually infinite sequences. This useful analogy leads to the general notion of a *vector space* that is the subject of this chapter.

6.1 Vector Spaces and Subspaces

In Chapters 1 and 3, we saw that the algebra of vectors and the algebra of matrices are similar in many respects. In particular, we can add both vectors and matrices, and we can multiply both by scalars. The properties that result from these two operations (Theorem 1.1 and Theorem 3.2) are identical in both settings. In this section, we use these properties to define generalized "vectors" that arise in a wide variety of examples. By proving general theorems about these "vectors," we will therefore simultaneously be proving results about all of these examples. This is the real power of algebra: its ability to take properties from a concrete setting, like $\mathbb{R}^n$, and *abstract* them into a general setting.

Definition Let V be a set on which two operations, called *addition* and *scalar multiplication*, have been defined. If $\mathbf{u}$ and $\mathbf{v}$ are in V, the *sum* of $\mathbf{u}$ and $\mathbf{v}$ is denoted by $\mathbf{u} + \mathbf{v}$, and if c is a scalar, the *scalar multiple* of $\mathbf{u}$ by c is denoted by $c\mathbf{u}$. If the following axioms hold for all $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and for all scalars c and d, then V is called a ***vector space*** and its elements are called ***vectors***.

1. $\mathbf{u} + \mathbf{v}$ is in V. — Closure under addition
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ — Commutativity
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ — Associativity
4. There exists an element $\mathbf{0}$ in V, called a ***zero vector,*** such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each $\mathbf{u}$ in V, there is an element $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. $c\mathbf{u}$ is in V. — Closure under scalar multiplication
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ — Distributivity
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ — Distributivity
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$

The German mathematician Hermann Grassmann (1809–1877) is generally credited with first introducing the idea of a vector space (although he did not call it that) in 1844. Unfortunately, his work was very difficult to read and did not receive the attention it deserved. One person who did study it was the Italian mathematician Giuseppe Peano (1858–1932). In his 1888 book *Calcolo Geometrico,* Peano clarified Grassmann's earlier work and laid down the axioms for a vector space as we know them today. Peano's book is also remarkable for introducing operations on sets. His notations ∪, ∩, and ∈ (for "union," "intersection," and "is an element of") are the ones we still use, although they were not immediately accepted by other mathematicians. Peano's axiomatic definition of a vector space also had very little influence for many years. Acceptance came in 1918, after Hermann Weyl (1885–1955) repeated it in his book *Space, Time, Matter,* an introduction to Einstein's general theory of relativity.

Remarks

- By "scalars" we will usually mean the real numbers. Accordingly, we should refer to V as a *real vector space* (or a *vector space over the real numbers*). It is also possible for scalars to be complex numbers or to belong to $\mathbb{Z}_p$, where p is prime. In these cases, V is called a *complex vector space* or a *vector space over* $\mathbb{Z}_p$, respectively. Most of our examples will be real vector spaces, so we will usually omit the adjective "real." If something is referred to as a "vector space," assume that we are working over the real number system.

  In fact, the scalars can be chosen from any number system in which, roughly speaking, we can add, subtract, multiply, and divide according to the usual laws of arithmetic. In abstract algebra, such a number system is called a ***field***.

- The definition of a vector space does not specify what the set V consists of. Neither does it specify what the operations called "addition" and "scalar multiplication' look like. Often, they will be familiar, but they need not be. See Example 6.6 and Exercises 5–7.

We will now look at several examples of vector spaces. In each case, we need to specify the set V and the operations of addition and scalar multiplication and to verify axioms 1 through 10. We need to pay particular attention to axioms 1 and 6 (closure), axiom 4 (the existence of a zero vector in V), and axiom 5 (each vector in V must have a negative in V).

Example 6.1

For any $n \geq 1$, $\mathbb{R}^n$ is a vector space with the usual operations of addition and scalar multiplication. Axioms 1 and 6 follow from the definitions of these operations, and the remaining axioms follow from Theorem 1.1.

Example 6.2

The set of all 2×3 matrices is a vector space with the usual operations of matrix addition and matrix scalar multiplication. Here the "vectors" are actually matrices. We know that the sum of two 2×3 matrices is also a 2×3 matrix and that multiplying a 2×3 matrix by a scalar gives another 2×3 matrix; hence, we have closure. The remaining axioms follow from Theorem 3.2. In particular, the zero vector $\mathbf{0}$ is the 2×3 zero matrix, and the negative of a 2×3 matrix A is just the 2×3 matrix $-A$.

There is nothing special about 2×3 matrices. For any positive integers m and n, the set of all $m \times n$ matrices forms a vector space with the usual operations of matrix addition and matrix scalar multiplication. This vector space is denoted M_{mn}.

Example 6.3

Let $\mathcal{P}_2$ denote the set of all polynomials of degree 2 or less with real coefficients. Define addition and scalar multiplication in the usual way. (See Appendix D.) If

$$p(x) = a_0 + a_1x + a_2x^2 \quad \text{and} \quad q(x) = b_0 + b_1x + b_2x^2$$

are in $\mathcal{P}_2$, then

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

has degree at most 2 and so is in $\mathcal{P}_2$. If c is a scalar, then

$$cp(x) = ca_0 + ca_1x + ca_2x^2$$

is also in $\mathcal{P}_2$. This verifies axioms 1 and 6.

The zero vector $\mathbf{0}$ is the zero polynomial—that is, the polynomial all of whose coefficients are zero. The negative of a polynomial $p(x) = a_0 + a_1x + a_2x^2$ is the polynomial $-p(x) = -a_0 - a_1x - a_2x^2$. It is now easy to verify the remaining axioms. We will check axiom 2 and leave the others for Exercise 12. With $p(x)$ and $q(x)$ as above, we have

$$\begin{aligned} p(x) + q(x) &= (a_0 + a_1x + a_2x^2) + (b_0 + b_1x + b_2x^2) \\ &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 \\ &= (b_0 + a_0) + (b_1 + a_1)x + (b_2 + a_2)x^2 \\ &= (b_0 + b_1x + b_2x^2) + (a_0 + a_1x + a_2x^2) \\ &= q(x) + p(x) \end{aligned}$$

where the third equality follows from the fact that addition of real numbers is commutative.

In general, for any fixed $n \geq 0$, the set $\mathscr{P}_n$ of all polynomials of degree less than or equal to n is a vector space, as is the set $\mathscr{P}$ of *all* polynomials.

Example 6.4

Let $\mathscr{F}$ denote the set of all real-valued functions defined on the real line. If f and g are two such functions and c is a scalar, then $f + g$ and cf are defined by

$$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (cf)(x) = cf(x)$$

In other words, the *value* of $f + g$ at x is obtained by adding together the values of f and g at x [Figure 6.1(a)]. Similarly, the value of cf at x is just the value of f at x multiplied by the scalar c [Figure 6.1(b)]. The zero vector in $\mathscr{F}$ is the constant function f_0 that is identically zero; that is, $f_0(x) = 0$ for all x. The negative of a function f is the function $-f$ defined by $(-f)(x) = -f(x)$ [Figure 6.1(c)].

Axioms 1 and 6 are obviously true. Verification of the remaining axioms is left as Exercise 13. Thus, $\mathscr{F}$ is a vector space.

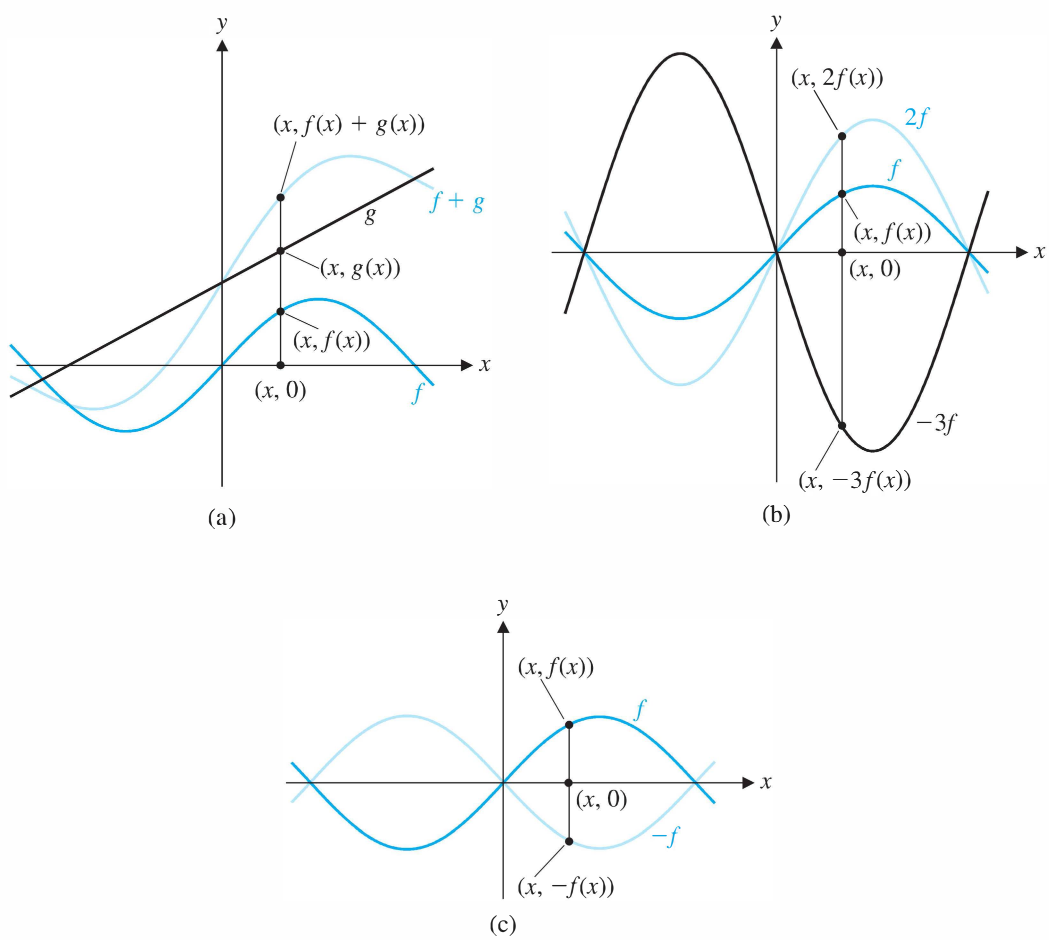


Figure 6.1
The graphs of (a) f, g, and $f + g$, (b) f, $2f$, and $-3f$, and (c) f and $-f$

In Example 6.4, we could also have considered only those functions defined on some *closed interval* $[a, b]$ of the real line. This approach also produces a vector space, denoted by $\mathscr{F}[a, b]$.

Example 6.5

The set $\mathbb{Z}$ of integers with the usual operations is *not* a vector space. To demonstrate this, it is enough to find that *one* of the ten axioms fails and to give a specific instance in which it fails (a *counterexample*). In this case, we find that we do not have closure under scalar multiplication. For example, the multiple of the integer 2 by the scalar $\frac{1}{3}$ is $(\frac{1}{3})(2) = \frac{2}{3}$, which is not an integer. Thus, it is not true that cx is in $\mathbb{Z}$ for *every* x in $\mathbb{Z}$ and *every* scalar c (i.e., axiom 6 fails).

Example 6.6

Let $V = \mathbb{R}^2$ with the usual definition of addition but the following definition of scalar multiplication:

$$c\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ 0 \end{bmatrix}$$

Then, for example,

$$1\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

so axiom 10 fails. [In fact, the other nine axioms are all true (check this), but we do not need to look into them, because V has already failed to be a vector space. This example shows the value of looking ahead, rather than working through the list of axioms in the order in which they have been given.]

a + bi

Example 6.7

Let $\mathbb{C}^2$ denote the set of all ordered pairs of complex numbers. Define addition and scalar multiplication as in $\mathbb{R}^2$, except here the scalars are complex numbers. For example,

$$\begin{bmatrix} 1+i \\ 2-3i \end{bmatrix} + \begin{bmatrix} -3+2i \\ 4 \end{bmatrix} = \begin{bmatrix} -2+3i \\ 6-3i \end{bmatrix}$$

and

$$(1-i)\begin{bmatrix} 1+i \\ 2-3i \end{bmatrix} = \begin{bmatrix} (1-i)(1+i) \\ (1-i)(2-3i) \end{bmatrix} = \begin{bmatrix} 2 \\ -1-5i \end{bmatrix}$$

Using properties of the complex numbers, it is straightforward to check that all ten axioms hold. Therefore, $\mathbb{C}^2$ is a complex vector space.

In general, $\mathbb{C}^n$ is a complex vector space for all $n \geq 1$.

Example 6.8

If p is prime, the set $\mathbb{Z}_p^n$ (with the usual definitions of addition and multiplication by scalars from $\mathbb{Z}_p$) is a vector space over $\mathbb{Z}_p$ for all $n \geq 1$.

Before we consider further examples, we state a theorem that contains some useful properties of vector spaces. It is important to note that, by proving this theorem for vector spaces in *general*, we are actually proving it for *every specific* vector space.

Theorem 6.1 Let V be a vector space, $\mathbf{u}$ a vector in V, and c a scalar.

a. $0\mathbf{u} = \mathbf{0}$
b. $c\mathbf{0} = \mathbf{0}$
c. $(-1)\mathbf{u} = -\mathbf{u}$
d. If $c\mathbf{u} = \mathbf{0}$, then $c = 0$ or $\mathbf{u} = \mathbf{0}$.

Proof We prove properties (b) and (d) and leave the proofs of the remaining properties as exercises.

(b) We have

$$c\mathbf{0} = c(\mathbf{0} + \mathbf{0}) = c\mathbf{0} + c\mathbf{0}$$

by vector space axioms 4 and 7. Adding the negative of $c\mathbf{0}$ to both sides produces

$$c\mathbf{0} + (-c\mathbf{0}) = (c\mathbf{0} + c\mathbf{0}) + (-c\mathbf{0})$$

which implies

$$\begin{aligned} \mathbf{0} &= c\mathbf{0} + (c\mathbf{0} + (-c\mathbf{0})) && \text{by axioms 5 and 3} \\ &= c\mathbf{0} + \mathbf{0} && \text{by axiom 5} \\ &= c\mathbf{0} && \text{by axiom 4} \end{aligned}$$

(d) Suppose $c\mathbf{u} = \mathbf{0}$. To show that either $c = 0$ or $\mathbf{u} = \mathbf{0}$, let's assume that $c \neq 0$. (If $c = 0$, there is nothing to prove.) Then, since $c \neq 0$, its reciprocal $1/c$ is defined, and

$$\begin{aligned} \mathbf{u} &= 1\mathbf{u} && \text{by axiom 10} \\ &= \left(\frac{1}{c}c\right)\mathbf{u} \\ &= \frac{1}{c}(c\mathbf{u}) && \text{by axiom 9} \\ &= \frac{1}{c}\mathbf{0} \\ &= \mathbf{0} && \text{by property (b)} \end{aligned}$$

We will write $\mathbf{u} - \mathbf{v}$ for $\mathbf{u} + (-\mathbf{v})$, thereby defining ***subtraction*** of vectors. We will also exploit the associativity property of addition to unambiguously write $\mathbf{u} + \mathbf{v} + \mathbf{w}$ for the sum of three vectors and, more generally,

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$$

for a ***linear combination*** of vectors.

Subspaces

We have seen that, in $\mathbb{R}^n$, it is possible for one vector space to sit inside another one, giving rise to the notion of a subspace. For example, a plane through the origin is a subspace of $\mathbb{R}^3$. We now extend this concept to general vector spaces.

Definition A subset W of a vector space V is called a ***subspace*** of V if W is itself a vector space with the same scalars, addition, and scalar multiplication as V.

As in $\mathbb{R}^n$, checking to see whether a subset W of a vector space V is a subspace of V involves testing only two of the ten vector space axioms. We prove this observation as a theorem.

Theorem 6.2 Let V be a vector space and let W be a nonempty subset of V. Then W is a subspace of V if and only if the following conditions hold:

a. If $\mathbf{u}$ and $\mathbf{v}$ are in W, then $\mathbf{u} + \mathbf{v}$ is in W.
b. If $\mathbf{u}$ is in W and c is a scalar, then $c\mathbf{u}$ is in W.

Proof Assume that W is a subspace of V. Then W satisfies vector space axioms 1 to 10. In particular, axiom 1 is condition (a) and axiom 6 is condition (b).

Conversely, assume that W is a subset of a vector space V, satisfying conditions (a) and (b). By hypothesis, axioms 1 and 6 hold. Axioms 2, 3, 7, 8, 9, and 10 hold in W because they are true for *all* vectors in V and thus are true in particular for those vectors in W. (We say that W *inherits* these properties from V.) This leaves axioms 4 and 5 to be checked.

Since W is nonempty, it contains at least one vector $\mathbf{u}$. Then condition (b) and Theorem 6.1(a) imply that $0\mathbf{u} = \mathbf{0}$ is also in W. This is axiom 4.

If $\mathbf{u}$ is in V, then, by taking $c = -1$ in condition (b), we have that $-\mathbf{u} = (-1)\mathbf{u}$ is also in W, using Theorem 6.1(c).

Remark Since Theorem 6.2 generalizes the notion of a subspace from the context of $\mathbb{R}^n$ to general vector spaces, all of the subspaces of $\mathbb{R}^n$ that we encountered in Chapter 3 are subspaces of $\mathbb{R}^n$ in the current context. In particular, lines and planes through the origin are subspaces of $\mathbb{R}^3$.

Example 6.9 We have already shown that the set $\mathscr{P}_n$ of all polynomials with degree at most n is a vector space. Hence, $\mathscr{P}_n$ is a subspace of the vector space $\mathscr{P}$ of *all* polynomials.

Example 6.10 Let W be the set of symmetric $n \times n$ matrices. Show that W is a subspace of M_{nn}.

Solution Clearly, W is nonempty, so we need only check conditions (a) and (b) in Theorem 6.2. Let A and B be in W and let c be a scalar. Then $A^T = A$ and $B^T = B$, from which it follows that

$$(A + B)^T = A^T + B^T = A + B$$

Therefore, $A + B$ is symmetric and, hence, is in W. Similarly,

$$(cA)^T = cA^T = cA$$

so cA is symmetric and, thus, is in W. We have shown that W is closed under addition and scalar multiplication. Therefore, it is a subspace of M_{nn}, by Theorem 6.2.

Example 6.11

Let $\mathscr{C}$ be the set of all continuous real-valued functions defined on $\mathbb{R}$ and let $\mathscr{D}$ be the set of all differentiable real-valued functions defined on $\mathbb{R}$. Show that $\mathscr{C}$ and $\mathscr{D}$ are subspaces of $\mathscr{F}$, the vector space of all real-valued functions defined on $\mathbb{R}$.

Solution From calculus, we know that if f and g are continuous functions and c is a scalar, then $f + g$ and cf are also continuous. Hence, $\mathscr{C}$ is closed under addition and scalar multiplication and so is a subspace of $\mathscr{F}$. If f and g are differentiable, then so are $f + g$ and cf. Indeed,

$$(f + g)' = f' + g' \quad \text{and} \quad (cf)' = c(f')$$

So $\mathscr{D}$ is also closed under addition and scalar multiplication, making it a subspace of $\mathscr{F}$.

It is a theorem of calculus that every differentiable function is continuous. Consequently, $\mathscr{D}$ is contained in $\mathscr{C}$ (denoted by $\mathscr{D} \subset \mathscr{C}$), making $\mathscr{D}$ a subspace of $\mathscr{C}$. It is also the case that every polynomial function is differentiable, so $\mathscr{P} \subset \mathscr{D}$, and thus $\mathscr{P}$ is a subspace of $\mathscr{D}$. We therefore have a *hierarchy* of subspaces of $\mathscr{F}$, one inside the other:

$$\mathscr{P} \subset \mathscr{D} \subset \mathscr{C} \subset \mathscr{F}$$

This hierarchy is depicted in Figure 6.2.

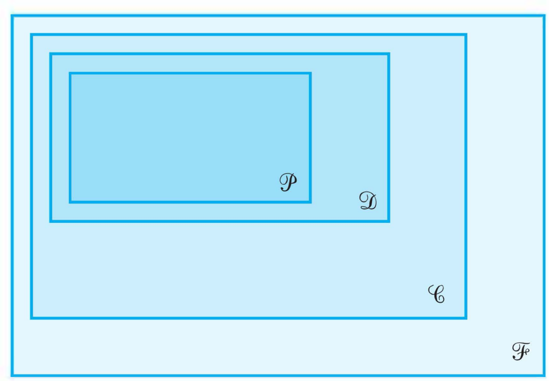


Figure 6.2
The hierarchy of subspaces of $\mathscr{F}$

There are other subspaces of $\mathscr{F}$ that can be placed into this hierarchy. Some of these are explored in the exercises.

In the preceding discussion, we could have restricted our attention to functions defined on a closed interval $[a, b]$. Then the corresponding subspaces of $\mathscr{F}[a, b]$ would be $\mathscr{C}[a, b]$, $\mathscr{D}[a, b]$, and $\mathscr{P}[a, b]$.

Example 6.12

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

$$f'' + f = 0 \tag{1}$$

That is, S is the solution set of Equation (1). Show that S is a subspace of $\mathscr{F}$.

Solution S is nonempty, since the zero function clearly satisfies Equation (1). Let f and g be in S and let c be a scalar. Then

$$\begin{aligned}(f+g)'' + (f+g) &= (f'' + g'') + (f+g)\\ &= (f''+f) + (g''+g)\\ &= 0 + 0\\ &= 0\end{aligned}$$

which shows that $f+g$ is in S. Similarly,

$$\begin{aligned}(cf)'' + cf &= cf'' + cf\\ &= c(f''+f)\\ &= c0\\ &= 0\end{aligned}$$

so cf is also in S.

Therefore, S is closed under addition and scalar multiplication and is a subspace of $\mathscr{F}$.

The differential Equation (1) is an example of a ***homogeneous linear differential equation.*** The solution sets of such equations are always subspaces of $\mathscr{F}$. Note that in Example 6.12 we did not actually solve Equation (1) (i.e., we did not find any specific solutions, other than the zero function). We will discuss techniques for finding solutions to this type of equation in Section 6.7.

As you gain experience working with vector spaces and subspaces, you will notice that certain examples tend to resemble one another. For example, consider the vector spaces $\mathbb{R}^4$, $\mathscr{P}_3$, and M_{22}. Typical elements of these vector spaces are, respectively,

$$\mathbf{u} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \quad p(x) = a + bx + cx^2 + dx^3, \quad \text{and} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

In the words of Yogi Berra, "It's déjà vu all over again."

Any calculations involving the vector space operations of addition and scalar multiplication are essentially the same in all three settings. To highlight the similarities, in the next example we will perform the necessary steps in the three vector spaces side by side.

Example 6.13

(a) Show that the set W of all vectors of the form

$$\begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix}$$

is a subspace of $\mathbb{R}^4$.

(b) Show that the set W of all polynomials of the form $a + bx - bx^2 + ax^3$ is a subspace of $\mathscr{P}_3$.

(c) Show that the set W of all matrices of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is a subspace of M_{22}.

Solution

(a) W is nonempty because it contains the zero vector $\mathbf{0}$. (Take $a = b = 0$.) Let $\mathbf{u}$ and $\mathbf{v}$ be in W—say,

$$\mathbf{u} = \begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} c \\ d \\ -d \\ c \end{bmatrix}$$

Then

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} a + c \\ b + d \\ -b - d \\ a + c \end{bmatrix} = \begin{bmatrix} a + c \\ b + d \\ -(b + d) \\ a + c \end{bmatrix}$$

so $\mathbf{u} + \mathbf{v}$ is also in W (because it has the right *form*).

Similarly, if k is a scalar, then

$$k\mathbf{u} = \begin{bmatrix} ka \\ kb \\ -kb \\ ka \end{bmatrix}$$

so $k\mathbf{u}$ is in W.

Thus, W is a nonempty subset of $\mathbb{R}^4$ that is closed under addition and scalar multiplication. Therefore, W is a subspace of $\mathbb{R}^4$, by Theorem 6.2.

(b) W is nonempty because it contains the zero polynomial. (Take $a = b = 0$.) Let $p(x)$ and $q(x)$ be in W—say,

$$p(x) = a + bx - bx^2 + ax^3$$

and

$$q(x) = c + dx - dx^2 + cx^3$$

Then

$$p(x) + q(x) = (a + c) + (b + d)x - (b + d)x^2 + (a + c)x^3$$

so $p(x) + q(x)$ is also in W (because it has the right *form*).

Similarly, if k is a scalar, then

$$kp(x) = ka + kbx - kbx^2 + kax^3$$

so $kp(x)$ is in W.

Thus, W is a nonempty subset of $\mathcal{P}_3$ that is closed under addition and scalar multiplication. Therefore, W is a subspace of $\mathcal{P}_3$ by Theorem 6.2.

(c) W is nonempty because it contains the zero matrix O. (Take $a = b = 0$.) Let A and B be in W—say,

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

and

$$B = \begin{bmatrix} c & d \\ -d & c \end{bmatrix}$$

Then

$$A + B = \begin{bmatrix} a + c & b + d \\ -(b + d) & a + c \end{bmatrix}$$

so $A + B$ is also in W (because it has the right *form*).

Similarly, if k is a scalar, then

$$kA = \begin{bmatrix} ka & kb \\ -kb & ka \end{bmatrix}$$

so kA is in W.

Thus, W is a nonempty subset of M_{22} that is closed under addition and scalar multiplication. Therefore, W is a subspace of M_{22}, by Theorem 6.2.

Example 6.13 shows that it is often possible to relate examples that, on the surface, appear to have nothing in common. Consequently, we can apply our knowledge of $\mathbb{R}^n$ to polynomials, matrices, and other examples. We will encounter this idea several times in this chapter and will make it precise in Section 6.5.

Example 6.14

If V is a vector space, then V is clearly a subspace of itself. The set $\{\mathbf{0}\}$, consisting of only the zero vector, is also a subspace of V, called the ***zero subspace.*** To show this, we simply note that the two closure conditions of Theorem 6.2 are satisfied:

$$\mathbf{0} + \mathbf{0} = \mathbf{0} \quad \text{and} \quad c\mathbf{0} = \mathbf{0} \quad \text{for any scalar } c$$

The subspaces $\{\mathbf{0}\}$ and V are called the ***trivial subspaces*** of V.

An examination of the proof of Theorem 6.2 reveals the following useful fact:

If W is a subspace of a vector space V, then W contains the zero vector $\mathbf{0}$ of V.

This fact is consistent with, and analogous to, the fact that lines and planes are subspaces of $\mathbb{R}^3$ if and only if they contain the origin. The requirement that every subspace must contain $\mathbf{0}$ is sometimes useful in showing that a set is *not* a subspace.

Example 6.15 Let W be the set of all 2×2 matrices of the form

$$\begin{bmatrix} a & a+1 \\ 0 & b \end{bmatrix}$$

Is W a subspace of M_{22}?

Solution Each matrix in W has the property that its $(1, 2)$ entry is one more than its $(1, 1)$ entry. Since the zero matrix

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

does not have this property, it is not in W. Hence, W is not a subspace of M_{22}.

Example 6.16 Let W be the set of all 2×2 matrices with determinant equal to 0. Is W a subspace of M_{22}? (Since $\det O = 0$, the zero matrix is in W, so the method of Example 6.15 is of no use to us.)

Solution Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then $\det A = \det B = 0$, so A and B are in W. But

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so $\det(A + B) = 1 \neq 0$, and therefore $A + B$ is not in W. Thus, W is not closed under addition and so is not a subspace of M_{22}.

Spanning Sets

The notion of a spanning set of vectors carries over easily from $\mathbb{R}^n$ to general vector spaces.

Definition If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a set of vectors in a vector space V, then the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called the ***span*** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ and is denoted by $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ or $\text{span}(S)$. If $V = \text{span}(S)$, then S is called a ***spanning set*** for V and V is said to be ***spanned*** by S.

Example 6.17

Show that the polynomials 1, x, and x^2 span $\mathscr{P}_2$.

Solution By its very definition, a polynomial $p(x) = a + bx + cx^2$ is a linear combination of 1, x, and x^2. Therefore, $\mathscr{P}_2 = \text{span}(1, x, x^2)$.

Example 6.17 can clearly be generalized to show that $\mathscr{P}_n = \text{span}(1, x, x^2, \ldots, x^n)$. However, no finite set of polynomials can possibly span $\mathscr{P}$, the vector space of all polynomials. (See Exercise 44 in Section 6.2.) But, if we allow a spanning set to be infinite, then clearly the set of *all* nonnegative powers of x will do. That is, $\mathscr{P} = \text{span}(1, x, x^2, \ldots)$.

Example 6.18

Show that $M_{23} = \text{span}(E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23})$, where

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(That is, E_{ij} is the matrix with a 1 in row i, column j and zeros elsewhere.)

Solution We need only observe that

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = a_{11}E_{11} + a_{12}E_{12} + a_{13}E_{13} + a_{21}E_{21} + a_{22}E_{22} + a_{23}E_{23}$$

Extending this example, we see that, in general, M_{mn} is spanned by the mn matrices E_{ij}, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Example 6.19

In $\mathscr{P}_2$, determine whether $r(x) = 1 - 4x + 6x^2$ is in $\text{span}(p(x), q(x))$, where

$$p(x) = 1 - x + x^2 \quad \text{and} \quad q(x) = 2 + x - 3x^2$$

Solution We are looking for scalars c and d such that $cp(x) + dq(x) = r(x)$. This means that

$$c(1 - x + x^2) + d(2 + x - 3x^2) = 1 - 4x + 6x^2$$

Regrouping according powers of x, we have

$$(c + 2d) + (-c + d)x + (c - 3d)x^2 = 1 - 4x + 6x^2$$

Equating the coefficients of like powers of x gives

$$\begin{aligned} c + 2d &= 1 \\ -c + d &= -4 \\ c - 3d &= 6 \end{aligned}$$

which is easily solved to give $c = 3$ and $d = -1$. Therefore, $r(x) = 3p(x) - q(x)$, so $r(x)$ is in $\text{span}(p(x), q(x))$. (Check this.)

Example 6.20

In $\mathscr{F}$, determine whether $\sin 2x$ is in $\text{span}(\sin x, \cos x)$.

Solution We set $c \sin x + d \cos x = \sin 2x$ and try to determine c and d so that this equation is true. Since these are functions, the equation must be true for *all* values of x. Setting $x = 0$, we have

$$c \sin 0 + d \cos 0 = \sin 0 \quad \text{or} \quad c(0) + d(1) = 0$$

from which we see that $d = 0$. Setting $x = \pi/2$, we get

$$c \sin(\pi/2) + d \cos(\pi/2) = \sin(\pi) \quad \text{or} \quad c(1) + d(0) = 0$$

giving $c = 0$. But this implies that $\sin 2x = 0(\sin x) + 0(\cos x) = 0$ for all x, which is absurd, since $\sin 2x$ is not the zero function. We conclude that $\sin 2x$ is not in $\text{span}(\sin x, \cos x)$.

Remark It *is* true that $\sin 2x$ can be written in terms of $\sin x$ and $\cos x$. For example, we have the double angle formula $\sin 2x = 2 \sin x \cos x$. However, this is not a *linear* combination.

Example 6.21

In M_{22}, describe the span of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Solution Every linear combination of A, B, and C is of the form

$$\begin{aligned} cA + dB + eC &= c\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + e\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} c + d & c + e \\ c + e & d \end{bmatrix} \end{aligned}$$

This matrix is symmetric, so $\text{span}(A, B, C)$ is contained within the subspace of symmetric 2×2 matrices. In fact, we have equality; that is, *every* symmetric 2×2 matrix is in $\text{span}(A, B, C)$. To show this, we let $\begin{bmatrix} x & y \\ y & z \end{bmatrix}$ be a symmetric 2×2 matrix. Setting

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} = \begin{bmatrix} c + d & c + e \\ c + e & d \end{bmatrix}$$

and solving for c and d, we find that $c = x - z$, $d = z$, and $e = -x + y + z$. Therefore,

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} = (x - z)\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + z\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (-x + y + z)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(Check this.) It follows that $\text{span}(A, B, C)$ is the subspace of symmetric 2×2 matrices.

As was the case in $\mathbb{R}^n$, the span of a set of vectors is always a subspace of the vector space that contains them. The next theorem makes this result precise. It generalizes Theorem 3.19.

Theorem 6.3 Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be vectors in a vector space V.

a. $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ is a subspace of V.
b. $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ is the smallest subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.

Proof (a) The proof of property (a) is identical to the proof of Theorem 3.19, with $\mathbb{R}^n$ replaced by V.

(b) To establish property (b), we need to show that any subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ also contains $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$. Accordingly, let W be a subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Then, since W is closed under addition and scalar multiplication, it contains every linear combination $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Therefore, $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ is contained in W.

Exercises 6.1

In Exercises 1–11, determine whether the given set, together with the specified operations of addition and scalar multiplication, is a vector space. If it is not, list all of the axioms that fail to hold.

1. The set of all vectors in $\mathbb{R}^2$ of the form $\begin{bmatrix} x \\ x \end{bmatrix}$, with the usual vector addition and scalar multiplication

2. The set of all vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ with $x \geq 0, y \geq 0$ (i.e., the first quadrant), with the usual vector addition and scalar multiplication

3. The set of all vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ with $xy \geq 0$ (i.e., the union of the first and third quadrants), with the usual vector addition and scalar multiplication

4. The set of all vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ with $x \geq y$, with the usual vector addition and scalar multiplication

5. $\mathbb{R}^2$, with the usual addition but scalar multiplication defined by

$$c\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ y \end{bmatrix}$$

6. $\mathbb{R}^2$, with the usual scalar multiplication but addition defined by

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 1 \\ y_1 + y_2 + 1 \end{bmatrix}$$

7. The set of all positive real numbers, with addition $\oplus$ defined by $x \oplus y = xy$ and scalar multiplication $\odot$ defined by $c \odot x = x^c$

8. The set of all rational numbers, with the usual addition and multiplication

9. The set of all upper triangular 2×2 matrices, with the usual matrix addition and scalar multiplication

10. The set of all 2×2 matrices of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where $ad = 0$, with the usual matrix addition and scalar multiplication

11. The set of all skew-symmetric $n \times n$ matrices, with the usual matrix addition and scalar multiplication (see page 162).

12. Finish verifying that $\mathscr{P}_2$ is a vector space (see Example 6.3).

13. Finish verifying that $\mathscr{F}$ is a vector space (see Example 6.4).

In Exercises 14–17, determine whether the given set, together with the specified operations of addition and scalar multiplication, is a complex vector space. If it is not, list all of the axioms that fail to hold.

14. The set of all vectors in $\mathbb{C}^2$ of the form $\begin{bmatrix} z \\ \bar{z} \end{bmatrix}$, with the usual vector addition and scalar multiplication

15. The set $M_{mn}(\mathbb{C})$ of all $m \times n$ complex matrices, with the usual matrix addition and scalar multiplication

16. The set $\mathbb{C}^2$, with the usual vector addition but scalar multiplication defined by $c\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \bar{c}z_1 \\ \bar{c}z_2 \end{bmatrix}$

17. $\mathbb{R}^n$, with the usual vector addition and scalar multiplication

In Exercises 18–21, determine whether the given set, together with the specified operations of addition and scalar multiplication, is a vector space over the indicated $\mathbb{Z}_p$. If it is not, list all of the axioms that fail to hold.

18. The set of all vectors in $\mathbb{Z}_2^n$ with an even number of 1s, over $\mathbb{Z}_2$ with the usual vector addition and scalar multiplication

19. The set of all vectors in $\mathbb{Z}_2^n$ with an odd number of 1s, over $\mathbb{Z}_2$ with the usual vector addition and scalar multiplication

20. The set $M_{mn}(\mathbb{Z}_p)$ of all $m \times n$ matrices with entries from $\mathbb{Z}_p$, over $\mathbb{Z}_p$ with the usual matrix addition and scalar multiplication

21. $\mathbb{Z}_6$, over $\mathbb{Z}_3$ with the usual addition and multiplication (Think this one through carefully!)

22. Prove Theorem 6.1(a).

23. Prove Theorem 6.1(c).

In Exercises 24–45, use Theorem 6.2 to determine whether W is a subspace of V.

24. $V = \mathbb{R}^3, W = \left\{\begin{bmatrix} a \\ 0 \\ a \end{bmatrix}\right\}$

25. $V = \mathbb{R}^3, W = \left\{\begin{bmatrix} a \\ -a \\ 2a \end{bmatrix}\right\}$

26. $V = \mathbb{R}^3, W = \left\{\begin{bmatrix} a \\ b \\ a+b+1 \end{bmatrix}\right\}$

27. $V = \mathbb{R}^3, W = \left\{\begin{bmatrix} a \\ b \\ |a| \end{bmatrix}\right\}$

28. $V = M_{22}, W = \left\{\begin{bmatrix} a & b \\ b & 2a \end{bmatrix}\right\}$

29. $V = M_{22}, W = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} : ad \geq bc\right\}$

30. $V = M_{nn}, W = \{A \text{ in } M_{nn} : \det A = 1\}$

31. $V = M_{nn}$, W is the set of diagonal $n \times n$ matrices

32. $V = M_{nn}$, W is the set of idempotent $n \times n$ matrices

33. $V = M_{nn}, W = \{A \text{ in } M_{nn} : AB = BA\}$, where B is a given (fixed) matrix

34. $V = \mathscr{P}_2, W = \{bx + cx^2\}$

35. $V = \mathscr{P}_2, W = \{a + bx + cx^2 : a + b + c = 0\}$

36. $V = \mathscr{P}_2, W = \{a + bx + cx^2 : abc = 0\}$

37. $V = \mathscr{P}$, W is the set of all polynomials of degree 3

38. $V = \mathscr{F}, W = \{f \text{ in } \mathscr{F} : f(-x) = f(x)\}$

39. $V = \mathscr{F}, W = \{f \text{ in } \mathscr{F} : f(-x) = -f(x)\}$

40. $V = \mathscr{F}, W = \{f \text{ in } \mathscr{F} : f(0) = 1\}$

41. $V = \mathscr{F}, W = \{f \text{ in } \mathscr{F} : f(0) = 0\}$

42. $V = \mathscr{F}$, W is the set of all integrable functions

43. $V = \mathscr{D}, W = \{f \text{ in } \mathscr{D} : f'(x) \geq 0 \text{ for all } x\}$

44. $V = \mathscr{F}, W = \mathscr{C}^{(2)}$, the set of all functions with continuous second derivatives

45. $V = \mathscr{F}, W = \{f \text{ in } \mathscr{F} : = \lim_{x \to 0} f(x) = \infty\}$

46. Let V be a vector space with subspaces U and W. Prove that $U \cap W$ is a subspace of V.

47. Let V be a vector space with subspaces U and W. Give an example with $V = \mathbb{R}^2$ to show that $U \cup W$ need not be a subspace of V.

48. Let V be a vector space with subspaces U and W. Define the ***sum of U and W*** to be

$$U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \text{ is in } U, \mathbf{w} \text{ is in } W\}$$

(a) If $V = \mathbb{R}^3$, U is the x-axis, and W is the y-axis, what is $U + W$?

(b) If U and W are subspaces of a vector space V, prove that $U + W$ is a subspace of V.

49. If U and V are vector spaces, define the ***Cartesian product*** of U and V to be

$$U \times V = \{(\mathbf{u}, \mathbf{v}) : \mathbf{u} \text{ is in } U \text{ and } \mathbf{v} \text{ is in } V\}$$

Prove that $U \times V$ is a vector space.

50. Let W be a subspace of a vector space V. Prove that $\Delta = \{(\mathbf{w}, \mathbf{w}) : \mathbf{w} \text{ is in } W\}$ is a subspace of $V \times V$.

In Exercises 51 and 52, let $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. Determine whether C is in span(A, B).

51. $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

52. $C = \begin{bmatrix} 3 & -5 \\ 5 & -1 \end{bmatrix}$

In Exercises 53 and 54, let $p(x) = 1 - 2x$, $q(x) = x - x^2$, *and* $r(x) = -2 + 3x + x^2$. *Determine whether* $s(x)$ *is in* span$(p(x), q(x), r(x))$.

53. $s(x) = 3 - 5x - x^2$

54. $s(x) = 1 + x + x^2$

In Exercises 55–58, let $f(x) = \sin^2 x$ *and* $g(x) = \cos^2 x$. *Determine whether* $h(x)$ *is in* span$(f(x), g(x))$.

55. $h(x) = 1$

56. $h(x) = \cos 2x$

57. $h(x) = \sin 2x$

58. $h(x) = \sin x$

59. Is M_{22} spanned by $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

60. Is M_{22} spanned by $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

61. Is $\mathscr{P}_2$ spanned by $1 + x, x + x^2, 1 + x^2$?

62. Is $\mathscr{P}_2$ spanned by $1 + x + 2x^2, 2 + x + 2x^2, -1 + x + 2x^2$?

63. Prove that every vector space has a unique zero vector.

64. Prove that for every vector $\mathbf{v}$ in a vector space V, there is a unique $\mathbf{v}'$ in V such that $\mathbf{v} + \mathbf{v}' = \mathbf{0}$.

Writing Project

The Rise of Vector Spaces

As noted in the sidebar on page 429, in the late 19th century, the mathematicians Hermann Grassmann and Giuseppe Peano were instrumental in introducing the idea of a vector space and the vector space axioms that we use today. Grassmann's work had its origins in barycentric coordinates, a technique invented in 1827 by August Ferdinand Möbius (of Möbius strip fame). However, widespread acceptance of the vector space concept did not come until the early 20th century.

Write a report on the history of vector spaces. Discuss the origins of the notion of a vector space and the contributions of Grassmann and Peano. Why was the mathematical community slow to adopt these ideas, and how did acceptance come about?

1. Carl B. Boyer and Uta C. Merzbach, *A History of Mathematics* (Third Edition) (Hoboken, NJ: Wiley, 2011).
2. Jean-Luc, Dorier (1995), A General Outline of the Genesis of Vector Space Theory, *Historia Mathematica* 22 (1995), pp. 227–261.
3. Victor J. Katz, *A History of Mathematics: An Introduction* (Third Edition) (Reading, MA: Addison Wesley Longman, 2008).

6.2 Linear Independence, Basis, and Dimension

In this section, we extend the notions of linear independence, basis, and dimension to general vector spaces, generalizing the results of Sections 2.3 and 3.5. In most cases, the proofs of the theorems carry over; we simply replace $\mathbb{R}^n$ by the vector space V.

Linear Independence

Definition A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in a vector space V is ***linearly dependent*** if there are scalars $c_1, c_2, \ldots, c_k$, *at least one of which is not zero*, such that

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

A set of vectors that is not linearly dependent is said to be ***linearly independent***.

As in $\mathbb{R}^n$, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is linearly independent in a vector space V if and only if

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

We also have the following useful alternative formulation of linear dependence.

Theorem 6.4 A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in a vector space V is linearly dependent if and only if at least one of the vectors can be expressed as a linear combination of the others.

Proof The proof is identical to that of Theorem 2.5.

As a special case of Theorem 6.4, note that a set of *two* vectors is linearly dependent if and only if one is a scalar multiple of the other.

Example 6.22 In $\mathscr{P}_2$, the set $\{1 + x + x^2, 1 - x + 3x^2, 1 + 3x - x^2\}$ is linearly dependent, since

$$2(1 + x + x^2) - (1 - x + 3x^2) = 1 + 3x - x^2$$

Example 6.23 In M_{22}, let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

Then $A + B = C$, so the set $\{A, B, C\}$ is linearly dependent.

Example 6.24 In $\mathscr{F}$, the set $\{\sin^2 x, \cos^2 x, \cos 2x\}$ is linearly dependent, since

$$\cos 2x = \cos^2 x - \sin^2 x$$

Example 6.25 Show that the set $\{1, x, x^2, \dots, x^n\}$ is linearly independent in $\mathscr{P}_n$.

Solution 1 Suppose that $c_0, c_1, \dots, c_n$ are scalars such that

$$c_0 \cdot 1 + c_1x + c_2x^2 + \cdots + c_nx^n = 0$$

Then the polynomial $p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ is zero for all values of x. But a polynomial of degree at most n cannot have more than n zeros (see Appendix D). So $p(x)$ must be the zero polynomial, meaning that $c_0 = c_1 = c_2 = \cdots = c_n = 0$. Therefore, $\{1, x, x^2, \dots, x^n\}$ is linearly independent.

Solution 2 We begin, as in the first solution, by assuming that

$$p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n = 0$$

Since this is true for all x, we can substitute $x = 0$ to obtain $c_0 = 0$. This leaves

$$c_1x + c_2x^2 + \cdots + c_nx^n = 0$$

Taking derivatives, we obtain

$$c_1 + 2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1} = 0$$

and setting $x = 0$, we see that $c_1 = 0$. Differentiating $2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1} = 0$ and setting $x = 0$, we find that $2c_2 = 0$, so $c_2 = 0$. Continuing in this fashion, we find that $k!c_k = 0$ for $k = 0, \ldots, n$. Therefore, $c_0 = c_1 = c_2 = \cdots = c_n = 0$, and $\{1, x, x^2, \ldots, x^n\}$ is linearly independent.

Example 6.26

In $\mathscr{P}_2$, determine whether the set $\{1 + x, x + x^2, 1 + x^2\}$ is linearly independent.

Solution Let c_1, c_2, and c_3 be scalars such that

$$c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2) = 0$$

Then

$$(c_1 + c_3) + (c_1 + c_2)x + (c_2 + c_3)x^2 = 0$$

This implies that

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

the solution to which is $c_1 = c_2 = c_3 = 0$. It follows that $\{1 + x, x + x^2, 1 + x^2\}$ is linearly independent.

Remark Compare Example 6.26 with Example 2.23(b). The system of equations that arises is exactly the same. This is because of the correspondence between $\mathscr{P}_2$ and $\mathbb{R}^3$ that relates

$$1 + x \leftrightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad x + x^2 \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad 1 + x^2 \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

and produces the columns of the coefficient matrix of the linear system that we have to solve. Thus, showing that $\{1 + x, x + x^2, 1 + x^2\}$ is linearly independent is equivalent to showing that

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

is linearly independent. This can be done simply by establishing that the matrix

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

has rank 3, by the Fundamental Theorem of Invertible Matrices.

Example 6.27

In $\mathscr{F}$, determine whether the set $\{\sin x, \cos x\}$ is linearly independent.

Solution The functions $f(x) = \sin x$ and $g(x) = \cos x$ are linearly *dependent* if and only if one of them is a scalar multiple of the other. But it is clear from their graphs that this is not the case, since, for example, any nonzero multiple of $f(x) = \sin x$ has the same zeros, none of which are zeros of $g(x) = \cos x$.

This approach may not always be appropriate to use, so we offer the following direct, more computational method. Suppose c and d are scalars such that

$$c \sin x + d \cos x = 0$$

Setting $x = 0$, we obtain $d = 0$, and setting $x = \pi/2$, we obtain $c = 0$. Therefore, the set $\{\sin x, \cos x\}$ is linearly independent.

Although the definitions of linear dependence and independence are phrased in terms of *finite* sets of vectors, we can extend the concepts to *infinite* sets as follows:

A set S of vectors in a vector space V is ***linearly dependent*** if it contains finitely many linearly dependent vectors. A set of vectors that is not linearly dependent is said to be ***linearly independent.***

Note that for finite sets of vectors, this is just the original definition. Following is an example of an infinite set of linearly independent vectors.

Example 6.28

In $\mathscr{P}$, show that $S = \{1, x, x^2, \ldots\}$ is linearly independent.

Solution Suppose there is a finite subset T of S that is linearly dependent. Let x^m be the highest power of x in T and let x^n be the lowest power of x in T. Then there are scalars $c_n, c_{n+1}, \ldots, c_m$, not all zero, such that

$$c_n x^n + c_{n+1} x^{n+1} + \cdots + c_m x^m = 0$$

But, by an argument similar to that used in Example 6.25, this implies that $c_n = c_{n+1} = \cdots = c_m = 0$, which is a contradiction. Hence, S cannot contain finitely many linearly dependent vectors, so it is linearly independent.

Bases

The important concept of a basis now can be extended easily to arbitrary vector spaces.

Definition A subset $\mathcal{B}$ of a vector space V is a ***basis*** for V if

1. $\mathcal{B}$ spans V and
2. $\mathcal{B}$ is linearly independent.

Example 6.29

If $\mathbf{e}_i$ is the ith column of the $n \times n$ identity matrix, then $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$, called the ***standard basis*** for $\mathbb{R}^n$.

Example 6.30

$\{1, x, x^2, \ldots, x^n\}$ is a basis for $\mathcal{P}_n$, called the ***standard basis*** for $\mathcal{P}_n$.

Example 6.31

The set $\mathcal{E} = \{E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, E_{m1}, \ldots, E_{mn}\}$ is a basis for M_{mn}, where the matrices E_{ij} are as defined in Example 6.18. $\mathcal{E}$ is called the ***standard basis*** for M_{mn}.

We have already seen that $\mathcal{E}$ spans M_{mn}. It is easy to show that $\mathcal{E}$ is linearly independent. (Verify this!) Hence, $\mathcal{E}$ is a basis for M_{mn}.

Example 6.32

Show that $\mathcal{B} = \{1 + x, x + x^2, 1 + x^2\}$ is a basis for $\mathcal{P}_2$.

Solution We have already shown that $\mathcal{B}$ is linearly independent, in Example 6.26. To show that $\mathcal{B}$ spans $\mathcal{P}_2$, let $a + bx + cx^2$ be an arbitrary polynomial in $\mathcal{P}_2$. We must show that there are scalars c_1, c_2, and c_3 such that

$$c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2) = a + bx + cx^2$$

or, equivalently,

$$(c_1 + c_3) + (c_1 + c_2)x + (c_2 + c_3)x^2 = a + bx + cx^2$$

Equating coefficients of like powers of x, we obtain the linear system

$$\begin{aligned} c_1 + \quad\; c_3 &= a \\ c_1 + c_2 \quad\;\; &= b \\ c_2 + c_3 &= c \end{aligned}$$

which has a solution, since the coefficient matrix $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ has rank 3 and, hence, is invertible. (We do not need to know *what* the solution is; we only need to know that it exists.) Therefore, $\mathcal{B}$ is a basis for $\mathcal{P}_2$.

Remark Observe that the matrix $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ is the key to Example 6.32. We can immediately obtain it using the correspondence between $\mathcal{P}_2$ and $\mathbb{R}^3$, as indicated in the Remark following Example 6.26.

Example 6.33

Show that $\mathcal{B} = \{1, x, x^2, \ldots\}$ is a basis for $\mathcal{P}$.

Solution In Example 6.28, we saw that $\mathcal{B}$ is linearly independent. It also spans $\mathcal{P}$, since clearly every polynomial is a linear combination of (finitely many) powers of x.

Example 6.34

Find bases for the three vector spaces in Example 6.13:

(a) $W_1 = \left\{ \begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix} \right\}$ (b) $W_2 = \{a + bx - bx^2 + ax^3\}$ (c) $W_3 = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \right\}$

Solution Once again, we will work the three examples side by side to highlight the similarities among them. In a strong sense, they are all the *same* example, but it will take us until Section 6.5 to make this idea perfectly precise.

(a) Since

$$\begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

we have $W_1 = \text{span}(\mathbf{u}, \mathbf{v})$, where

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

Since $\{\mathbf{u}, \mathbf{v}\}$ is clearly linearly independent, it is also a basis for W_1.

(b) Since

$$a + bx - bx^2 + ax^3 = a(1 + x^3) + b(x - x^2)$$

we have $W_2 = \text{span}(u(x), v(x))$, where

$$u(x) = 1 + x^3$$

and $$v(x) = x - x^2$$

Since $\{u(x), v(x)\}$ is clearly linearly independent, it is also a basis for W_2.

(c) Since

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

we have $W_3 = \text{span}(U, V)$, where

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

Since $\{U, V\}$ is clearly linearly independent, it is also a basis for W_3.

Coordinates

Section 3.5 introduced the idea of the coordinates of a vector with respect to a basis for subspaces of $\mathbb{R}^n$. We now extend this concept to arbitrary vector spaces.

Theorem 6.5

Let V be a vector space and let $\mathcal{B}$ be a basis for V. For every vector $\mathbf{v}$ in V, there is exactly one way to write $\mathbf{v}$ as a linear combination of the basis vectors in $\mathcal{B}$.

Proof The proof is the same as the proof of Theorem 3.29. It works even if the basis $\mathcal{B}$ is infinite, since linear combinations are, by definition, finite.

The converse of Theorem 6.5 is also true. That is, if $\mathcal{B}$ is a set of vectors in a vector space V with the property that every vector in V can be written uniquely as a linear combination of the vectors in $\mathcal{B}$, then $\mathcal{B}$ is a basis for V (see Exercise 30). In this sense, the *unique representation property* characterizes a basis.

Since representation of a vector with respect to a basis is unique, the next definition makes sense.

Definition Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V. Let $\mathbf{v}$ be a vector in V, and write $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. Then $c_1, c_2, \ldots, c_n$ are called the ***coordinates of* v *with respect to* $\mathcal{B}$,** and the column vector

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

is called the ***coordinate vector of* v *with respect to* $\mathcal{B}$.**

Observe that if the basis $\mathcal{B}$ of V has n vectors, then $[\mathbf{v}]_{\mathcal{B}}$ is a (column) vector in $\mathbb{R}^n$.

Example 6.35

Find the coordinate vector $[p(x)]_{\mathcal{B}}$ of $p(x) = 2 - 3x + 5x^2$ with respect to the standard basis $\mathcal{B} = \{1, x, x^2\}$ of $\mathcal{P}_2$.

Solution The polynomial $p(x)$ is already a linear combination of 1, x, and x^2, so

$$[p(x)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$$

This is the correspondence between $\mathcal{P}_2$ and $\mathbb{R}^3$ that we remarked on after Example 6.26, and it can easily be generalized to show that the coordinate vector of a polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \quad \text{in } \mathcal{P}_n$$

with respect to the standard basis $\mathcal{B} = \{1, x, x^2, \ldots, x^n\}$ is just the vector

$$[p(x)]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{in } \mathbb{R}^{n+1}$$

Remark The *order* in which the basis vectors appear in $\mathcal{B}$ affects the order of the entries in a coordinate vector. For example, in Example 6.35, assume that the standard basis vectors are ordered as $\mathcal{B}' = \{x^2, x, 1\}$. Then the coordinate vector of $p(x) = 2 - 3x + 5x^2$ with respect to $\mathcal{B}'$ is

$$[p(x)]_{\mathcal{B}'} = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix}$$

Example 6.36

Find the coordinate vector $[A]_{\mathcal{B}}$ of $A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix}$ with respect to the standard basis $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ of M_{22}.

Solution Since

$$A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} = 2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 4\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 3\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= 2E_{11} - E_{12} + 4E_{21} + 3E_{22}$$

we have

$$[A]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \\ 4 \\ 3 \end{bmatrix}$$

This is the correspondence between M_{22} and $\mathbb{R}^4$ that we noted before the introduction to Example 6.13. It too can easily be generalized to give a correspondence between M_{mn} and $\mathbb{R}^{mn}$.

Example 6.37

Find the coordinate vector $[p(x)]_{\mathcal{B}}$ of $p(x) = 1 + 2x - x^2$ with respect to the basis $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\}$ of $\mathscr{P}_2$.

Solution We need to find c_1, c_2, and c_3 such that

$$c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2) = 1 + 2x - x^2$$

or, equivalently,

$$(c_1 + c_3) + (c_1 + c_2)x + (c_2 + c_3)x^2 = 1 + 2x - x^2$$

As in Example 6.32, this means we need to solve the system

$$\begin{aligned} c_1 + \quad\quad c_3 &= 1 \\ c_1 + c_2 \quad\quad &= 2 \\ c_2 + c_3 &= -1 \end{aligned}$$

whose solution is found to be $c_1 = 2$, $c_2 = 0$, $c_3 = -1$. Therefore,

$$[p(x)]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

[Since this result says that $p(x) = 2(1 + x) - (1 + x^2)$, it is easy to check that it is correct.]

The next theorem shows that the process of forming coordinate vectors is compatible with the vector space operations of addition and scalar multiplication.

Theorem 6.6 Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in V and let c be a scalar. Then

a. $[\mathbf{u} + \mathbf{v}]_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}$
b. $[c\mathbf{u}]_{\mathcal{B}} = c[\mathbf{u}]_{\mathcal{B}}$

Proof We begin by writing $\mathbf{u}$ and $\mathbf{v}$ in terms of the basis vectors—say, as

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \quad \text{and} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n$$

Then, using vector space properties, we have

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2 + \cdots + (c_n + d_n)\mathbf{v}_n$$

and

$$c\mathbf{u} = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 + \cdots + (cc_n)\mathbf{v}_n$$

so

$$[\mathbf{u} + \mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}$$

and

$$[c\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} cc_1 \\ cc_2 \\ \vdots \\ cc_n \end{bmatrix} = c\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c[\mathbf{u}]_{\mathcal{B}}$$

An easy corollary to Theorem 6.6 states that coordinate vectors preserve linear combinations:

$$[c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_{\mathcal{B}} = c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{u}_k]_{\mathcal{B}} \tag{1}$$

You are asked to prove this corollary in Exercise 31.

The most useful aspect of coordinate vectors is that they allow us to transfer information from a general vector space to $\mathbb{R}^n$, where we have the tools of Chapters 1 to 3 at our disposal. We will explore this idea in some detail in Sections 6.3 and 6.6. For now, we have the following useful theorem.

Theorem 6.7 Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V and let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be vectors in V. Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is linearly independent in V if and only if $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_k]_{\mathcal{B}}\}$ is linearly independent in $\mathbb{R}^n$.

Proof Assume that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is linearly independent in V and let

$$c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{u}_k]_{\mathcal{B}} = \mathbf{0}$$

in $\mathbb{R}^n$. But then we have

$$[c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_{\mathcal{B}} = \mathbf{0}$$

using Equation (1), so the coordinates of the vector $c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k$ with respect to $\mathcal{B}$ are all zero. That is,

$$c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n = \mathbf{0}$$

The linear independence of $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ now forces $c_1 = c_2 = \cdots = c_k = 0$, so $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_k]_{\mathcal{B}}\}$ is linearly independent.

The converse implication, which uses similar ideas, is left as Exercise 32.

Observe that, in the special case where $\mathbf{u}_i = \mathbf{v}_i$, we have

$$\mathbf{v}_i = 0 \cdot \mathbf{v}_1 + \cdots + 1 \cdot \mathbf{v}_i + \cdots + 0 \cdot \mathbf{v}_n$$

so $[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i$ and $\{[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}\} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis in $\mathbb{R}^n$.

Dimension

The definition of dimension is the same for a vector space as for a subspace of $\mathbb{R}^n$—the number of vectors in a basis for the space. Since a vector space can have more than one basis, we need to show that this definition makes sense; that is, we need to establish that different bases for the same vector space contain the same number of vectors.

Part (a) of the next theorem generalizes Theorem 2.8.

Theorem 6.8 Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V.

a. Any set of more than n vectors in V must be linearly dependent.
b. Any set of fewer than n vectors in V cannot span V.

Proof (a) Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a set of vectors in V, with $m > n$. Then $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ is a set of more than n vectors in $\mathbb{R}^n$ and, hence, is linearly dependent, by Theorem 2.8. This means that $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is linearly dependent as well, by Theorem 6.7.

(b) Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a set of vectors in V, with $m < n$. Then $S = \{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ is a set of fewer than n vectors in $\mathbb{R}^n$. Now $\text{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = V$ if and only if $\text{span}(S) = \mathbb{R}^n$ (see Exercise 33). But $\text{span}(S)$ is just the column space of the $n \times m$ matrix

$$A = [\,[\mathbf{u}_1]_{\mathcal{B}} \cdots [\mathbf{u}_m]_{\mathcal{B}}\,]$$

so $\dim(\text{span}(S)) = \dim(\text{col}(A)) \le m < n$. Hence, S cannot span $\mathbb{R}^n$, so $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ does not span V.

Now we extend Theorem 3.23.

Theorem 6.9 **The Basis Theorem**

If a vector space V has a basis with n vectors, then every basis for V has exactly n vectors.

The proof of Theorem 3.23 also works here, virtually word for word. However, it is easier to make use of Theorem 6.8.

Proof Let $\mathcal{B}$ be a basis for V with n vectors and let $\mathcal{B}'$ be another basis for V with m vectors. By Theorem 6.8, $m \leq n$; otherwise, $\mathcal{B}'$ would be linearly dependent.

Now use Theorem 6.8 with the roles of $\mathcal{B}$ and $\mathcal{B}'$ interchanged. Since $\mathcal{B}'$ is a basis of V with m vectors, Theorem 6.8 implies that any set of more than m vectors in V is linearly dependent. Hence, $n \leq m$, since $\mathcal{B}$ is a basis and is, therefore, linearly independent.

Since $n \leq m$ and $m \leq n$, we must have $n = m$, as required.

The following definition now makes sense, since the number of vectors in a (finite) basis does not depend on the choice of basis.

Definition A vector space V is called ***finite-dimensional*** if it has a basis consisting of finitely many vectors. The ***dimension*** of V, denoted by dim V, is the number of vectors in a basis for V. The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero. A vector space that has no finite basis is called ***infinite-dimensional.***

Example 6.38

Since the standard basis for $\mathbb{R}^n$ has n vectors, dim $\mathbb{R}^n = n$. In the case of $\mathbb{R}^3$, a one-dimensional subspace is just the span of a single nonzero vector and thus is a line through the origin. A two-dimensional subspace is spanned by its basis of two linearly independent (i.e., nonparallel) vectors and therefore is a plane through the origin. Any three linearly independent vectors must span $\mathbb{R}^3$, by the Fundamental Theorem. The subspaces of $\mathbb{R}^3$ are now completely classified according to dimension, as shown in Table 6.1.

Table 6.1

dim V	V
3	$\mathbb{R}^3$
2	Plane through the origin
1	Line through the origin
0	$\{\mathbf{0}\}$

Example 6.39

The standard basis for $\mathscr{P}_n$ contains $n + 1$ vectors (see Example 6.30), so dim $\mathscr{P}_n = n + 1$.

Example 6.40

The standard basis for M_{mn} contains mn vectors (see Example 6.31), so $\dim M_{mn} = mn$.

Example 6.41

Both $\mathscr{P}$ and $\mathscr{F}$ are infinite-dimensional, since they each contain the infinite linearly independent set $\{1, x, x^2, \ldots\}$ (see Exercise 44).

Example 6.42

Find the dimension of the vector space W of symmetric 2×2 matrices (see Example 6.10).

Solution A symmetric 2×2 matrix is of the form

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

so W is spanned by the set

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

If S is linearly independent, then it will be a basis for W. Setting

$$a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

we obtain

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

from which it immediately follows that $a = b = c = 0$. Hence, S is linearly independent and is, therefore, a basis for W. We conclude that $\dim W = 3$.

The dimension of a vector space is its "magic number." Knowing the dimension of a vector space V provides us with much information about V and can greatly simplify the work needed in certain types of calculations, as the next few theorems and examples illustrate.

Theorem 6.10

Let V be a vector space with $\dim V = n$. Then:

a. Any linearly independent set in V contains at most n vectors.
b. Any spanning set for V contains at least n vectors.
c. Any linearly independent set of exactly n vectors in V is a basis for V.
d. Any spanning set for V consisting of exactly n vectors is a basis for V.
e. Any linearly independent set in V can be extended to a basis for V.
f. Any spanning set for V can be reduced to a basis for V.

Proof The proofs of properties (a) and (b) follow from parts (a) and (b) of Theorem 6.8, respectively.

(c) Let S be a linearly independent set of exactly n vectors in V. If S does not span V, then there is some vector $\mathbf{v}$ in V that is not a linear combination of the vectors in S. Inserting $\mathbf{v}$ into S produces a set S' with $n + 1$ vectors that is still linearly independent (see Exercise 54). But this is impossible, by Theorem 6.8(a). We conclude that S must span V and therefore be a basis for V.

(d) Let S be a spanning set for V consisting of exactly n vectors. If S is linearly dependent, then some vector $\mathbf{v}$ in S is a linear combination of the others. Throwing $\mathbf{v}$ away leaves a set S' with $n - 1$ vectors that still spans V (see Exercise 55). But this is impossible, by Theorem 6.8(b). We conclude that S must be linearly independent and therefore be a basis for V.

(e) Let S be a linearly independent set of vectors in V. If S spans V, it is a basis for V and so consists of exactly n vectors, by the Basis Theorem. If S does not span V, then, as in the proof of property (c), there is some vector $\mathbf{v}$ in V that is not a linear combination of the vectors in S. Inserting $\mathbf{v}$ into S produces a set S' that is still linearly independent. If S' still does not span V, we can repeat the process and expand it into a larger, linearly independent set. Eventually, this process must stop, since no linearly independent set in V can contain more than n vectors, by Theorem 6.8(a). When the process stops, we have a linearly independent set S^* that contains S and also spans V. Therefore, S^* is a basis for V that extends S.

(f) You are asked to prove this property in Exercise 56.

You should view Theorem 6.10 as, in part, a labor-saving device. In many instances, it can dramatically decrease the amount of work needed to check that a set of vectors is linearly independent, a spanning set, or a basis.

Example 6.43

In each case, determine whether S is a basis for V.

(a) $V = \mathcal{P}_2$, $S = \{1 + x, 2 - x + x^2, 3x - 2x^2, -1 + 3x + x^2\}$

(b) $V = M_{22}$, $S = \left\{\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\0 & -1\end{bmatrix}\right\}$

(c) $V = \mathcal{P}_2$, $S = \{1 + x, x + x^2, 1 + x^2\}$

Solution (a) Since $\dim(\mathcal{P}_2) = 3$ and S contains four vectors, S is linearly dependent, by Theorem 6.10(a). Hence, S is not a basis for $\mathcal{P}_2$.

(b) Since $\dim(M_{22}) = 4$ and S contains three vectors, S cannot span M_{22}, by Theorem 6.10(b). Hence, S is not a basis for M_{22}.

(c) Since $\dim(\mathcal{P}_2) = 3$ and S contains three vectors, S will be a basis for $\mathcal{P}_2$ if it is linearly independent or if it spans $\mathcal{P}_2$, by Theorem 6.10(c) or (d). It is easier to show that S is linearly independent; we did this in Example 6.26. Therefore, S is a basis for $\mathcal{P}_2$. (This is the same problem as in Example 6.32—but see how much easier it becomes using Theorem 6.10!)

Example 6.44

Extend $\{1 + x, 1 - x\}$ to a basis for $\mathcal{P}_2$.

Solution First note that $\{1 + x, 1 - x\}$ is linearly independent. (Why?) Since $\dim(\mathcal{P}_2) = 3$, we need a third vector—one that is not linearly dependent on the first two.

We could proceed, as in the proof of Theorem 6.10(e), to find such a vector using trial and error. However, it is easier in practice to proceed in a different way.

We enlarge the given set of vectors by throwing in the *entire* standard basis for $\mathcal{P}_2$. This gives

$$S = \{1 + x, 1 - x, 1, x, x^2\}$$

Now S is linearly dependent, by Theorem 6.10(a), so we need to throw away some vectors—in this case, two. Which ones? We use Theorem 6.10(f), starting with the first vector that was added, 1. Since $1 = \frac{1}{2}(1 + x) + \frac{1}{2}(1 - x)$, the set $\{1 + x, 1 - x, 1\}$ is linearly dependent, so we throw away 1. Similarly, $x = \frac{1}{2}(1 + x) - \frac{1}{2}(1 - x)$, so $\{1 + x, 1 - x, x\}$ is linearly dependent also. Finally, we check that $\{1 + x, 1 - x, x^2\}$ is linearly independent. (Can you see a quick way to tell this?) Therefore, $\{1 + x, 1 - x, x^2\}$ is a basis for $\mathcal{P}_2$ that extends $\{1 + x, 1 - x\}$.

In Example 6.42, the vector space W of symmetric 2×2 matrices is a subspace of the vector space M_{22} of all 2×2 matrices. As we showed, $\dim W = 3 \leq 4 = \dim M_{22}$. This is an example of a general result, as the final theorem of this section shows.

Theorem 6.11 Let W be a subspace of a finite-dimensional vector space V. Then:

a. W is finite-dimensional and $\dim W \leq \dim V$.
b. $\dim W = \dim V$ if and only if $W = V$.

Proof (a) Let $\dim V = n$. If $W = \{\mathbf{0}\}$, then $\dim(W) = 0 \leq n = \dim V$. If W is nonzero, then any basis $\mathcal{B}$ for V (containing n vectors) certainly spans W, since W is contained in V. But $\mathcal{B}$ can be reduced to a basis $\mathcal{B}'$ for W (containing at most n vectors), by Theorem 6.10(f). Hence, W is finite-dimensional and $\dim(W) \leq n = \dim V$.

(b) If $W = V$, then certainly $\dim W = \dim V$. On the other hand, if $\dim W = \dim V = n$, then any basis $\mathcal{B}$ for W consists of exactly n vectors. But these are then n linearly independent vectors in V and, hence, a basis for V, by Theorem 6.10(c). Therefore, $V = \text{span}(\mathcal{B}) = W$.

Exercises 6.2

In Exercises 1–4, test the sets of matrices for linear independence in M_{22}. For those that are linearly dependent, express one of the matrices as a linear combination of the others.

1. $\left\{\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}\right\}$

2. $\left\{\begin{bmatrix} 2 & -3 \\ 4 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 3 & 3 \end{bmatrix}, \begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix}\right\}$

3. $\left\{\begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ -1 & 7 \end{bmatrix}\right\}$

4. $\left\{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right\}$

In Exercises 5–9, test the sets of polynomials for linear independence. For those that are linearly dependent, express one of the polynomials as a linear combination of the others.

5. $\{x, 1 + x\}$ in $\mathcal{P}_1$

6. $\{1 + x, 1 + x^2, 1 - x + x^2\}$ in $\mathcal{P}_2$

7. $\{x, 2x - x^2, 3x + 2x^2\}$ in $\mathcal{P}_2$

8. $\{2x, x - x^2, 1 + x^3, 2 - x^2 + x^3\}$ in $\mathscr{P}_3$

9. $\{1 - 2x, 3x + x^2 - x^3, 1 + x^2 + 2x^3, 3 + 2x + 3x^3\}$ in $\mathscr{P}_3$

In Exercises 10–14, test the sets of functions for linear independence in $\mathscr{F}$. For those that are linearly dependent, express one of the functions as a linear combination of the others.

10. $\{1, \sin x, \cos x\}$

11. $\{1, \sin^2 x, \cos^2 x\}$

12. $\{e^x, e^{-x}\}$

13. $\{1, \ln(2x), \ln(x^2)\}$

14. $\{\sin x, \sin 2x, \sin 3x\}$

15. If f and g are in $\mathscr{C}^{(1)}$, the vector space of all functions with continuous derivatives, then the determinant

$$W(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix}$$

is called the ***Wronskian*** of f and g [named after the Polish-French mathematician Jósef Maria Hoëné-Wronski (1776–1853), who worked on the theory of determinants and the philosophy of mathematics]. Show that f and g are linearly independent if their Wronskian is not identically zero (that is, if there is some x such that $W(x) \neq 0$).

16. In general, the Wronskian of $f_1, \ldots, f_n$ in $\mathscr{C}^{(n-1)}$ is the determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

and $f_1, \ldots, f_n$ are linearly independent, provided $W(x)$ is not identically zero. Repeat Exercises 10–14 using the Wronskian test.

17. Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a linearly independent set of vectors in a vector space V.

(a) Is $\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{u} + \mathbf{w}\}$ linearly independent? Either prove that it is or give a counterexample to show that it is not.

(b) Is $\{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{u} - \mathbf{w}\}$ linearly independent? Either prove that it is or give a counterexample to show that it is not.

In Exercises 18–25, determine whether the set $\mathcal{B}$ is a basis for the vector space V.

18. $V = M_{22}$, $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \right\}$

19. $V = M_{22}$, $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right\}$

20. $V = M_{22}$,

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\}$$

21. $V = M_{22}$,

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \right\}$$

22. $V = \mathscr{P}_2$, $\mathcal{B} = \{x, 1 + x, x - x^2\}$

23. $V = \mathscr{P}_2$, $\mathcal{B} = \{1 - x, 1 - x^2, x - x^2\}$

24. $V = \mathscr{P}_2$, $\mathcal{B} = \{1, 1 + 2x + 3x^2\}$

25. $V = \mathscr{P}_2$, $\mathcal{B} = \{1, 2 - x, 3 - x^2, x + 2x^2\}$

26. Find the coordinate vector of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ with respect to the basis $\mathcal{B} = \{E_{22}, E_{21}, E_{12}, E_{11}\}$ of M_{22}.

27. Find the coordinate vector of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ with respect to the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ of M_{22}.

28. Find the coordinate vector of $p(x) = 1 + 2x + 3x^2$ with respect to the basis $\mathcal{B} = \{1 + x, 1 - x, x^2\}$ of $\mathscr{P}_2$.

29. Find the coordinate vector of $p(x) = 2 - x + 3x^2$ with respect to the basis $\mathcal{B} = \{1, 1 + x, -1 + x^2\}$ of $\mathscr{P}_2$.

30. Let $\mathcal{B}$ be a set of vectors in a vector space V with the property that every vector in V can be written uniquely as a linear combination of the vectors in $\mathcal{B}$. Prove that $\mathcal{B}$ is a basis for V.

31. Let $\mathcal{B}$ be a basis for a vector space V, let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be vectors in V, and let $c_1, \ldots, c_k$ be scalars. Show that $[c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_{\mathcal{B}} = c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{u}_k]_{\mathcal{B}}$.

32. Finish the proof of Theorem 6.7 by showing that if $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_k]_{\mathcal{B}}\}$ is linearly independent in $\mathbb{R}^n$ then $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is linearly independent in V.

33. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a set of vectors in an n-dimensional vector space V and let $\mathcal{B}$ be a basis for V. Let $S = \{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ be the set of coordinate vectors of $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ with respect to $\mathcal{B}$. Prove that $\text{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = V$ if and only if $\text{span}(S) = \mathbb{R}^n$.

In Exercises 34–39, find the dimension of the vector space V and give a basis for V.

34. $V = \{p(x) \text{ in } \mathscr{P}_2 : p(0) = 0\}$

35. $V = \{p(x) \text{ in } \mathscr{P}_2 : p(1) = 0\}$

36. $V = \{p(x) \text{ in } \mathscr{P}_2 : xp'(x) = p(x)\}$

37. $V = \{A \text{ in } M_{22} : A \text{ is upper triangular}\}$

38. $V = \{A \text{ in } M_{22} : A \text{ is skew-symmetric}\}$

39. $V = \{A \text{ in } M_{22} : AB = BA\}$, where $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

40. Find a formula for the dimension of the vector space of symmetric $n \times n$ matrices.

41. Find a formula for the dimension of the vector space of skew-symmetric $n \times n$ matrices.

42. Let U and W be subspaces of a finite-dimensional vector space V. Prove ***Grassmann's Identity:***

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

[*Hint:* The subspace $U + W$ is defined in Exercise 48 of Section 6.1. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for $U \cap W$. Extend $\mathcal{B}$ to a basis $\mathcal{C}$ of U and a basis $\mathcal{D}$ of W. Prove that $\mathcal{C} \cup \mathcal{D}$ is a basis for $U + W$.]

43. Let U and V be finite-dimensional vector spaces.

(a) Find a formula for $\dim(U \times V)$ in terms of $\dim U$ and $\dim V$. (See Exercise 49 in Section 6.1.)

(b) If W is a subspace of V, show that $\dim \Delta = \dim W$, where $\Delta = \{(\mathbf{w}, \mathbf{w}) : \mathbf{w} \text{ is in } W\}$.

44. Prove that the vector space $\mathcal{P}$ is infinite-dimensional. [*Hint:* Suppose it has a finite basis. Show that there is some polynomial that is not a linear combination of this basis.]

45. Extend $\{1 + x, 1 + x + x^2\}$ to a basis for $\mathcal{P}_2$.

46. Extend $\left\{\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right\}$ to a basis for M_{22}.

47. Extend $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right\}$ to a basis for M_{22}.

48. Extend $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\}$ to a basis for the vector space of symmetric 2×2 matrices.

49. Find a basis for $\operatorname{span}(1, 1 + x, 2x)$ in $\mathcal{P}_1$.

50. Find a basis for $\operatorname{span}(1 - 2x, 2x - x^2, 1 - x^2, 1 + x^2)$ in $\mathcal{P}_2$.

51. Find a basis for $\operatorname{span}(1 - x, x - x^2, 1 - x^2, 1 - 2x + x^2)$ in $\mathcal{P}_2$.

52. Find a basis for $\operatorname{span}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right)$ in M_{22}.

53. Find a basis for $\operatorname{span}(\sin^2 x, \cos^2 x, \cos 2x)$ in $\mathcal{F}$.

54. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a linearly independent set in a vector space V. Show that if $\mathbf{v}$ is a vector in V that is not in $\operatorname{span}(S)$, then $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}\}$ is still linearly independent.

55. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a spanning set for a vector space V. Show that if $\mathbf{v}_n$ is in $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$, then $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ is still a spanning set for V.

56. Prove Theorem 6.10(f).

57. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V and let $c_1, \ldots, c_n$ be nonzero scalars. Prove that $\{c_1\mathbf{v}_1, \ldots, c_n\mathbf{v}_n\}$ is also a basis for V.

58. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V. Prove that

$$\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3, \ldots, \mathbf{v}_1 + \cdots + \mathbf{v}_n\}$$

is also a basis for V.

Let $a_0, a_1, \ldots, a_n$ be $n + 1$ distinct real numbers. Define polynomials $p_0(x), p_1(x), \ldots, p_n(x)$ by

$$p_i(x) = \frac{(x - a_0)\cdots(x - a_{i-1})(x - a_{i+1})\cdots(x - a_n)}{(a_i - a_0)\cdots(a_i - a_{i-1})(a_i - a_{i+1})\cdots(a_i - a_n)}$$

*These are called the **Lagrange polynomials** associated with $a_0, a_1, \ldots, a_n$. [Joseph-Louis Lagrange (1736–1813) was born in Italy but spent most of his life in Germany and France. He made important contributions to such fields as number theory, algebra, astronomy, mechanics, and the calculus of variations. In 1773, Lagrange was the first to give the volume interpretation of a determinant (see Chapter 4).]*

59. **(a)** Compute the Lagrange polynomials associated with $a_0 = 1, a_1 = 2, a_2 = 3$.

(b) Show, in general, that

$$p_i(a_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

60. **(a)** Prove that the set $\mathcal{B} = \{p_0(x), p_1(x), \ldots, p_n(x)\}$ of Lagrange polynomials is linearly independent in $\mathcal{P}_n$. [*Hint:* Set $c_0p_0(x) + \cdots + c_np_n(x) = 0$ and use Exercise 59(b).]

(b) Deduce that $\mathcal{B}$ is a basis for $\mathcal{P}_n$.

61. If $q(x)$ is an arbitrary polynomial in $\mathcal{P}_n$, it follows from Exercise 60(b) that

$$q(x) = c_0p_0(x) + \cdots + c_np_n(x) \qquad (1)$$

for some scalars $c_0, \ldots, c_n$.

(a) Show that $c_i = q(a_i)$ for $i = 0, \ldots, n$, and deduce that $q(x) = q(a_0)p_0(x) + \cdots + q(a_n)p_n(x)$ is the unique representation of $q(x)$ with respect to the basis $\mathcal{B}$.

(b) Show that for any $n + 1$ points $(a_0, c_0), (a_1, c_1), \ldots, (a_n, c_n)$ with distinct first components, the function $q(x)$ defined by Equation (1) is the unique polynomial of degree at most n that passes through all of the points. This formula is known as the ***Lagrange interpolation formula.*** (Compare this formula with Problem 19 in Exploration: Geometric Applications of Determinants in Chapter 4.)

(c) Use the Lagrange interpolation formula to find the polynomial of degree at most 2 that passes through the points

(i) $(1, 6), (2, -1)$, and $(3, -2)$

(ii) $(-1, 10), (0, 5)$, and $(3, 2)$

62. Use the Lagrange interpolation formula to show that if a polynomial in $\mathscr{P}_n$ has $n + 1$ zeros, then it must be the zero polynomial.

63. Find a formula for the number of invertible matrices in $M_{nn}(\mathbb{Z}_p)$. [*Hint:* This is the same as determining the number of different bases for $\mathbb{Z}_p^n$. (Why?) Count the number of ways to construct a basis for $\mathbb{Z}_p^n$, one vector at a time.]

Exploration

Magic Squares

The engraving shown on page 461 is Albrecht Dürer's *Melancholia I* (1514). Among the many mathematical artifacts in this engraving is the chart of numbers that hangs on the wall in the upper right-hand corner. (It is enlarged in the detail shown.) Such an array of numbers is known as a *magic square*. We can think of it as a 4×4 matrix

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}$$

Observe that the numbers in each row, in each column, and in both diagonals have the same sum: 34. Observe further that the entries are the integers $1, 2, \ldots, 16$. (Note that Dürer cleverly placed the 15 and 14 adjacent to each other in the last row, giving the date of the engraving.) These observations lead to the following definition.

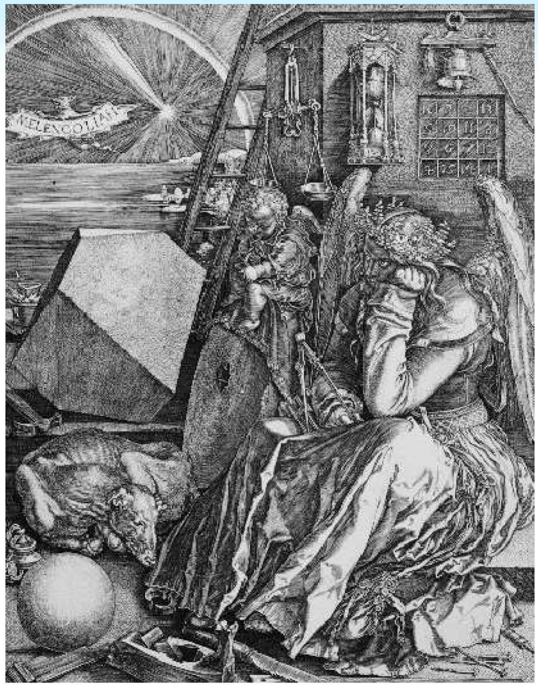
Definition An $n \times n$ matrix M is called a ***magic square*** if the sum of the entries is the same in each row, each column, and both diagonals. This common sum is called the ***weight*** of M, denoted $\text{wt}(M)$. If M is an $n \times n$ magic square that contains each of the entries $1, 2, \ldots, n^2$ exactly once, then M is called a ***classical magic square.***

1. If M is a classical $n \times n$ magic square, show that

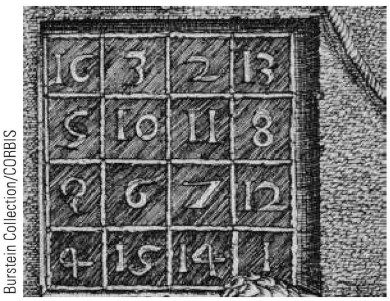
$$\text{wt}(M) = \frac{n(n^2 + 1)}{2}$$

[*Hint:* Use Exercise 51 in Section 2.4.]

2. Find a classical 3×3 magic square. Find a different one. Are your two examples related in any way?



Burstein Collection/CORBIS



Burstein Collection/CORBIS

3. Clearly, the 3×3 matrix with all entries equal to $\frac{1}{3}$ is a magic square with weight 1. Using your answer to Problem 2, find a 3×3 magic square with weight 1, *all of whose entries are different.* Describe a method for constructing a 3×3 magic square with distinct entries and weight w for any real number w.

Let Mag_n denote the set of all $n \times n$ magic squares, and let Mag_n^0 denote the set of all $n \times n$ magic squares of weight 0.

4. (a) Prove that Mag_3 is a subspace of M_{33}.
   (b) Prove that Mag_3^0 is a subspace of Mag_3.

5. Use Problems 3 and 4 to show that if M is a 3×3 magic square with weight w, then we can write M as

$$M = M_0 + kJ$$

where M_0 is a 3×3 magic square of weight 0, J is the 3×3 matrix consisting entirely of ones, and k is a scalar. What must k be? [*Hint:* Show that $M - kJ$ is in Mag_3^0 for an appropriate value of k.]

Let's try to find a way of describing *all* 3×3 magic squares. Let

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

be a magic square with weight 0. The conditions on the rows, columns, and diagonals give rise to a system of eight homogeneous linear equations in the variables a, $b, \ldots, i$.

6. Write out this system of equations and solve it. [*Note:* Using a CAS will facilitate the calculations.]

7. Find the dimension of Mag_3^0. *Hint:* By doing a substitution, if necessary, use your solution to Problem 6 to show that M can be written in the form

$$M = \begin{bmatrix} s & -s-t & t \\ -s+t & 0 & s-t \\ -t & s+t & -s \end{bmatrix}$$

8. Find the dimension of Mag_3. [*Hint:* Combine the results of Problems 5 and 7.]

9. Can you find a direct way of showing that the (2, 2) entry of a 3×3 magic square with weight w must be $w/3$? [*Hint:* Add and subtract certain rows, columns, and diagonals to leave a multiple of the central entry.]

10. Let M be a 3×3 magic square of weight 0, obtained from a classical 3×3 magic square as in Problem 5. If M has the form given in Problem 7, write out an equation for the sum of the squares of the entries of M. Show that this is the equation of a circle in the variables s and t, and carefully plot it. Show that there are exactly eight points (s, t) on this circle with both s and t integers. Using Problem 8, show that these eight points give rise to eight classical 3×3 magic squares. How are these magic squares related to one another?

6.3 Change of Basis

In many applications, a problem described using one coordinate system may be solved more easily by switching to a new coordinate system. This switch is usually accomplished by performing a change of variables, a process that you have probably encountered in other mathematics courses. In linear algebra, a basis provides us with a coordinate system for a vector space, via the notion of coordinate vectors. Choosing the right basis will often greatly simplify a particular problem. For example, consider the molecular structure of zinc, shown in Figure 6.3(a). A scientist studying zinc might wish to measure the lengths of the bonds between the atoms, the angles between these bonds, and so on. Such an analysis will be greatly facilitated by introducing coordinates and making use of the tools of linear algebra. The standard basis and the associated standard xyz coordinate axes are not always the best choice. As Figure 6.3(b) shows, in this case $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is probably a better choice of basis for $\mathbb{R}^3$ than the standard basis, since these vectors align nicely with the bonds between the atoms of zinc.

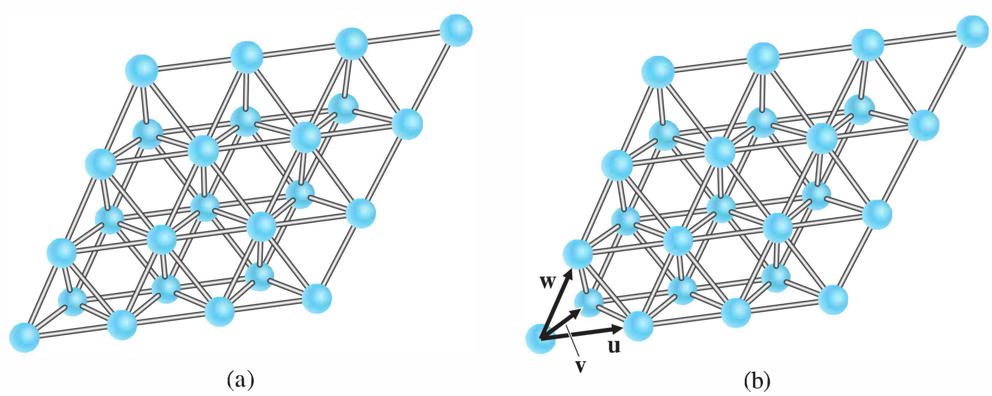


Figure 6.3

Change-of-Basis Matrices

Figure 6.4 shows two different coordinate systems for $\mathbb{R}^2$, each arising from a different basis. Figure 6.4(a) shows the coordinate system related to the basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$, while Figure 6.4(b) arises from the basis $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

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

The same vector $\mathbf{x}$ is shown relative to each coordinate system. It is clear from the diagrams that the coordinate vectors of $\mathbf{x}$ with respect to $\mathcal{B}$ and $\mathcal{C}$ are

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{and} \quad [\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$$

respectively. It turns out that there is a direct connection between the two coordinate vectors. One way to find the relationship is to use $[\mathbf{x}]_{\mathcal{B}}$ to calculate

$$\mathbf{x} = \mathbf{u}_1 + 3\mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

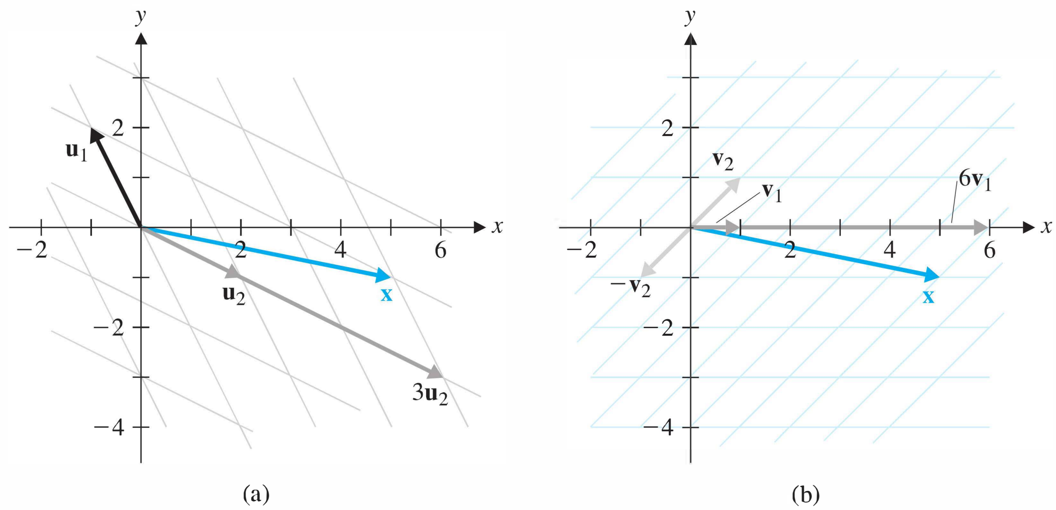


Figure 6.4

Then we can find $[\mathbf{x}]_{\mathcal{C}}$ by writing $\mathbf{x}$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. However, there is a better way to proceed—one that will provide us with a general mechanism for such problems. We illustrate this approach in the next example.

Example 6.45

Using the bases $\mathcal{B}$ and $\mathcal{C}$ above, find $[\mathbf{x}]_{\mathcal{C}}$, given that $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Solution Since $\mathbf{x} = \mathbf{u}_1 + 3\mathbf{u}_2$, writing $\mathbf{u}_1$ and $\mathbf{u}_2$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$ will give us the required coordinates of $\mathbf{x}$ with respect to $\mathcal{C}$. We find that

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -3\mathbf{v}_1 + 2\mathbf{v}_2$$

and

$$\mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\mathbf{v}_1 - \mathbf{v}_2$$

so

$$\begin{aligned} \mathbf{x} &= \mathbf{u}_1 + 3\mathbf{u}_2 \\ &= (-3\mathbf{v}_1 + 2\mathbf{v}_2) + 3(3\mathbf{v}_1 - \mathbf{v}_2) \\ &= 6\mathbf{v}_1 - \mathbf{v}_2 \end{aligned}$$

This gives

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$$

in agreement with Figure 6.4(b).

This method may not look any easier than the one suggested prior to Example 6.45, but it has one big advantage: We can now find $[\mathbf{y}]_{\mathcal{C}}$ from $[\mathbf{y}]_{\mathcal{B}}$ for *any* vector $\mathbf{y}$ in $\mathbb{R}^2$

with very little additional work. Let's look at the calculations in Example 6.45 from a different point of view. From $\mathbf{x} = \mathbf{u}_1 + 3\mathbf{u}_2$, we have

$$[\mathbf{x}]_{\mathcal{C}} = [\mathbf{u}_1 + 3\mathbf{u}_2]_{\mathcal{C}} = [\mathbf{u}_1]_{\mathcal{C}} + 3[\mathbf{u}_2]_{\mathcal{C}}$$

by Theorem 6.6. Thus,

$$\begin{aligned}[\mathbf{x}]_{\mathcal{C}} &= [[\mathbf{u}_1]_{\mathcal{C}}[\mathbf{u}_2]_{\mathcal{C}}]\begin{bmatrix}1\\3\end{bmatrix}\\ &= \begin{bmatrix}-3 & 3\\ 2 & -1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}\\ &= P[\mathbf{x}]_{\mathcal{B}}\end{aligned}$$

where P is the matrix whose columns are $[\mathbf{u}_1]_{\mathcal{C}}$ and $[\mathbf{u}_2]_{\mathcal{C}}$. This procedure generalizes very nicely.

Definition Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be bases for a vector space V. The $n \times n$ matrix whose columns are the coordinate vectors $[\mathbf{u}_1]_{\mathcal{C}}, \ldots, [\mathbf{u}_n]_{\mathcal{C}}$ of the vectors in $\mathcal{B}$ with respect to $\mathcal{C}$ is denoted by $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and is called the ***change-of-basis matrix*** from $\mathcal{B}$ to $\mathcal{C}$. That is,

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = [\,[\mathbf{u}_1]_{\mathcal{C}}[\mathbf{u}_2]_{\mathcal{C}} \cdots [\mathbf{u}_n]_{\mathcal{C}}]$$

Think of $\mathcal{B}$ as the "old" basis and $\mathcal{C}$ as the "new" basis. Then the columns of $P_{\mathcal{C}\leftarrow\mathcal{B}}$ are just the coordinate vectors obtained by writing the old basis vectors in terms of the new ones. Theorem 6.12 shows that Example 6.45 is a special case of a general result.

Theorem 6.12 Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be bases for a vector space V and let $P_{\mathcal{C}\leftarrow\mathcal{B}}$ be the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$. Then

a. $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ for all $\mathbf{x}$ in V.
b. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is the unique matrix P with the property that $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ for all $\mathbf{x}$ in V.
c. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible and $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

Proof (a) Let $\mathbf{x}$ be in V and let

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

That is, $\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$. Then

$$\begin{aligned}[\mathbf{x}]_{\mathcal{C}} &= [c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n]_{\mathcal{C}}\\ &= c_1[\mathbf{u}_1]_{\mathcal{C}} + \cdots + c_n[\mathbf{u}_n]_{\mathcal{C}}\\ &= [\,[\mathbf{u}_1]_{\mathcal{C}} \cdots [\mathbf{u}_n]_{\mathcal{C}}]\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}\\ &= P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}\end{aligned}$$

(b) Suppose that P is an $n \times n$ matrix with the property that $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ for all $\mathbf{x}$ in V. Taking $\mathbf{x} = \mathbf{u}_i$, the ith basis vector in $\mathcal{B}$, we see that $[\mathbf{x}]_{\mathcal{B}} = [\mathbf{u}_i]_{\mathcal{B}} = \mathbf{e}_i$, so the ith column of P is

$$\mathbf{p}_i = P\mathbf{e}_i = P[\mathbf{u}_i]_{\mathcal{B}} = [\mathbf{u}_i]_{\mathcal{C}}$$

which is the ith column of $P_{\mathcal{C}\leftarrow\mathcal{B}}$, by definition. It follows that $P = P_{\mathcal{C}\leftarrow\mathcal{B}}$.

(c) Since $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is linearly independent in V, the set $\{[\mathbf{u}_1]_{\mathcal{C}}, \dots, [\mathbf{u}_n]_{\mathcal{C}}\}$ is linearly independent in $\mathbb{R}^n$, by Theorem 6.7. Hence, $P_{\mathcal{C}\leftarrow\mathcal{B}} = [[\mathbf{u}_1]_{\mathcal{C}} \ \cdots \ [\mathbf{u}]_{\mathcal{C}}$ is invertible, by the Fundamental Theorem.

For all $\mathbf{x}$ in V, we have $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$. Solving for $[\mathbf{x}]_{\mathcal{B}}$, we find that

$$[\mathbf{x}]_{\mathcal{B}} = (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}[\mathbf{x}]_{\mathcal{C}}$$

for all $\mathbf{x}$ in V. Therefore, $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}$ is a matrix that changes bases from $\mathcal{C}$ to $\mathcal{B}$. Thus, by the uniqueness property (b), we must have $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

Remarks

- You may find it helpful to think of change of basis as a transformation (indeed, it is a linear transformation) from $\mathbb{R}^n$ to itself that simply switches from one coordinate system to another. The transformation corresponding to $P_{\mathcal{C}\leftarrow\mathcal{B}}$ accepts $[\mathbf{x}]_{\mathcal{B}}$ as input and returns $[\mathbf{x}]_{\mathcal{C}}$ as output; $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$ does just the opposite. Figure 6.5 gives a schematic representation of the process.

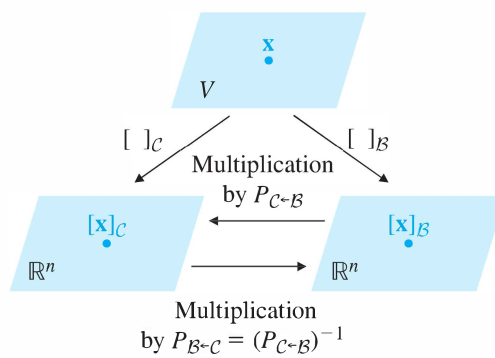


Figure 6.5
Change of basis

- The columns of $P_{\mathcal{C}\leftarrow\mathcal{B}}$ are the coordinate vectors of one basis with respect to the other basis. To remember which basis is which, think of the notation $\mathcal{C} \leftarrow \mathcal{B}$ as saying "$\mathcal{B}$ in terms of $\mathcal{C}$." It is also helpful to remember that $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$ is a linear combination of the columns of $P_{\mathcal{C}\leftarrow\mathcal{B}}$. But since the result of this combination is $[\mathbf{x}]_{\mathcal{C}}$, the columns of $P_{\mathcal{C}\leftarrow\mathcal{B}}$ must themselves be coordinate vectors with respect to $\mathcal{C}$.

Example 6.46

Find the change-of-basis matrices $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$ for the bases $\mathcal{B} = \{1, x, x^2\}$ and $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\}$ of $\mathcal{P}_2$. Then find the coordinate vector of $p(x) = 1 + 2x - x^2$ with respect to $\mathcal{C}$.

Solution Changing *to* a standard basis is easy, so we find $P_{\mathcal{B}\leftarrow\mathcal{C}}$ first. Observe that the coordinate vectors for $\mathcal{C}$ in terms of $\mathcal{B}$ are

$$[1 + x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad [x + x^2]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad [1 + x^2]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

(Look back at the Remark following Example 6.26.) It follows that

$$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

To find $P_{\mathcal{C}\leftarrow\mathcal{B}}$, we could express each vector in $\mathcal{B}$ as a linear combination of the vectors in $\mathcal{C}$ (do this), but it is much easier to use the fact that $P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}$, by Theorem 6.12(c). We find that

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

It now follows that

$$\begin{aligned} [p(x)]_{\mathcal{C}} &= P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \end{aligned}$$

which agrees with Example 6.37.

Remark If we do not need $P_{\mathcal{C}\leftarrow\mathcal{B}}$ explicitly, we can find $[p(x)]_{\mathcal{C}}$ from $[p(x)]_{\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$ using Gaussian elimination. Row reduction produces

$$[P_{\mathcal{B}\leftarrow\mathcal{C}} \mid [p(x)]_{\mathcal{B}}] \longrightarrow [I \mid (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}[p(x)]_{\mathcal{B}}] = [I \mid P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}}] = [I \mid [p(x)]_{\mathcal{C}}]$$

(See the next section on using Gauss-Jordan elimination.)

It is worth repeating the observation in Example 6.46: Changing *to* a standard basis is easy. If $\mathcal{E}$ is the standard basis for a vector space V and $\mathcal{B}$ is any other basis, then the columns of $P_{\mathcal{E}\leftarrow\mathcal{B}}$ are the coordinate vectors of $\mathcal{B}$ with respect to $\mathcal{E}$, and these are usually "visible." We make use of this observation again in the next example.

Example 6.47

In M_{22}, let $\mathcal{B}$ be the basis $\{E_{11}, E_{21}, E_{12}, E_{22}\}$ and let $\mathcal{C}$ be the basis $\{A, B, C, D\}$, where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Find the change-of-basis matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and verify that $[X]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}}$ for $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Solution 1 To solve this problem directly, we must find the coordinate vectors of $\mathcal{B}$ with respect to $\mathcal{C}$. This involves solving four linear combination problems of the form $X = aA + bB + cC + dD$, where X is in $\mathcal{B}$ and we must find a, b, c, and d. However, here we are lucky, since we can find the required coefficients by inspection.

Clearly, $E_{11} = A$, $E_{21} = -B + C$, $E_{12} = -A + B$, and $E_{22} = -C + D$. Thus,

$$[E_{11}]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad [E_{21}]_{\mathcal{C}} = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad [E_{12}]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad [E_{22}]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

so
$$P_{\mathcal{C}\leftarrow\mathcal{B}} = [[E_{11}]_{\mathcal{C}} \;\; [E_{21}]_{\mathcal{C}} \;\; [E_{12}]_{\mathcal{C}} \;\; [E_{22}]_{\mathcal{C}}] = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then

$$[X]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$$

and
$$P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 4 \end{bmatrix}$$

This is the coordinate vector with respect to $\mathcal{C}$ of the matrix

$$\begin{aligned} -A - B - C + 4D &= -\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + 4\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = X \end{aligned}$$

as it should be.

Solution 2 We can compute $P_{\mathcal{C}\leftarrow\mathcal{B}}$ in a different way, as follows. As you will be asked to prove in Exercise 21, if $\mathcal{E}$ is another basis for M_{22}, then $P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{E}}P_{\mathcal{E}\leftarrow\mathcal{B}} = (P_{\mathcal{E}\leftarrow\mathcal{C}})^{-1}P_{\mathcal{E}\leftarrow\mathcal{B}}$. If $\mathcal{E}$ is the standard basis, then $P_{\mathcal{E}\leftarrow\mathcal{B}}$ and $P_{\mathcal{E}\leftarrow\mathcal{C}}$ can be found by inspection. We have

$$P_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Do you see why?) Therefore,

$$
\begin{aligned}
P_{\mathcal{C}\leftarrow\mathcal{B}} &= (P_{\mathcal{E}\leftarrow\mathcal{C}})^{-1}P_{\mathcal{E}\leftarrow\mathcal{B}} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

which agrees with the first solution.

Remark The second method has the advantage of not requiring the computation of any linear combinations. It has the disadvantage of requiring that we find a matrix inverse. However, using a CAS will facilitate finding a matrix inverse, so in general the second method is preferable to the first. For certain problems, though, the first method may be just as easy to use. In any event, we are about to describe yet a third approach, which you may find best of all.

The Gauss-Jordan Method for Computing a Change-of-Basis Matrix

Finding the change-of-basis matrix to a standard basis is easy and can be done by inspection. Finding the change-of-basis matrix from a standard basis is almost as easy, but requires the calculation of a matrix inverse, as in Example 6.46. If we do it by hand, then (except for the 2×2 case) we will usually find the necessary inverse by Gauss-Jordan elimination. We now look at a modification of the Gauss-Jordan method that can be used to find the change-of-basis matrix between two nonstandard bases, as in Example 6.47.

Suppose $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are bases for a vector space V and $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$. The ith column of P is

$$[\mathbf{u}_i]_{\mathcal{C}} = \begin{bmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{bmatrix}$$

so $\mathbf{u}_i = p_{1i}\mathbf{v}_1 + \cdots + p_{ni}\mathbf{v}_n$. If $\mathcal{E}$ is any basis for V, then

$$[\mathbf{u}_i]_{\mathcal{E}} = [p_{1i}\mathbf{v}_1 + \cdots + p_{ni}\mathbf{v}_n]_{\mathcal{E}} = p_{1i}[\mathbf{v}_1]_{\mathcal{E}} + \cdots + p_{ni}[\mathbf{v}_n]_{\mathcal{E}}$$

This can be rewritten in matrix form as

$$\big[\,[\mathbf{v}_1]_{\mathcal{E}} \;\cdots\; [\mathbf{v}_n]_{\mathcal{E}}\,\big] \begin{bmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{bmatrix} = [\mathbf{u}_i]_{\mathcal{E}}$$

which we can solve by applying Gauss-Jordan elimination to the augmented matrix

$$[[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}} \mid [\mathbf{u}_i]_{\mathcal{E}}]$$

There are n such systems of equations to be solved, one for each column of $P_{\mathcal{C}\leftarrow\mathcal{B}}$, but *the coefficient matrix* $[[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}}]$ *is the same in each case*. Hence, we can solve all the systems simultaneously by row reducing the $n \times 2n$ augmented matrix

$$[[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}} \mid [\mathbf{u}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{u}_n]_{\mathcal{E}}] = [C \mid B]$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, so is $\{[\mathbf{v}_1]_{\mathcal{E}}, \ldots, [\mathbf{v}_n]_{\mathcal{E}}\}$, by Theorem 6.7. Therefore, the matrix C whose columns are $[\mathbf{v}_1]_{\mathcal{E}}, \ldots, [\mathbf{v}_n]_{\mathcal{E}}$ has the $n \times n$ identity matrix I for its reduced row echelon form, by the Fundamental Theorem. It follows that Gauss-Jordan elimination will necessarily produce

$$[C \mid B] \rightarrow [I \mid P]$$

where $P = P_{\mathcal{C}\leftarrow\mathcal{B}}$.

We have proved the following theorem.

Theorem 6.13

Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be bases for a vector space V. Let $B = [[\mathbf{u}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{u}_n]_{\mathcal{E}}]$ and $C = [[\mathbf{v}_1]_{\mathcal{E}} \ldots [\mathbf{v}_n]_{\mathcal{E}}]$, where $\mathcal{E}$ is any basis for V. Then row reduction applied to the $n \times 2n$ augmented matrix $[C \mid B]$ produces

$$[C \mid B] \rightarrow [I \mid P_{\mathcal{C}\leftarrow\mathcal{B}}]$$

If $\mathcal{E}$ is a standard basis, this method is particularly easy to use, since in that case $B = P_{\mathcal{E}\leftarrow\mathcal{B}}$ and $C = P_{\mathcal{E}\leftarrow\mathcal{C}}$. We illustrate this method by reworking the problem in Example 6.47.

Example 6.48

Rework Example 6.47 using the Gauss-Jordan method.

Solution Taking $\mathcal{E}$ to be the standard basis for M_{22}, we see that

$$B = P_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad C = P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Row reduction produces

$$[C \mid B] = \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

(Verify this row reduction.) It follows that

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

as we found before.

Exercises 6.3

In Exercises 1–4:
(a) Find the coordinate vectors $[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{C}}$ of $\mathbf{x}$ with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively.
(b) Find the change-of-basis matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$ from $\mathcal{B}$ to $\mathcal{C}$.
(c) Use your answer to part (b) to compute $[\mathbf{x}]_{\mathcal{C}}$, and compare your answer with the one found in part (a).
(d) Find the change-of-basis matrix $P_{\mathcal{B}\leftarrow\mathcal{C}}$ from $\mathcal{C}$ to $\mathcal{B}$.
(e) Use your answers to parts (c) and (d) to compute $[\mathbf{x}]_{\mathcal{B}}$, and compare your answer with the one found in part (a).

1. $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ in $\mathbb{R}^2$

2. $\mathbf{x} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ in $\mathbb{R}^2$

3. $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

4. $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

In Exercises 5–8, follow the instructions for Exercises 1–4 using $p(x)$ instead of $\mathbf{x}$.

5. $p(x) = 2 - x, \mathcal{B} = \{1, x\}, \mathcal{C} = \{x, 1 + x\}$ in $\mathcal{P}_1$

6. $p(x) = 1 + 3x, \mathcal{B} = \{1 + x, 1 - x\},$
$\mathcal{C} = \{2x, 4\}$ in $\mathcal{P}_1$

7. $p(x) = 1 + x^2, \mathcal{B} = \{1 + x + x^2, x + x^2, x^2\},$
$\mathcal{C} = \{1, x, x^2\}$ in $\mathcal{P}_2$

8. $p(x) = 4 - 2x - x^2, \mathcal{B} = \{x, 1 + x^2, x + x^2\},$
$\mathcal{C} = \{1, 1 + x, x^2\}$ in $\mathcal{P}_2$

In Exercises 9 and 10, follow the instructions for Exercises 1–4 using A instead of $\mathbf{x}$.

9. $A = \begin{bmatrix} 4 & 2 \\ 0 & -1 \end{bmatrix}$, $\mathcal{B} =$ the standard basis,
$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ in M_{22}

10. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$
$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ in M_{22}

In Exercises 11 and 12, follow the instructions for Exercises 1–4 using $f(x)$ instead of $\mathbf{x}$.

11. $f(x) = 2 \sin x - 3 \cos x, \mathcal{B} = \{\sin x + \cos x, \cos x\},$
$\mathcal{C} = \{\sin x + \cos x, \sin x - \cos x\}$ in span$(\sin x, \cos x)$

12. $f(x) = \sin x, \mathcal{B} = \{\sin x + \cos x, \cos x\},$
$\mathcal{C} = \{\cos x - \sin x, \sin x + \cos x\}$ in span$(\sin x, \cos x)$

13. Rotate the xy-axes in the plane counterclockwise through an angle $\theta = 60°$ to obtain new $x'y'$-axes. Use the methods of this section to find (a) the $x'y'$-coordinates of the point whose xy-coordinates are $(3, 2)$ and (b) the xy-coordinates of the point whose $x'y'$-coordinates are $(4, -4)$.

14. Repeat Exercise 13 with $\theta = 135°$.

15. Let $\mathcal{B}$ and $\mathcal{C}$ be bases for $\mathbb{R}^2$. If $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ and the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$ is

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

find $\mathcal{B}$.

16. Let $\mathcal{B}$ and $\mathcal{C}$ be bases for $\mathcal{P}_2$. If $\mathcal{B} = \{x, 1 + x, 1 - x + x^2\}$ and the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$ is

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

find $\mathcal{C}$.

In calculus, you learn that a ***Taylor polynomial of degree n about a*** *is a polynomial of the form*

$$p(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n$$

where $a_n \neq 0$. In other words, it is a polynomial that has been expanded in terms of powers of $x - a$ instead of powers of x. Taylor polynomials are very useful for approximating functions that are "well behaved" near $x = a$.

The set $\mathcal{B} = \{1, x - a, (x - a)^2, \ldots, (x - a)^n\}$ is a basis for $\mathscr{P}_n$ for any real number a. (Do you see a quick way to show this? Try using Theorem 6.7.) This fact allows us to use the techniques of this section to rewrite a polynomial as a Taylor polynomial about a given a.

17. Express $p(x) = 1 + 2x - 5x^2$ as a Taylor polynomial about $a = 1$.

18. Express $p(x) = 1 + 2x - 5x^2$ as a Taylor polynomial about $a = -2$.

19. Express $p(x) = x^3$ as a Taylor polynomial about $a = -1$.

20. Express $p(x) = x^3$ as a Taylor polynomial about $a = \frac{1}{2}$.

21. Let $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ be bases for a finite-dimensional vector space V. Prove that

$$P_{\mathcal{D}\leftarrow\mathcal{C}}P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{D}\leftarrow\mathcal{B}}$$

22. Let V be an n-dimensional vector space with basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Let P be an invertible $n \times n$ matrix and set

$$\mathbf{u}_i = p_{1i}\mathbf{v}_1 + \cdots + p_{ni}\mathbf{v}_n$$

for $i = 1, \ldots, n$. Prove that $\mathcal{C} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a basis for V and show that $P = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

6.4 Linear Transformations

We encountered linear transformations in Section 3.6 in the context of matrix transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. In this section, we extend this concept to linear transformations between arbitrary vector spaces.

Definition A ***linear transformation*** from a vector space V to a vector space W is a mapping $T: V \to W$ such that, for all $\mathbf{u}$ and $\mathbf{v}$ in V and for all scalars c,

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
2. $T(c\mathbf{u}) = cT(\mathbf{u})$

It is straightforward to show that this definition is equivalent to the requirement that T preserve all linear combinations. That is,

$T: V \to W$ is a linear transformation if and only if

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_kT(\mathbf{v}_k)$$

for all $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in V and scalars $c_1, \ldots, c_k$.

Example 6.49

Every matrix transformation is a linear transformation. That is, if A is an $m \times n$ matrix, then the transformation $T_A: \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_A(\mathbf{x}) = A\mathbf{x} \quad \text{for } \mathbf{x} \text{ in } \mathbb{R}^n$$

is a linear transformation. This is a restatement of Theorem 3.30.

Example 6.50

Define $T : M_{nn} \to M_{nn}$ by $T(A) = A^T$. Show that T is a linear transformation.

Solution We check that, for A and B in M_{nn} and scalars c,

$$T(A + B) = (A + B)^T = A^T + B^T = T(A) + T(B)$$

and

$$T(cA) = (cA)^T = cA^T = cT(A)$$

Therefore, T is a linear transformation.

Example 6.51

Let D be the ***differential operator*** $D : \mathscr{D} \to \mathscr{F}$ defined by $D(f) = f'$. Show that D is a linear transformation.

Solution Let f and g be differentiable functions and let c be a scalar. Then, from calculus, we know that

$$D(f + g) = (f + g)' = f' + g' = D(f) + D(g)$$

and

$$D(cf) = (cf)' = cf' = cD(f)$$

Hence, D is a linear transformation.

In calculus, you learn that every continuous function on $[a, b]$ is integrable. The next example shows that integration is a linear transformation.

Example 6.52

Define $S : \mathscr{C}[a, b] \to \mathbb{R}$ by $S(f) = \int_a^b f(x)\,dx$. Show that S is a linear transformation.

Solution Let f and g be in $\mathscr{C}[a, b]$. Then

$$\begin{aligned} S(f + g) &= \int_a^b (f + g)(x)\,dx \\ &= \int_a^b (f(x) + g(x))\,dx \\ &= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \\ &= S(f) + S(g) \end{aligned}$$

and

$$\begin{aligned} S(cf) &= \int_a^b (cf)(x)\,dx \\ &= \int_a^b cf(x)\,dx \\ &= c\int_a^b f(x)\,dx \\ &= cS(f) \end{aligned}$$

It follows that S is linear.

Example 6.53

Show that none of the following transformations is linear:

(a) $T : M_{22} \to \mathbb{R}$ defined by $T(A) = \det A$

(b) $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 2^x$

(c) $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = x + 1$

Solution In each case, we give a specific counterexample to show that one of the properties of a linear transformation fails to hold.

(a) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Then $A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, so

$$T(A + B) = \det(A + B) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

But

$$T(A) + T(B) = \det A + \det B = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} = 0 + 0 = 0$$

so $T(A + B) \neq T(A) + T(B)$ and T is not linear.

(b) Let $x = 1$ and $y = 2$. Then

$$T(x + y) = T(3) = 2^3 = 8 \neq 6 = 2^1 + 2^2 = T(x) + T(y)$$

so T is not linear.

(c) Let $x = 1$ and $y = 2$. Then

$$T(x + y) = T(3) = 3 + 1 = 4 \neq 5 = (1 + 1) + (2 + 1) = T(x) + T(y)$$

Therefore, T is not linear.

Remark Example 6.53(c) shows that you need to be careful when you encounter the word "linear." As a *function*, $T(x) = x + 1$ is linear, since its graph is a straight line. However, it is not a *linear transformation* from the vector space $\mathbb{R}$ to itself, since it fails to satisfy the definition. (Which linear functions from $\mathbb{R}$ to $\mathbb{R}$ will also be linear transformations?)

There are two special linear transformations that deserve to be singled out.

Example 6.54

(a) For any vector spaces V and W, the transformation $T_0 : V \to W$ that maps every vector in V to the zero vector in W is called the ***zero transformation.*** That is,

$$T_0(\mathbf{v}) = \mathbf{0} \quad \text{for all } \mathbf{v} \text{ in } V$$

(b) For any vector space V, the transformation $I : V \to V$ that maps every vector in V to itself is called the ***identity transformation.*** That is,

$$I(\mathbf{v}) = \mathbf{v} \quad \text{for all } \mathbf{v} \text{ in } V$$

(If it is important to identify the vector space V, we may write I_V for clarity.) The proofs that the zero and identity transformations are linear are left as easy exercises.

Properties of Linear Transformations

In Chapter 3, all linear transformations were matrix transformations, and their properties were directly related to properties of the matrices involved. The following theorem is easy to prove for matrix transformations. (Do it!) The full proof for linear transformations in general takes a bit more care, but it is still straightforward.

Theorem 6.14 Let $T : V \to W$ be a linear transformation. Then:

a. $T(\mathbf{0}) = \mathbf{0}$
b. $T(-\mathbf{v}) = -T(\mathbf{v})$ for all $\mathbf{v}$ in V.
c. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for all $\mathbf{u}$ and $\mathbf{v}$ in V.

Proof We prove properties (a) and (c) and leave the proof of property (b) for Exercise 21.

(a) Let $\mathbf{v}$ be any vector in V. Then $T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}$, as required. (Can you give a reason for each step?)

(c) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u} + (-1)\mathbf{v}) = T(\mathbf{u}) + (-1)T(\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$

Remark Property (a) can be useful in showing that certain transformations are *not* linear. As an illustration, consider Example 6.53(b). If $T(x) = 2^x$, then $T(0) = 2^0 = 1 \neq 0$, so T is not linear, by Theorem 6.14(a). Be warned, however, that there are lots of transformations that *do* map the zero vector to the zero vector but that are still *not* linear. Example 6.53(a) is a case in point: The zero vector is the 2×2 zero matrix O, so $T(O) = \det O = 0$, but we have seen that $T(A) = \det A$ is not linear.

The most important property of a linear transformation $T : V \to W$ is that T is completely determined by its effect on a basis for V. The next example shows what this means.

Example 6.55 Suppose T is a linear transformation from $\mathbb{R}^2$ to $\mathscr{P}_2$ such that

$$T\begin{bmatrix}1\\1\end{bmatrix} = 2 - 3x + x^2 \quad \text{and} \quad T\begin{bmatrix}2\\3\end{bmatrix} = 1 - x^2$$

Find $T\begin{bmatrix}-1\\2\end{bmatrix}$ and $T\begin{bmatrix}a\\b\end{bmatrix}$.

Solution Since $\mathcal{B} = \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}\right\}$ is a basis for $\mathbb{R}^2$ (why?), every vector in $\mathbb{R}^2$ is in span$(\mathcal{B})$. Solving

$$c_1\begin{bmatrix}1\\1\end{bmatrix} + c_2\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}-1\\2\end{bmatrix}$$

we find that $c_1 = -7$ and $c_2 = 3$. Therefore,

$$\begin{aligned} T\begin{bmatrix} -1 \\ 2 \end{bmatrix} &= T\left(-7\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= -7T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3T\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= -7(2 - 3x + x^2) + 3(1 - x^2) \\ &= -11 + 21x - 10x^2 \end{aligned}$$

Similarly, we discover that

$$\begin{bmatrix} a \\ b \end{bmatrix} = (3a - 2b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

so

$$\begin{aligned} T\begin{bmatrix} a \\ b \end{bmatrix} &= T\left((3a - 2b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= (3a - 2b)T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)T\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= (3a - 2b)(2 - 3x + x^2) + (b - a)(1 - x^2) \\ &= (5a - 3b) + (-9a + 6b)x + (4a - 3b)x^2 \end{aligned}$$

(Note that by setting $a = -1$ and $b = 2$, we recover the solution $T\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -11 + 21x - 10x^2$.)

The proof of the general theorem is quite straightforward.

Theorem 6.15 Let $T : V \to W$ be a linear transformation and let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a spanning set for V. Then $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ spans the range of T.

Proof The range of T is the set of all vectors in W that are of the form $T(\mathbf{v})$, where $\mathbf{v}$ is in V. Let $T(\mathbf{v})$ be in the range of T. Since $\mathcal{B}$ spans V, there are scalars $c_1, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Applying T and using the fact that it is a linear transformation, we see that

$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$$

In other words, $T(\mathbf{v})$ is in span$(T(\mathcal{B}))$, as required.

Theorem 6.15 applies, in particular, when $\mathcal{B}$ is a basis for V. You might guess that, in this case, $T(\mathcal{B})$ would then be a basis for the range of T. Unfortunately, this is not always the case. We will address this issue in Section 6.5.

Composition of Linear Transformations

In Section 3.6, we defined the composition of matrix transformations. The definition extends to general linear transformations in an obvious way.

$S \circ T$ is read "S of T."

Definition If $T : U \to V$ and $S : V \to W$ are linear transformations, then the ***composition of S with T*** is the mapping $S \circ T$, defined by

$$(S \circ T)(\mathbf{u}) = S(T(\mathbf{u}))$$

where $\mathbf{u}$ is in U.

Observe that $S \circ T$ is a mapping from U to W (see Figure 6.6). Notice also that for the definition to make sense, the range of T must be contained in the domain of S.

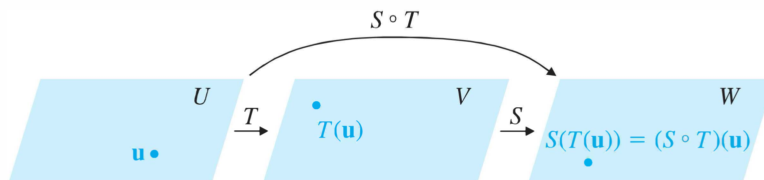


Figure 6.6
Composition of linear transformations

Example 6.56

Let $T : \mathbb{R}^2 \to \mathscr{P}_1$ and $S : \mathscr{P}_1 \to \mathscr{P}_2$ be the linear transformations defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x \quad \text{and} \quad S(p(x)) = xp(x)$$

Find $(S \circ T)\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ and $(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}$.

Solution We compute

$$(S \circ T)\begin{bmatrix} 3 \\ -2 \end{bmatrix} = S\left(T\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) = S(3 + (3 - 2)x) = S(3 + x) = x(3 + x)$$
$$= 3x + x^2$$

and

$$(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix} = S\left(T\begin{bmatrix} a \\ b \end{bmatrix}\right) = S(a + (a + b)x) = x(a + (a + b)x)$$
$$= ax + (a + b)x^2$$

Chapter 3 showed that the composition of two matrix transformations was another matrix transformation. In general, we have the following theorem.

Theorem 6.16 If $T : U \to V$ and $S : V \to W$ are linear transformations, then $S \circ T : U \to W$ is a linear transformation.

Proof Let $\mathbf{u}$ and $\mathbf{v}$ be in U and let c be a scalar. Then

$$\begin{aligned}(S \circ T)(\mathbf{u} + \mathbf{v}) &= S(T(\mathbf{u} + \mathbf{v}))\\ &= S(T(\mathbf{u}) + T(\mathbf{v})) && \text{since } T \text{ is linear}\\ &= S(T(\mathbf{u})) + S(T(\mathbf{v})) && \text{since } S \text{ is linear}\\ &= (S \circ T)(\mathbf{u}) + (S \circ T)(\mathbf{v})\end{aligned}$$

and

$$\begin{aligned}(S \circ T)(c\mathbf{u}) &= S(T(c\mathbf{u}))\\ &= S(cT(\mathbf{u})) && \text{since } T \text{ is linear}\\ &= cS(T(\mathbf{u})) && \text{since } S \text{ is linear}\\ &= c(S \circ T)(\mathbf{u})\end{aligned}$$

Therefore, $S \circ T$ is a linear transformation.

The algebraic properties of linear transformations mirror those of matrix transformations, which, in turn, are related to the algebraic properties of matrices. For example, composition of linear transformations is associative. That is, if R, S, and T are linear transformations, then

$$R \circ (S \circ T) = (R \circ S) \circ T$$

provided these compositions make sense. The proof of this property is identical to that given in Section 3.6.

The next example gives another useful (but not surprising) property of linear transformations.

Example 6.57

Let $S : U \to V$ and $T : V \to W$ be linear transformations and let $I : V \to V$ be the identity transformation. Then for every $\mathbf{v}$ in V, we have

$$(T \circ I)(\mathbf{v}) = T(I(\mathbf{v})) = T(\mathbf{v})$$

Since $T \circ I$ and T have the same value at every $\mathbf{v}$ in their domain, it follows that $T \circ I = T$. Similarly, $I \circ S = S$.

Remark The method of Example 6.57 is worth noting. Suppose we want to show that two linear transformations T_1 and T_2 (both from V to W) are equal. It suffices to show that $T_1(\mathbf{v}) = T_2(\mathbf{v})$ for every $\mathbf{v}$ in V.

Further properties of linear transformations are explored in the exercises.

Inverses of Linear Transformations

Definition A linear transformation $T : V \to W$ is ***invertible*** if there is a linear transformation $T' : W \to V$ such that

$$T' \circ T = I_V \quad \text{and} \quad T \circ T' = I_W$$

In this case, T' is called an ***inverse*** for T.

Remarks

- The domain V and codomain W of T do not have to be the same, as they do in the case of invertible matrix transformations. However, we will see in the next section that V and W must be very closely related.
- The requirement that T' be linear could have been omitted from this definition. For, as we will see in Theorem 6.24, if T' is *any* mapping from W to V such that $T' \circ T = I_V$ and $T \circ T' = I_W$, then T' is forced to be linear as well.
- If T' is an inverse for T, then the definition implies that T is an inverse for T'. Hence, T' is invertible too.

Example 6.58

Verify that the mappings $T: \mathbb{R}^2 \to \mathcal{P}_1$ and $T': \mathcal{P}_1 \to \mathbb{R}^2$ defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x \quad \text{and} \quad T'(c + dx) = \begin{bmatrix} c \\ d - c \end{bmatrix}$$

are inverses.

Solution We compute

$$(T' \circ T)\begin{bmatrix} a \\ b \end{bmatrix} = T'\left(T\begin{bmatrix} a \\ b \end{bmatrix}\right) = T'(a + (a + b)x) = \begin{bmatrix} a \\ (a + b) - a \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

and

$$(T \circ T')(c + dx) = T(T'(c + dx)) = T\begin{bmatrix} c \\ d - c \end{bmatrix} = c + (c + (d - c))x = c + dx$$

Hence, $T' \circ T = I_{\mathbb{R}^2}$ and $T \circ T' = I_{\mathcal{P}_1}$. Therefore, T and T' are inverses of each other.

As was the case for invertible matrices, inverses of linear transformations are unique if they exist. The following theorem is the analogue of Theorem 3.6.

Theorem 6.17 If T is an invertible linear transformation, then its inverse is unique.

Proof The proof is the same as that of Theorem 3.6, with products of matrices replaced by compositions of linear transformations. (You are asked to complete this proof in Exercise 31.)

Thanks to Theorem 6.17, if T is invertible, we can refer to *the* inverse of T. It will be denoted by T^{-1} (pronounced "T inverse"). In the next two sections, we will address the issue of determining when a given linear transformation is invertible and finding its inverse when it exists.

Exercises 6.4

In Exercises 1–12, determine whether T is a linear transformation.

1. $T: M_{22} \to M_{22}$ defined by

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & c+d \end{bmatrix}$$

2. $T: M_{22} \to M_{22}$ defined by

$$T\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & w-z \\ x-y & 1 \end{bmatrix}$$

3. $T: M_{nn} \to M_{nn}$ defined by $T(A) = AB$, where B is a fixed $n \times n$ matrix

4. $T: M_{nn} \to M_{nn}$ defined by $T(A) = AB - BA$, where B is a fixed $n \times n$ matrix

5. $T: M_{nn} \to \mathbb{R}$ defined by $T(A) = \operatorname{tr}(A)$

6. $T: M_{nn} \to \mathbb{R}$ defined by $T(A) = a_{11}a_{22}\cdots a_{nn}$

7. $T: M_{nn} \to \mathbb{R}$ defined by $T(A) = \operatorname{rank}(A)$

8. $T: \mathscr{P}_2 \to \mathscr{P}_2$ defined by $T(a + bx + cx^2) = (a+1) + (b+1)x + (c+1)x^2$

9. $T: \mathscr{P}_2 \to \mathscr{P}_2$ defined by $T(a + bx + cx^2) = a + b(x+1) + b(x+1)^2$

10. $T: \mathscr{F} \to \mathscr{F}$ defined by $T(f) = f(x^2)$

11. $T: \mathscr{F} \to \mathscr{F}$ defined by $T(f) = (f(x))^2$

12. $T: \mathscr{F} \to \mathbb{R}$ defined by $T(f) = f(c)$, where c is a fixed scalar

13. Show that the transformations S and T in Example 6.56 are both linear.

14. Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation for which

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \text{and} \quad T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$$

Find $T\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $T\begin{bmatrix} a \\ b \end{bmatrix}$.

15. Let $T: \mathbb{R}^2 \to \mathscr{P}_2$ be a linear transformation for which

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 - 2x \quad \text{and} \quad T\begin{bmatrix} 3 \\ -1 \end{bmatrix} = x + 2x^2$$

Find $T\begin{bmatrix} -7 \\ 9 \end{bmatrix}$ and $T\begin{bmatrix} a \\ b \end{bmatrix}$.

16. Let $T: \mathscr{P}_2 \to \mathscr{P}_2$ be a linear transformation for which

$$T(1) = 3 - 2x, \quad T(x) = 4x - x^2, \quad \text{and} \quad T(x^2) = 2 + 2x^2$$

Find $T(6 + x - 4x^2)$ and $T(a + bx + cx^2)$.

17. Let $T: \mathscr{P}_2 \to \mathscr{P}_2$ be a linear transformation for which

$$T(1 + x) = 1 + x^2, \quad T(x + x^2) = x - x^2,$$
$$T(1 + x^2) = 1 + x + x^2$$

Find $T(4 - x + 3x^2)$ and $T(a + bx + cx^2)$.

18. Let $T: M_{22} \to \mathbb{R}$ be a linear transformation for which

$$T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 1, \quad T\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 2,$$
$$T\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = 3, \quad T\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 4$$

Find $T\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$ and $T\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

19. Let $T: M_{22} \to \mathbb{R}$ be a linear transformation. Show that there are scalars a, b, c, and d such that

$$T\begin{bmatrix} w & x \\ y & z \end{bmatrix} = aw + bx + cy + dz$$

for all $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ in M_{22}.

20. Show that there is no linear transformation $T: \mathbb{R}^3 \to \mathscr{P}_2$ such that

$$T\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 1 + x, \quad T\begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} = 2 - x + x^2,$$
$$T\begin{bmatrix} 0 \\ 6 \\ -8 \end{bmatrix} = -2 + 2x^2$$

21. Prove Theorem 6.14(b).

22. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V and let $T: V \to V$ be a linear transformation. Prove that if $T(\mathbf{v}_1) = \mathbf{v}_1, T(\mathbf{v}_2) = \mathbf{v}_2 \ldots, T(\mathbf{v}_n) = \mathbf{v}_n$, then T is the identity transformation on V.

23. Let $T: \mathscr{P}_n \to \mathscr{P}_n$ be a linear transformation such that $T(x^k) = kx^{k-1}$ for $k = 0, 1, \ldots, n$. Show that T must be the differential operator D.

24. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a vector space V and let $T: V \to W$ be a linear transformation.

(a) If $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is linearly independent in W, show that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent in V.

(b) Show that the converse of part (a) is false. That is, it is not necessarily true that if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent in V, then $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is linearly independent in W. Illustrate this with an example $T: \mathbb{R}^2 \to \mathbb{R}^2$.

25. Define linear transformations $S: \mathbb{R}^2 \to M_{22}$ and $T: \mathbb{R}^2 \to \mathbb{R}^2$ by

$$S\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b & b \\ 0 & a-b \end{bmatrix} \quad \text{and} \quad T\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 2c+d \\ -d \end{bmatrix}$$

Compute $(S \circ T)\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $(S \circ T)\begin{bmatrix} x \\ y \end{bmatrix}$. Can you compute $(T \circ S)\begin{bmatrix} x \\ y \end{bmatrix}$? If so, compute it.

26. Define linear transformations $S: \mathcal{P}_1 \to \mathcal{P}_2$ and $T: \mathcal{P}_2 \to \mathcal{P}_1$ by

$$S(a + bx) = a + (a+b)x + 2bx^2$$

and
$$T(a + bx + cx^2) = b + 2cx$$

Compute $(S \circ T)(3 + 2x - x^2)$ and $(S \circ T)(a + bx + cx^2)$. Can you compute $(T \circ S)(a + bx)$? If so, compute it.

27. Define linear transformations $S: \mathcal{P}_n \to \mathcal{P}_n$ and $T: \mathcal{P}_n \to \mathcal{P}_n$ by

$$S(p(x)) = p(x+1) \quad \text{and} \quad T(p(x)) = p'(x)$$

Find $(S \circ T)(p(x))$ and $(T \circ S)(p(x))$. [*Hint:* Remember the Chain Rule.]

28. Define linear transformations $S: \mathcal{P}_n \to \mathcal{P}_n$ and $T: \mathcal{P}_n \to \mathcal{P}_n$ by

$$S(p(x)) = p(x+1) \quad \text{and} \quad T(p(x)) = xp'(x)$$

Find $(S \circ T)(p(x))$ and $(T \circ S)(p(x))$.

In Exercises 29 and 30, verify that S and T are inverses.

29. $S: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x + y \\ 3x + y \end{bmatrix}$ and $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ -3x + 4y \end{bmatrix}$

30. $S: \mathcal{P}_1 \to \mathcal{P}_1$ defined by $S(a + bx) = (-4a + b) + 2ax$ and $T: \mathcal{P}_1 \to \mathcal{P}_1$ defined by

$$T(a + bx) = b/2 + (a + 2b)x$$

31. Prove Theorem 6.17.

32. Let $T: V \to V$ be a linear transformation such that $T \circ T = I$.

(a) Show that $\{\mathbf{v}, T(\mathbf{v})\}$ is linearly dependent if and only if $T(\mathbf{v}) = \pm\mathbf{v}$.

(b) Give an example of such a linear transformation with $V = \mathbb{R}^2$.

33. Let $T: V \to V$ be a linear transformation such that $T \circ T = T$.

(a) Show that $\{\mathbf{v}, T(\mathbf{v})\}$ is linearly dependent if and only if $T(\mathbf{v}) = \mathbf{v}$ or $T(\mathbf{v}) = \mathbf{0}$.

(b) Give an example of such a linear transformation with $V = \mathbb{R}^2$.

*The set of all linear transformations from a vector space V to a vector space W is denoted by $\mathcal{L}(V, W)$. If S and T are in $\mathcal{L}(V, W)$, we can define the **sum** $S + T$ of S and T by*

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$$

*for all $\mathbf{v}$ in V. If c is a scalar, we define the **scalar multiple** cT of T by c to be*

$$(cT)(\mathbf{v}) = cT(\mathbf{v})$$

for all $\mathbf{v}$ in V. Then $S + T$ and cT are both transformations from V to W.

34. Prove that $S + T$ and cT are linear transformations.

35. Prove that $\mathcal{L}(V, W)$ is a vector space with this addition and scalar multiplication.

36. Let R, S, and T be linear transformations such that the following operations make sense. Prove that:

(a) $R \circ (S + T) = R \circ S + R \circ T$

(b) $c(R \circ S) = (cR) \circ S = R \circ (cS)$ for any scalar c

6.5 The Kernel and Range of a Linear Transformation

The null space and column space are two of the fundamental subspaces associated with a matrix. In this section, we extend these notions to the kernel and range of a linear transformation.

The word *kernel* is derived from the Old English word *cyrnel*, a form of the word *corn*, meaning “seed” or “grain.” Like a kernel of corn, the kernel of a linear transformation is its “core” or “seed” in the sense that it carries information about many of the important properties of the transformation.

Definition Let $T: V \to W$ be a linear transformation. The ***kernel*** of T, denoted $\ker(T)$, is the set of all vectors in V that are mapped by T to $\mathbf{0}$ in W. That is,

$$\ker(T) = \{\mathbf{v} \text{ in } V: T(\mathbf{v}) = \mathbf{0}\}$$

The ***range*** of T, denoted $\text{range}(T)$, is the set of all vectors in W that are images of vectors in V under T. That is,

$$\begin{aligned} \text{range}(T) &= \{T(\mathbf{v}) : \mathbf{v} \text{ in } V\} \\ &= \{\mathbf{w} \text{ in } W : \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \text{ in } V\} \end{aligned}$$

Example 6.59

Let A be an $m \times n$ matrix and let $T = T_A$ be the corresponding matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $T(\mathbf{v}) = A\mathbf{v}$. Then, as we saw in Chapter 3, the range of T is the column space of A.

The kernel of T is

$$\begin{aligned} \ker(T) &= \{\mathbf{v} \text{ in } \mathbb{R}^n : T(\mathbf{v}) = \mathbf{0}\} \\ &= \{\mathbf{v} \text{ in } \mathbb{R}^n : A\mathbf{v} = \mathbf{0}\} \\ &= \text{null}(A) \end{aligned}$$

In words, the kernel of a matrix transformation is just the null space of the corresponding matrix.

Example 6.60

Find the kernel and range of the differential operator $D : \mathscr{P}_3 \to \mathscr{P}_2$ defined by $D(p(x)) = p'(x)$.

Solution Since $D(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2$, we have

$$\begin{aligned} \ker(D) &= \{a + bx + cx^2 + dx^3 : D(a + bx + cx^2 + dx^3) = 0\} \\ &= \{a + bx + cx^2 + dx^3 : b + 2cx + 3dx^2 = 0\} \end{aligned}$$

But $b + 2cx + 3dx^2 = 0$ if and only if $b = 2c = 3d = 0$, which implies that $b = c = d = 0$. Therefore,

$$\begin{aligned} \ker(D) &= \{a + bx + cx^2 + dx^3 : b = c = d = 0\} \\ &= \{a : a \text{ in } \mathbb{R}\} \end{aligned}$$

In other words, the kernel of D is the set of constant polynomials.

The range of D is all of $\mathscr{P}_2$, since *every* polynomial in $\mathscr{P}_2$ is the image under D (i.e., the derivative) of *some* polynomial in $\mathscr{P}_3$. To be specific, if $a + bx + cx^2$ is in $\mathscr{P}_2$, then

$$a + bx + cx^2 = D\left(ax + \left(\frac{b}{2}\right)x^2 + \left(\frac{c}{3}\right)x^3\right)$$

Example 6.61

Let $S : \mathscr{P}_1 \to \mathbb{R}$ be the linear transformation defined by

$$S(p(x)) = \int_0^1 p(x)\, dx$$

Find the kernel and range of S.

Solution In detail, we have

$$\begin{aligned} S(a + bx) &= \int_0^1 (a + bx)\, dx \\ &= \left[ax + \frac{b}{2}x^2 \right]_0^1 \\ &= \left(a + \frac{b}{2} \right) - 0 = a + \frac{b}{2} \end{aligned}$$

Therefore,

$$\begin{aligned} \ker(S) &= \{a + bx : S(a + bx) = 0\} \\ &= \left\{ a + bx : a + \frac{b}{2} = 0 \right\} \\ &= \left\{ a + bx : a = -\frac{b}{2} \right\} \\ &= \left\{ -\frac{b}{2} + bx \right\} \end{aligned}$$

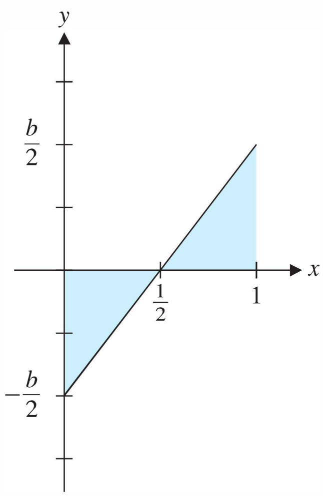


Figure 6.7

If $y = -\frac{b}{2} + bx$, then $\int_0^1 y\, dx = 0$

Geometrically, $\ker(S)$ consists of all those linear polynomials whose graphs have the property that the area between the line and the x-axis is equally distributed above and below the axis on the interval $[0, 1]$ (see Figure 6.7).

The range of S is $\mathbb{R}$, since every real number can be obtained as the image under S of some polynomial in $\mathscr{P}_1$. For example, if a is an arbitrary real number, then

$$\int_0^1 a\, dx = [ax]_0^1 = a - 0 = a$$

so $a = S(a)$.

Example 6.62

Let $T : M_{22} \to M_{22}$ be the linear transformation defined by taking transposes: $T(A) = A^T$. Find the kernel and range of T.

Solution We see that

$$\begin{aligned} \ker(T) &= \{A \text{ in } M_{22} : T(A) = O\} \\ &= \{A \text{ in } M_{22} : A^T = O\} \end{aligned}$$

But if $A^T = O$, then $A = (A^T)^T = O^T = O$. It follows that $\ker(T) = \{O\}$.

Since, for any matrix A in M_{22}, we have $A = (A^T)^T = T(A^T)$ (and A^T is in M_{22}), we deduce that $\text{range}(T) = M_{22}$.

In all of these examples, the kernel and range of a linear transformation are subspaces of the domain and codomain, respectively, of the transformation. Since we are generalizing the null space and column space of a matrix, this is perhaps not surprising. Nevertheless, we should not take anything for granted, so we need to prove that it is not a coincidence.

Theorem 6.18 Let $T: V \to W$ be a linear transformation. Then:

a. The kernel of T is a subspace of V.
b. The range of T is a subspace of W.

Proof (a) Since $T(\mathbf{0}) = \mathbf{0}$, the zero vector of V is in $\ker(T)$, so $\ker(T)$ is nonempty. Let $\mathbf{u}$ and $\mathbf{v}$ be in $\ker(T)$ and let c be a scalar. Then $T(\mathbf{u}) = T(\mathbf{v}) = \mathbf{0}$, so

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

and

$$T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{0} = \mathbf{0}$$

Therefore, $\mathbf{u} + \mathbf{v}$ and $c\mathbf{u}$ are in $\ker(T)$, and $\ker(T)$ is a subspace of V.

(b) Since $\mathbf{0} = T(\mathbf{0})$, the zero vector of W is in $\text{range}(T)$, so $\text{range}(T)$ is nonempty. Let $T(\mathbf{u})$ and $T(\mathbf{v})$ be in the range of T and let c be a scalar. Then $T(\mathbf{u}) + T(\mathbf{v}) = T(\mathbf{u} + \mathbf{v})$ is the image of the vector $\mathbf{u} + \mathbf{v}$. Since $\mathbf{u}$ and $\mathbf{v}$ are in V, so is $\mathbf{u} + \mathbf{v}$, and hence $T(\mathbf{u}) + T(\mathbf{v})$ is in range (T). Similarly, $cT(\mathbf{u}) = T(c\mathbf{u})$. Since $\mathbf{u}$ is in V, so is $c\mathbf{u}$, and hence $cT(\mathbf{u})$ is in $\text{range}(T)$. Therefore, $\text{range}(T)$ is a nonempty subset of W that is closed under addition and scalar multiplication, and thus it is a subspace of W.

Figure 6.8 gives a schematic representation of the kernel and range of a linear transformation.

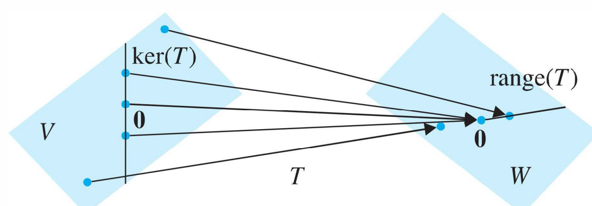


Figure 6.8
The kernel and range of $T: V \to W$

In Chapter 3, we defined the rank of a matrix to be the dimension of its column space and the nullity of a matrix to be the dimension of its null space. We now extend these definitions to linear transformations.

Definition Let $T: V \to W$ be a linear transformation. The ***rank*** of T is the dimension of the range of T and is denoted by $\text{rank}(T)$. The ***nullity*** of T is the dimension of the kernel of T and is denoted by $\text{nullity}(T)$.

Example 6.63 If A is a matrix and $T = T_A$ is the matrix transformation defined by $T(\mathbf{v}) = A\mathbf{v}$, then the range and kernel of T are the column space and the null space of A, respectively, by Example 6.59. Hence, from Section 3.5, we have

$$\text{rank}(T) = \text{rank}(A) \quad \text{and} \quad \text{nullity}(T) = \text{nullity}(A)$$

Example 6.64

Find the rank and the nullity of the linear transformation $D : \mathscr{P}_3 \to \mathscr{P}_2$ defined by $D(p(x)) = p'(x)$.

Solution In Example 6.60, we computed range $(D) = \mathscr{P}_2$, so

$$\text{rank}(D) = \dim \mathscr{P}_2 = 3$$

The kernel of D is the set of all constant polynomials: $\ker(D) = \{a : a \text{ in } \mathbb{R}\} = \{a \cdot 1 : a \text{ in } \mathbb{R}\}$. Hence, $\{1\}$ is a basis for $\ker(D)$, so

$$\text{nullity}(D) = \dim(\ker(D)) = 1$$

Example 6.65

Find the rank and the nullity of the linear transformation $S : \mathscr{P}_1 \to \mathbb{R}$ defined by

$$S(p(x)) = \int_0^1 p(x)\, dx$$

Solution From Example 6.61, $\text{range}(S) = \mathbb{R}$ and $\text{rank}(S) = \dim \mathbb{R} = 1$. Also,

$$\begin{aligned}\ker(S) &= \left\{-\frac{b}{2} + bx : b \text{ in } \mathbb{R}\right\}\\ &= \{b(-\tfrac{1}{2} + x) : b \text{ in } \mathbb{R}\}\\ &= \text{span}(-\tfrac{1}{2} + x)\end{aligned}$$

so $\{-\frac{1}{2} + x\}$ is a basis for $\ker(S)$. Therefore, $\text{nullity}(S) = \dim(\ker(S)) = 1$.

Example 6.66

Find the rank and the nullity of the linear transformation $T : M_{22} \to M_{22}$ defined by $T(A) = A^T$.

Solution In Example 6.62, we found that $\text{range}(T) = M_{22}$ and $\ker(T) = \{O\}$. Hence,

$$\text{rank}(T) = \dim M_{22} = 4 \quad \text{and} \quad \text{nullity}(T) = \dim\{O\} = 0$$

In Chapter 3, we saw that the rank and nullity of an $m \times n$ matrix A are related by the formula $\text{rank}(A) + \text{nullity}(A) = n$. This is the Rank Theorem (Theorem 3.26). Since the matrix transformation $T = T_A$ has $\mathbb{R}^n$ as its domain, we could rewrite the relationship as

$$\text{rank}(A) + \text{nullity}(A) = \dim \mathbb{R}^n$$

This version of the Rank Theorem extends very nicely to general linear transformations, as you can see from the last three examples:

$$\text{rank}(D) + \text{nullity}(D) = 3 + 1 = 4 = \dim \mathscr{P}_3 \qquad \text{Example 6.64}$$

$$\text{rank}(S) + \text{nullity}(S) = 1 + 1 = 2 = \dim \mathscr{P}_1 \qquad \text{Example 6.65}$$

$$\text{rank}(T) + \text{nullity}(T) = 4 + 0 = 4 = \dim M_{22} \qquad \text{Example 6.66}$$

Theorem 6.19 **The Rank Theorem**

Let $T : V \to W$ be a linear transformation from a finite-dimensional vector space V into a vector space W. Then

$$\text{rank}(T) + \text{nullity}(T) = \dim V$$

In the next section, you will see how to adapt the proof of Theorem 3.26 to prove this version of the result. For now, we give an alternative proof that does not use matrices.

Proof Let $\dim V = n$ and let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for $\ker(T)$[so that $\text{nullity}(T) = \dim(\ker(T)) = k$]. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent set, it can be extended to a basis for V, by Theorem 6.28. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ be such a basis. If we can show that the set $\mathcal{C} = \{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\}$ is a basis for $\text{range}(T)$, then we will have $\text{rank}(T) = \dim(\text{range}(T)) = n - k$ and thus

$$\text{rank}(T) + \text{nullity}(T) = k + (n - k) = n = \dim V$$

as required.

Certainly $\mathcal{C}$ is contained in the range of T. To show that $\mathcal{C}$ spans the range of T, let $T(\mathbf{v})$ be a vector in the range of T. Then $\mathbf{v}$ is in V, and since $\mathcal{B}$ is a basis for V, we can find scalars $c_1, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are in the kernel of T, we have $T(\mathbf{v}_1) = \cdots = T(\mathbf{v}_k) = \mathbf{0}$, so

$$\begin{aligned} T(\mathbf{v}) &= T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n) \\ &= c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) + c_{k+1}T(\mathbf{v}_{k+1}) + \cdots + c_nT(\mathbf{v}_n) \\ &= c_{k+1}T(\mathbf{v}_{k+1}) + \cdots + c_nT(\mathbf{v}_n) \end{aligned}$$

This shows that the range of T is spanned by $\mathcal{C}$.

To show that $\mathcal{C}$ is linearly independent, suppose that there are scalars $c_{k+1}, \ldots, c_n$ such that

$$c_{k+1}T(\mathbf{v}_{k+1}) + \cdots + c_nT(\mathbf{v}_n) = \mathbf{0}$$

Then $T(c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n) = \mathbf{0}$, which means that $c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n$ is in the kernel of T and is, hence, expressible as a linear combination of the basis vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of $\ker(T)$—say,

$$c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

But now $$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k - c_{k+1}\mathbf{v}_{k+1} - \cdots - c_n\mathbf{v}_n = \mathbf{0}$$

and the linear independence of $\mathcal{B}$ forces $c_1 = \cdots = c_n = 0$. In particular, $c_{k+1} = \cdots = c_n = 0$, which means $\mathcal{C}$ is linearly independent.

We have shown that $\mathcal{C}$ is a basis for the range of T, so, by our comments above, the proof is complete.

We have verified the Rank Theorem for Examples 6.64, 6.65, and 6.66. In practice, this theorem allows us to find the rank and nullity of a linear transformation with only half the work. The following examples illustrate the process.

Example 6.67

Find the rank and nullity of the linear transformation $T: \mathscr{P}_2 \to \mathscr{P}_3$ defined by $T(p(x)) = xp(x)$. (Check that T really is linear.)

Solution In detail, we have

$$T(a + bx + cx^2) = ax + bx^2 + cx^3$$

It follows that

$$\begin{aligned}\ker(T) &= \{a + bx + cx^2 : T(a + bx + cx^2) = 0\}\\ &= \{a + bx + cx^2 : ax + bx^2 + cx^3 = 0\}\\ &= \{a + bx + cx^2 : a = b = c = 0\}\\ &= \{0\}\end{aligned}$$

so we have $\text{nullity}(T) = \dim(\ker(T)) = 0$. The Rank Theorem implies that

$$\text{rank}(T) = \dim \mathscr{P}_2 - \text{nullity}(T) = 3 - 0 = 3$$

Remark In Example 6.67, it would be just as easy to find the rank of T first, since $\{x, x^2, x^3\}$ is easily seen to be a basis for the range of T. Usually, though, one of the two (the rank or the nullity of a linear transformation) will be easier to compute; the Rank Theorem can then be used to find the other. With practice, you will become better at knowing which way to proceed.

Example 6.68

Let W be the vector space of all symmetric 2×2 matrices. Define a linear transformation $T: W \to \mathscr{P}_2$ by

$$T\begin{bmatrix} a & b \\ b & c \end{bmatrix} = (a - b) + (b - c)x + (c - a)x^2$$

(Check that T is linear.) Find the rank and nullity of T.

Solution The nullity of T is easier to compute directly than the rank, so we proceed as follows:

$$\begin{aligned}\ker(T) &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : T\begin{bmatrix} a & b \\ b & c \end{bmatrix} = 0\right\}\\ &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : (a - b) + (b - c)x + (c - a)x^2 = 0\right\}\\ &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : (a - b) = (b - c) = (c - a) = 0\right\}\\ &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : a = b = c\right\}\\ &= \left\{\begin{bmatrix} c & c \\ c & c \end{bmatrix}\right\} = \text{span}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)\end{aligned}$$

Therefore, $\left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ is a basis for the kernel of T, so $\text{nullity}(T) = \dim(\ker(T)) = 1$. The Rank Theorem and Example 6.42 tell us that $\text{rank}(T) = \dim W - \text{nullity}(T) = 3 - 1 = 2$.

One-to-One and Onto Linear Transformations

We now investigate criteria for a linear transformation to be invertible. The keys to the discussion are the very important properties one-to-one and onto.

Definition A linear transformation $T : V \to W$ is called ***one-to-one*** if T maps distinct vectors in V to distinct vectors in W. If range$(T) = W$, then T is called ***onto.***

Remarks

- The definition of one-to-one may be written more formally as follows:

$T : V \to W$ is one-to-one if, for all $\mathbf{u}$ and $\mathbf{v}$ in V,

$$\mathbf{u} \neq \mathbf{v} \text{ implies that } T(\mathbf{u}) \neq T(\mathbf{v})$$

The above statement is equivalent to the following:

$T : V \to W$ is one-to-one if, for all $\mathbf{u}$ and $\mathbf{v}$ in V,

$$T(\mathbf{u}) = T(\mathbf{v}) \text{ implies that } \mathbf{u} = \mathbf{v}$$

Figure 6.9 illustrates these two statements.

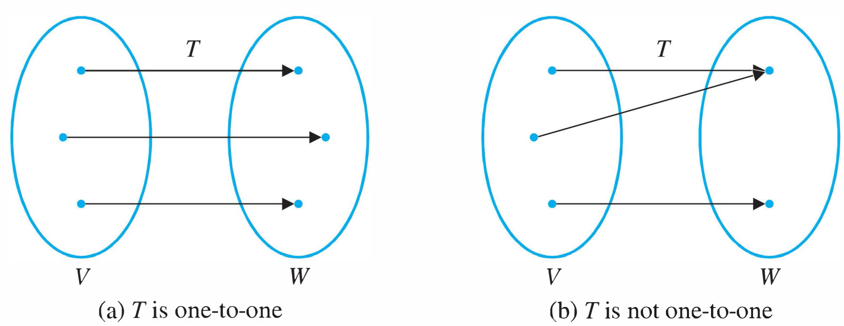


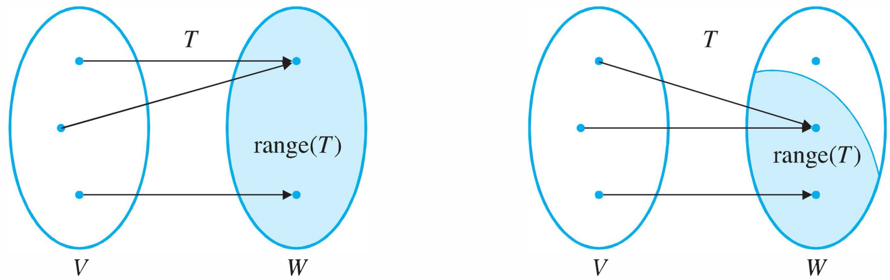
Figure 6.9

- Another way to write the definition of onto is as follows:

$T : V \to W$ is onto if, for all $\mathbf{w}$ in W, there is at least one $\mathbf{v}$ in V such that

$$\mathbf{w} = T(\mathbf{v})$$

In other words, *given* $\mathbf{w}$ in W, does there exist some $\mathbf{v}$ in V such that $\mathbf{w} = T(\mathbf{v})$? If, for an arbitrary $\mathbf{w}$, we can solve this equation for $\mathbf{v}$, then T is onto (see Figure 6.10).



(a) T is onto (b) T is not onto

Figure 6.10

Example 6.69

Which of the following linear transformations are one-to-one? onto?

(a) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ x - y \\ 0 \end{bmatrix}$

(b) $D : \mathscr{P}_3 \to \mathscr{P}_2$ defined by $D(p(x)) = p'(x)$

(c) $T : M_{22} \to M_{22}$ defined by $T(A) = A^T$

Solution (a) Let $T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$. Then

$$\begin{bmatrix} 2x_1 \\ x_1 - y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 - y_2 \\ 0 \end{bmatrix}$$

so $2x_1 = 2x_2$ and $x_1 - y_1 = x_2 - y_2$. Solving these equations, we see that $x_1 = x_2$ and $y_1 = y_2$. Hence, $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, so T is one-to-one.

T is not onto, since its range is not all of $\mathbb{R}^3$. To be specific, there is no vector $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ such that $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. (Why not?)

(b) In Example 6.60, we showed that range$(D) = \mathscr{P}_2$, so D is onto. D is not one-to-one, since distinct polynomials in $\mathscr{P}_3$ can have the same derivative. For example, $x^3 \neq x^3 + 1$, but $D(x^3) = 3x^2 = D(x^3 + 1)$.

(c) Let A and B be in M_{22}, with $T(A) = T(B)$. Then $A^T = B^T$, so $A = (A^T)^T = (B^T)^T = B$. Hence, T is one-to-one. In Example 6.62, we showed that range$(T) = M_{22}$. Hence, T is onto.

It turns out that there is a very simple criterion for determining whether a linear transformation is one-to-one.

Theorem 6.20

A linear transformation $T : V \to W$ is one-to-one if and only if $\ker(T) = \{\mathbf{0}\}$.

Proof Assume that T is one-to-one. If $\mathbf{v}$ is in the kernel of T, then $T(\mathbf{v}) = \mathbf{0}$. But we also know that $T(\mathbf{0}) = \mathbf{0}$, so $T(\mathbf{v}) = T(\mathbf{0})$. Since T is one-to-one, this implies that $\mathbf{v} = \mathbf{0}$, so the only vector in the kernel of T is the zero vector.

Conversely, assume that $\ker(T) = \{\mathbf{0}\}$. To show that T is one-to-one, let $\mathbf{u}$ and $\mathbf{v}$ be in V with $T(\mathbf{u}) = T(\mathbf{v})$. Then $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0}$, which implies that $\mathbf{u} - \mathbf{v}$ is in the kernel of T. But $\ker(T) = \{\mathbf{0}\}$, so we must have $\mathbf{u} - \mathbf{v} = \mathbf{0}$ or, equivalently, $\mathbf{u} = \mathbf{v}$. This proves that T is one-to-one.

Example 6.70

Show that the linear transformation $T : \mathbb{R}^2 \to \mathscr{P}_1$ defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

is one-to-one and onto.

Solution If $\begin{bmatrix} a \\ b \end{bmatrix}$ is in the kernel of T, then

$$0 = T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

It follows that $a = 0$ and $a + b = 0$. Hence, $b = 0$, and therefore $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Consequently, $\ker(T) = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$, and T is one-to-one, by Theorem 6.20.

By the Rank Theorem,

$$\operatorname{rank}(T) = \dim \mathbb{R}^2 - \operatorname{nullity}(T) = 2 - 0 = 2$$

Therefore, the range of T is a two-dimensional subspace of $\mathbb{R}^2$, and hence $\operatorname{range}(T) = \mathbb{R}^2$. It follows that T is onto.

For linear transformations between two n-dimensional vector spaces, the properties of one-to-one and onto are closely related. Observe first that for a linear transformation $T : V \to W$, $\ker(T) = \{\mathbf{0}\}$ if and only if $\operatorname{nullity}(T) = 0$, and T is onto if and only if $\operatorname{rank}(T) = \dim W$. (Why?) The proof of the next theorem essentially uses the method of Example 6.70.

Theorem 6.21

Let $\dim V = \dim W = n$. Then a linear transformation $T : V \to W$ is one-to-one if and only if it is onto.

Proof Assume that T is one-to-one. Then $\operatorname{nullity}(T) = 0$ by Theorem 6.20 and the remark preceding Theorem 6.21. The Rank Theorem implies that

$$\operatorname{rank}(T) = \dim V - \operatorname{nullity}(T) = n - 0 = n$$

Therefore, T is onto.

Conversely, assume that T is onto. Then $\operatorname{rank}(T) = \dim W = n$. By the Rank Theorem,

$$\operatorname{nullity}(T) = \dim V - \operatorname{rank}(T) = n - n = 0$$

Hence, $\ker(T) = \{\mathbf{0}\}$, and T is one-to-one.

In Section 6.4, we pointed out that if $T: V \to W$ is a linear transformation, then the image of a basis for V under T need not be a basis for the range of T. We can now give a condition that ensures that a basis for V will be mapped by T to a basis for W.

Theorem 6.22 Let $T: V \to W$ be a one-to-one linear transformation. If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent set in V, then $T(S) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ is a linearly independent set in W.

Proof Let $c_1, \ldots, c_k$ be scalars such that

$$c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) = \mathbf{0}$$

Then $T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = \mathbf{0}$, which implies that $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ is in the kernel of T. But, since T is one-to-one, $\ker(T) = \{\mathbf{0}\}$, by Theorem 6.20. Hence,

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

But, since $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent, all of the scalars c_i must be 0. Therefore, $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ is linearly independent.

Corollary 6.23 Let $\dim V = \dim W = n$. Then a one-to-one linear transformation $T: V \to W$ maps a basis for V to a basis for W.

Proof Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V. By Theorem 6.22, $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is a linearly independent set in W, so we need only show that $T(\mathcal{B})$ spans W. But, by Theorem 6.15, $T(\mathcal{B})$ spans the range of T. Moreover, T is onto, by Theorem 6.21, so $\text{range}(T) = W$. Therefore, $T(\mathcal{B})$ spans W, which completes the proof.

Example 6.71 Let $T: \mathbb{R}^2 \to \mathcal{P}_1$ be the linear transformation from Example 6.70, defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

Then, by Corollary 6.23, the standard basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$ for $\mathbb{R}^2$ is mapped to a basis $T(\mathcal{E}) = \{T(\mathbf{e}_1), T(\mathbf{e}_2)\}$ of $\mathcal{P}_1$. We find that

$$T(\mathbf{e}_1) = T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 + x \quad \text{and} \quad T(\mathbf{e}_2) = T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = x$$

It follows that $\{1 + x, x\}$ is a basis for $\mathcal{P}_1$.

We can now determine which linear transformations $T: V \to W$ are invertible.

Theorem 6.24 A linear transformation $T: V \to W$ is invertible if and only if it is one-to-one and onto.

Proof Assume that T is invertible. Then there exists a linear transformation $T^{-1}: W \to V$ such that

$$T^{-1} \circ T = I_V \quad \text{and} \quad T \circ T^{-1} = I_W$$

To show that T is one-to-one, let $\mathbf{v}$ be in the kernel of T. Then $T(\mathbf{v}) = \mathbf{0}$. Therefore,

$$\begin{aligned} T^{-1}(T(\mathbf{v})) = T^{-1}(\mathbf{0}) &\Rightarrow (T^{-1} \circ T)(\mathbf{v}) = \mathbf{0} \\ &\Rightarrow I(\mathbf{v}) = \mathbf{0} \\ &\Rightarrow \mathbf{v} = \mathbf{0} \end{aligned}$$

which establishes that $\ker(T) = \{\mathbf{0}\}$. Therefore, T is one-to-one, by Theorem 6.20.

To show that T is onto, let $\mathbf{w}$ be in W and let $\mathbf{v} = T^{-1}(\mathbf{w})$. Then

$$\begin{aligned} T(\mathbf{v}) &= T(T^{-1}(\mathbf{w})) \\ &= (T \circ T^{-1})(\mathbf{w}) \\ &= I(\mathbf{w}) \\ &= \mathbf{w} \end{aligned}$$

which shows that $\mathbf{w}$ is the image of $\mathbf{v}$ under T. Since $\mathbf{v}$ is in V, this shows that T is onto.

Conversely, assume that T is one-to-one and onto. This means that $\text{nullity}(T) = 0$ and $\text{rank}(T) = \dim W$. We need to show that there exists a linear transformation $T': W \to V$ such that $T' \circ T = I_V$ and $T \circ T' = I_W$.

Let $\mathbf{w}$ be in W. Since T is onto, there exists some vector $\mathbf{v}$ in V such that $T(\mathbf{v}) = \mathbf{w}$. There is only one such vector $\mathbf{v}$, since, if $\mathbf{v}'$ is another vector in V such that $T(\mathbf{v}') = \mathbf{w}$, then $T(\mathbf{v}) = T(\mathbf{v}')$; the fact that T is one-to-one then implies that $\mathbf{v} = \mathbf{v}'$. It therefore makes sense to define a mapping $T': W \to V$ by setting $T'(\mathbf{w}) = \mathbf{v}$.

It follows that

$$(T' \circ T)(\mathbf{v}) = T'(T(\mathbf{v})) = T'(\mathbf{w}) = \mathbf{v}$$

and

$$(T \circ T')(\mathbf{w}) = T(T'(\mathbf{w})) = T(\mathbf{v}) = \mathbf{w}$$

It then follows that $T' \circ T = I_V$ and $T \circ T' = I_W$. Now we must show that T' is a *linear* transformation.

To this end, let $\mathbf{w}_1$ and $\mathbf{w}_2$ be in W and let c_1 and c_2 be scalars. As above, let $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$. Then $\mathbf{v}_1 = T'(\mathbf{w}_1)$ and $\mathbf{v}_2 = T'(\mathbf{w}_2)$ and

$$\begin{aligned} T'(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) &= T'(c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)) \\ &= T'(T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) \\ &= I(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) \\ &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \\ &= c_1T'(\mathbf{w}_1) + c_2T'(\mathbf{w}_2) \end{aligned}$$

Consequently, T' is linear, so, by Theorem 6.17, $T' = T^{-1}$.

The words *isomorphism* and *isomorphic* are derived from the Greek words *isos*, meaning "equal," and *morph*, meaning "shape." Thus, figuratively speaking, isomorphic vector spaces have "equal shapes."

Isomorphisms of Vector Spaces

We now are in a position to describe, in concrete terms, what it means for two vector spaces to be "essentially the same."

Definition A linear transformation $T: V \to W$ is called an ***isomorphism*** if it is one-to-one and onto. If V and W are two vector spaces such that there is an isomorphism from V to W, then we say that V is ***isomorphic*** to W and write $V \cong W$.

Example 6.72

Show that $\mathcal{P}_{n-1}$ and $\mathbb{R}^n$ are isomorphic.

Solution The process of forming the coordinate vector of a polynomial provides us with one possible isomorphism (as we observed already in Section 6.2, although we did not use the term *isomorphism* there). Specifically, define $T: \mathcal{P}_{n-1} \to \mathbb{R}^n$ by $T(p(x)) = [p(x)]_{\mathcal{E}}$, where $\mathcal{E} = \{1, x, \ldots, x^{n-1}\}$ is the standard basis for $\mathcal{P}_{n-1}$. That is,

$$T(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix}$$

Theorem 6.6 shows that T is a linear transformation. If $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ is in the kernel of T, then

$$\begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix} = T(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Hence, $a_0 = a_1 = \cdots = a_{n-1} = 0$, so $p(x) = 0$. Therefore, $\ker(T) = \{0\}$, and T is one-to-one. Since $\dim \mathcal{P}_{n-1} = \dim \mathbb{R}^n = n$, T is also onto, by Theorem 6.21. Thus, T is an isomorphism, and $\mathcal{P}_{n-1} \cong \mathbb{R}^n$.

Example 6.73

Show that M_{mn} and $\mathbb{R}^{mn}$ are isomorphic.

Solution Once again, the coordinate mapping from M_{mn} to $\mathbb{R}^{mn}$ (as in Example 6.36) is an isomorphism. The details of the proof are left as an exercise.

In fact, the easiest way to tell if two vector spaces are isomorphic is simply to check their dimensions, as the next theorem shows.

Theorem 6.25 Let V and W be two finite-dimensional vector spaces (over the same field of scalars). Then V is isomorphic to W if and only if $\dim V = \dim W$.

Proof Let $n = \dim V$. If V is isomorphic to W, then there is an isomorphism $T : V \to W$. Since T is one-to-one, $\text{nullity}(T) = 0$. The Rank Theorem then implies that

$$\text{rank}(T) = \dim V - \text{nullity}(T) = n - 0 = n$$

Therefore, the range of T is an n-dimensional subspace of W. But, since T is onto, $W = \text{range}(T)$, so $\dim W = n$, as we wished to show.

Conversely, assume that V and W have the same dimension, n. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V and let $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be a basis for W. We will define a linear transformation $T : V \to W$ and then show that T is one-to-one and onto. An arbitrary vector $\mathbf{v}$ in V can be written uniquely as a linear combination of the vectors in the basis $\mathcal{B}$—say,

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

We define T by

$$T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$$

It is straightforward to check that T is linear. (Do so.) To see that T is one-to-one, suppose $\mathbf{v}$ is in the kernel of T. Then

$$c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n = T(\mathbf{v}) = \mathbf{0}$$

and the linear independence of $\mathcal{C}$ forces $c_1 = \cdots = c_n = 0$. But then

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

so $\ker(T) = \{\mathbf{0}\}$, meaning that T is one-to-one. Since $\dim V = \dim W$, T is also onto, by Theorem 6.21. Therefore, T is an isomorphism, and $V \cong W$.

Example 6.74 Show that $\mathbb{R}^n$ and $\mathcal{P}_n$ are not isomorphic.

Solution Since $\dim \mathbb{R}^n = n \neq n + 1 = \dim \mathcal{P}_n$, $\mathbb{R}^n$ and $\mathcal{P}_n$ are not isomorphic, by Theorem 6.25.

Example 6.75 Let W be the vector space of all symmetric 2×2 matrices. Show that W is isomorphic to $\mathbb{R}^3$.

Solution In Example 6.42, we showed that $\dim W = 3$. Hence, $\dim W = \dim \mathbb{R}^3$, so $W \cong \mathbb{R}^3$, by Theorem 6.25. (There is an obvious candidate for an isomorphism $T : W \to \mathbb{R}^3$. What is it?)

Remark Our examples have all been *real* vector spaces, but the theorems we have proved are true for vector spaces over the complex numbers $\mathbb{C}$ or $\mathbb{Z}_p$, where p is prime. For example, the vector space $M_{22}(\mathbb{Z}_2)$ of all 2×2 matrices with entries from $\mathbb{Z}_2$ has dimension 4 as a vector space over $\mathbb{Z}_2$, and hence $M_{22}(\mathbb{Z}_2) \cong \mathbb{Z}_2^4$.

Exercises 6.5

1. Let $T: M_{22} \to M_{22}$ be the linear transformation defined by

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

(a) Which, if any, of the following matrices are in $\ker(T)$?

(i) $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 0 & 4 \\ 2 & 0 \end{bmatrix}$ **(iii)** $\begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$

(b) Which, if any, of the matrices in part (a) are in $\text{range}(T)$?

(c) Describe $\ker(T)$ and $\text{range}(T)$.

2. Let $T: M_{22} \to \mathbb{R}$ be the linear transformation defined by $T(A) = \text{tr}(A)$.

(a) Which, if any, of the following matrices are in $\ker(T)$?

(i) $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 0 & 4 \\ 2 & 0 \end{bmatrix}$ **(iii)** $\begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix}$

(b) Which, if any, of the following scalars are in $\text{range}(T)$?

(i) 0 **(ii)** 2 **(iii)** $\sqrt{2}/2$

(c) Describe $\ker(T)$ and $\text{range}(T)$.

3. Let $T: \mathscr{P}_2 \to \mathbb{R}^2$ be the linear transformation defined by

$$T(a + bx + cx^2) = \begin{bmatrix} a - b \\ b + c \end{bmatrix}$$

(a) Which, if any, of the following polynomials are in $\ker(T)$?

(i) $1 + x$ **(ii)** $x - x^2$ **(iii)** $1 + x - x^2$

(b) Which, if any, of the following vectors are in $\text{range}(T)$?

(i) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ **(iii)** $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(c) Describe $\ker(T)$ and $\text{range}(T)$.

4. Let $T: \mathscr{P}_2 \to \mathscr{P}_2$ be the linear transformation defined by $T(p(x)) = xp'(x)$.

(a) Which, if any, of the following polynomials are in $\ker(T)$?

(i) 1 **(ii)** x **(iii)** x^2

(b) Which, if any, of the polynomials in part (a) are in $\text{range}(T)$?

(c) Describe $\ker(T)$ and $\text{range}(T)$.

In Exercises 5–8, find bases for the kernel and range of the linear transformations T in the indicated exercises. In each case, state the nullity and rank of T and verify the Rank Theorem.

5. Exercise 1

6. Exercise 2

7. Exercise 3

8. Exercise 4

In Exercises 9–14, find either the nullity or the rank of T and then use the Rank Theorem to find the other.

9. $T: M_{22} \to \mathbb{R}^2$ defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a - b \\ c - d \end{bmatrix}$

10. $T: \mathscr{P}_2 \to \mathbb{R}^2$ defined by $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$

11. $T: M_{22} \to M_{22}$ defined by $T(A) = AB$, where $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

12. $T: M_{22} \to M_{22}$ defined by $T(A) = AB - BA$, where $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

13. $T: \mathscr{P}_2 \to \mathbb{R}$ defined by $T(p(x)) = p'(0)$

14. $T: M_{33} \to M_{33}$ defined by $T(A) = A - A^T$

In Exercises 15–20, determine whether the linear transformation T is (a) one-to-one and (b) onto.

15. $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - y \\ x + 2y \end{bmatrix}$

16. $T : \mathbb{R}^2 \to \mathscr{P}_2$ defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = (a - 2b) + (3a + b)x + (a + b)x^2$$

17. $T : \mathscr{P}_2 \to \mathbb{R}^3$ defined by

$$T(a + bx + cx^2) = \begin{bmatrix} 2a - b \\ a + b - 3c \\ c - a \end{bmatrix}$$

18. $T : \mathscr{P}_2 \to \mathbb{R}^2$ defined by $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$

19. $T : \mathbb{R}^3 \to \mathbb{M}_{22}$ defined by $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a - b & b - c \\ a + b & b + c \end{bmatrix}$

20. $T : \mathbb{R}^3 \to W$ defined by $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} =$

$\begin{bmatrix} a + b + c & b - 2c \\ b - 2c & a - c \end{bmatrix}$, where W is the vector space of all symmetric 2×2 matrices

In Exercises 21–26, determine whether V and W are isomorphic. If they are, give an explicit isomorphism $T : V \to W$.

21. $V = D_3$ (diagonal 3×3 matrices), $W = \mathbb{R}^3$

22. $V = S_3$ (symmetric 3×3 matrices), $W = U_3$ (upper triangular 3×3 matrices)

23. $V = S_3$ (symmetric 3×3 matrices), $W = S_3'$ (skew-symmetric 3×3 matrices)

24. $V = \mathscr{P}_2$, $W = \{p(x) \text{ in } \mathscr{P}_3 : p(0) = 0\}$

25. $V = \mathbb{C}$, $W = \mathbb{R}^2$

26. $V = \{A \text{ in } M_{22} : \text{tr}(A) = 0\}$, $W = \mathbb{R}^2$

27. Show that $T : \mathscr{P}_n \to \mathscr{P}_n$ defined by $T(p(x)) = p(x) + p'(x)$ is an isomorphism.

28. Show that $T : \mathscr{P}_n \to \mathscr{P}_n$ defined by $T(p(x)) = p(x - 2)$ is an isomorphism.

29. Show that $T : \mathscr{P}_n \to \mathscr{P}_n$ defined by $T(p(x)) = x^n p\left(\dfrac{1}{x}\right)$ is an isomorphism.

30. **(a)** Show that $\mathscr{C}[0, 1] \cong \mathscr{C}[2, 3]$. [*Hint:* Define $T : \mathscr{C}[0, 1] \to \mathscr{C}[2, 3]$ by letting $T(f)$ be the function whose value at x is $(T(f))(x) = f(x - 2)$ for x in $[2, 3]$.]

(b) Show that $\mathscr{C}[0, 1] \cong \mathscr{C}[a, a + 1]$ for all a.

31. Show that $\mathscr{C}[0, 1] \cong \mathscr{C}[0, 2]$.

32. Show that $\mathscr{C}[a, b] \cong \mathscr{C}[c, d]$ for all $a < b$ and $c < d$.

33. Let $S : V \to W$ and $T : U \to V$ be linear transformations.

(a) Prove that if S and T are both one-to-one, so is $S \circ T$.

(b) Prove that if S and T are both onto, so is $S \circ T$.

34. Let $S : V \to W$ and $T : U \to V$ be linear transformations.

(a) Prove that if $S \circ T$ is one-to-one, so is T.

(b) Prove that if $S \circ T$ is onto, so is S.

35. Let $T : V \to W$ be a linear transformation between two finite-dimensional vector spaces.

(a) Prove that if $\dim V < \dim W$, then T cannot be onto.

(b) Prove that if $\dim V > \dim W$, then T cannot be one-to-one.

36. Let $a_0, a_1, \ldots, a_n$ be $n + 1$ distinct real numbers. Define $T : \mathscr{P}_n \to \mathbb{R}^{n+1}$ by

$$T(p(x)) = \begin{bmatrix} p(a_0) \\ p(a_1) \\ \vdots \\ p(a_n) \end{bmatrix}$$

Prove that T is an isomorphism.

37. If V is a finite-dimensional vector space and $T : V \to V$ is a linear transformation such that $\text{rank}(T) = \text{rank}(T^2)$, prove that $\text{range}(T) \cap \ker(T) = \{\mathbf{0}\}$. [*Hint:* T^2 denotes $T \circ T$. Use the Rank Theorem to help show that the kernels of T and T^2 are the same.]

38. Let U and W be subspaces of a finite-dimensional vector space V. Define $T : U \times W \to V$ by $T(\mathbf{u}, \mathbf{w}) = \mathbf{u} - \mathbf{w}$.

(a) Prove that T is a linear transformation.

(b) Show that $\text{range}(T) = U + W$.

(c) Show that $\ker(T) \cong U \cap W$. [*Hint:* See Exercise 50 in Section 6.1.]

(d) Prove ***Grassmann's Identity:***

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

[*Hint:* Apply the Rank Theorem, using results (a) and (b) and Exercise 43(b) in Section 6.2.]

6.6 The Matrix of a Linear Transformation

Theorem 6.15 showed that a linear transformation $T : V \to W$ is completely determined by its effect on a spanning set for V. In particular, if we know how T acts on a basis for V, then we can compute $T(\mathbf{v})$ for any vector $\mathbf{v}$ in V. Example 6.55 illustrated the process. We implicitly used this important property of linear transformations in Theorem 3.31 to help us compute the standard matrix of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$. In this section, we will show that every linear transformation between finite-dimensional vector spaces can be represented as a matrix transformation.

Suppose that V is an n-dimensional vector space, W is an m-dimensional vector space, and $T : V \to W$ is a linear transformation. Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Then the coordinate vector mapping $R(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$ defines an isomorphism $R : V \to \mathbb{R}^n$. At the same time, we have an isomorphism $S : W \to \mathbb{R}^m$ given by $S(\mathbf{w}) = [\mathbf{w}]_{\mathcal{C}}$, which allows us to associate the image $T(\mathbf{v})$ with the vector $[T(\mathbf{v})]_{\mathcal{C}}$ in $\mathbb{R}^m$. Figure 6.11 illustrates the relationships.

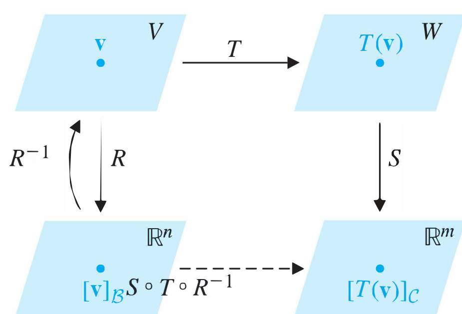


Figure 6.11

Since R is an isomorphism, it is invertible, so we may form the composite mapping

$$S \circ T \circ R^{-1} : \mathbb{R}^n \to \mathbb{R}^m$$

which maps $[\mathbf{v}]_{\mathcal{B}}$ to $[T(\mathbf{v})]_{\mathcal{C}}$. Since this mapping goes from $\mathbb{R}^n$ to $\mathbb{R}^m$, we know from Chapter 3 that it is a matrix transformation. What, then, is the standard matrix of $S \circ T \circ R^{-1}$? We would like to find the $m \times n$ matrix A such that $A[\mathbf{v}]_{\mathcal{B}} = (S \circ T \circ R^{-1})([\mathbf{v}]_{\mathcal{B}})$. Or, since $(S \circ T \circ R^{-1})([\mathbf{v}]_{\mathcal{B}}) = [T(\mathbf{v})]_{\mathcal{C}}$, we require

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

It turns out to be surprisingly easy to find. The basic idea is that of Theorem 3.31. The columns of A are the images of the standard basis vectors for $\mathbb{R}^n$ under $S \circ T \circ R^{-1}$. But, if $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V, then

$$\begin{aligned} R(\mathbf{v}_i) &= [\mathbf{v}_i]_{\mathcal{B}} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \leftarrow i\text{th entry} \\ &= \mathbf{e}_i \end{aligned}$$

so $R^{-1}(\mathbf{e}_i) = \mathbf{v}_i$. Therefore, the ith column of the matrix A we seek is given by

$$\begin{aligned}(S \circ T \circ R^{-1})(\mathbf{e}_i) &= S(T(R^{-1}(\mathbf{e}_i))) \\ &= S(T(\mathbf{v}_i)) \\ &= [T(\mathbf{v}_i)]_{\mathcal{C}}\end{aligned}$$

which is the coordinate vector of $T(\mathbf{v}_i)$ with respect to the basis $\mathcal{C}$ of W.

We summarize this discussion as a theorem.

Theorem 6.26 Let V and W be two finite-dimensional vector spaces with bases $\mathcal{B}$ and $\mathcal{C}$, respectively, where $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. If $T : V \to W$ is a linear transformation, then the $m \times n$ matrix A defined by

$$A = [[T(\mathbf{v}_1)]_{\mathcal{C}} \mid [T(\mathbf{v}_2)]_{\mathcal{C}} \mid \cdots \mid [T(\mathbf{v}_n)]_{\mathcal{C}}]$$

satisfies

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

for every vector $\mathbf{v}$ in V.

The matrix A in Theorem 6.26 is called the ***matrix of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$.*** The relationship is illustrated below. (Recall that T_A denotes multiplication by A.)

$$\begin{array}{ccc} \mathbf{v} & \xrightarrow{T} & T(\mathbf{v}) \\ \downarrow & & \downarrow \\ [\mathbf{v}]_{\mathcal{B}} & \xrightarrow{T_A} & A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}} \end{array}$$

Remarks

- The matrix of a linear transformation T with respect to bases $\mathcal{B}$ and $\mathcal{C}$ is sometimes denoted by $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$. Note the direction of the arrow: right-to-left (not left-to-right, as for $T: V \to W$). With this notation, the final equation in Theorem 6.26 becomes

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

Observe that the $\mathcal{B}$s in the subscripts appear side by side and appear to "cancel" each other. In words, this equation says, "The matrix for T times the coordinate vector for $\mathbf{v}$ gives the coordinate vector for $T(\mathbf{v})$."

In the special case where $V = W$ and $\mathcal{B} = \mathcal{C}$, we write $[T]_{\mathcal{B}}$ (instead of $[T]_{\mathcal{B}\leftarrow\mathcal{B}}$). Theorem 6.26 then states that

$$[T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$$

- The matrix of a linear transformation with respect to given bases is unique. That is, for every vector $\mathbf{v}$ in V, there is only *one* matrix A with the property specified by Theorem 6.26—namely,

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

(You are asked to prove this in Exercise 39.)

- The diagram that follows Theorem 6.26 is sometimes called a *commutative diagram* because we can start in the upper left-hand corner with the vector $\mathbf{v}$ and get to $[T(\mathbf{v})]_{\mathcal{C}}$ in the lower right-hand corner in two different, but equivalent, ways. If, as before, we denote the coordinate mappings that map $\mathbf{v}$ to $[\mathbf{v}]_{\mathcal{B}}$ and $\mathbf{w}$ to $[\mathbf{w}]_{\mathcal{C}}$ by R and S, respectively, then we can summarize this "commutativity" by

$$S \circ T = T_A \circ R$$

The reason for the term *commutative* becomes clearer when $V = W$ and $\mathcal{B} = \mathcal{C}$, for then $R = S$ too, and we have

$$R \circ T = T_A \circ R$$

suggesting that the coordinate mapping R commutes with the linear transformation T (provided we use the matrix version of T—namely, $T_A = T_{[T]_{\mathcal{B}}}$—where it is required).

- The matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ depends on the *order* of the vectors in the bases $\mathcal{B}$ and $\mathcal{C}$. Rearranging the vectors within either basis will affect the matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$. [See Example 6.77(b).]

Example 6.76

Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - 2y \\ x + y - 3z \end{bmatrix}$$

and let $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\mathcal{C} = \{\mathbf{e}_2, \mathbf{e}_1\}$ be bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Find the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ and verify Theorem 6.26 for $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$.

Solution First, we compute

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad T(\mathbf{e}_3) = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

Next, we need their coordinate vectors with respect to $\mathcal{C}$. Since

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{e}_2 + \mathbf{e}_1, \quad \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \mathbf{e}_2 - 2\mathbf{e}_1, \quad \begin{bmatrix} 0 \\ -3 \end{bmatrix} = -3\mathbf{e}_2 + 0\mathbf{e}_1$$

we have

$$[T(\mathbf{e}_1)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad [T(\mathbf{e}_2)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad [T(\mathbf{e}_3)]_{\mathcal{C}} = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

Therefore, the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ is

$$\begin{aligned} A = [T]_{\mathcal{C}\leftarrow\mathcal{B}} &= [\,[T(\mathbf{e}_1)]_{\mathcal{C}} \;\; [T(\mathbf{e}_2)]_{\mathcal{C}} \;\; [T(\mathbf{e}_3)]_{\mathcal{C}}\,] \\ &= \begin{bmatrix} 1 & 1 & -3 \\ 1 & -2 & 0 \end{bmatrix} \end{aligned}$$

To verify Theorem 6.26 for $\mathbf{v}$, we first compute

$$T(\mathbf{v}) = T\begin{bmatrix}1\\3\\-2\end{bmatrix} = \begin{bmatrix}-5\\10\end{bmatrix}$$

Then

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix}1\\3\\-2\end{bmatrix}$$

and

$$[T(\mathbf{v})]_{\mathcal{C}} = \begin{bmatrix}-5\\10\end{bmatrix}_{\mathcal{C}} = \begin{bmatrix}10\\-5\end{bmatrix}$$

(Check these.)

Using all of these facts, we confirm that

$$A[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix}1 & 1 & -3\\1 & -2 & 0\end{bmatrix}\begin{bmatrix}1\\3\\-2\end{bmatrix} = \begin{bmatrix}10\\-5\end{bmatrix} = [T(\mathbf{v})]_{\mathcal{C}}$$

Example 6.77

Let $D : \mathscr{P}_3 \to \mathscr{P}_2$ be the differential operator $D(p(x)) = p'(x)$. Let $\mathcal{B} = \{1, x, x^2, x^3\}$ and $\mathcal{C} = \{1, x, x^2\}$ be bases for $\mathscr{P}_3$ and $\mathscr{P}_2$, respectively.

(a) Find the matrix A of D with respect to $\mathcal{B}$ and $\mathcal{C}$.

(b) Find the matrix A' of D with respect to $\mathcal{B}'$ and $\mathcal{C}$, where $\mathcal{B}' = \{x^3, x^2, x, 1\}$.

(c) Using part (a), compute $D(5 - x + 2x^3)$ and $D(a + bx + cx^2 + dx^3)$ to verify Theorem 6.26.

Solution First note that $D(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2$. (See Example 6.60.)

(a) Since the images of the basis $\mathcal{B}$ under D are $D(1) = 0$, $D(x) = 1$, $D(x^2) = 2x$, and $D(x^3) = 3x^2$, their coordinate vectors with respect to $\mathcal{C}$ are

$$[D(1)]_{\mathcal{C}} = \begin{bmatrix}0\\0\\0\end{bmatrix}, \quad [D(x)]_{\mathcal{C}} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad [D(x^2)]_{\mathcal{C}} = \begin{bmatrix}0\\2\\0\end{bmatrix}, \quad [D(x^3)]_{\mathcal{C}} = \begin{bmatrix}0\\0\\3\end{bmatrix}$$

Consequently,

$$A = [D]_{\mathcal{C}\leftarrow\mathcal{B}} = [\,[D(1)]_{\mathcal{C}} \mid [D(x)]_{\mathcal{C}} \mid [D(x^2)]_{\mathcal{C}} \mid [D(x^3)]_{\mathcal{C}}\,]$$
$$= \begin{bmatrix}0 & 1 & 0 & 0\\0 & 0 & 2 & 0\\0 & 0 & 0 & 3\end{bmatrix}$$

(b) Since the basis $\mathcal{B}'$ is just $\mathcal{B}$ in the *reverse* order, we see that

$$A' = [D]_{\mathcal{C}\leftarrow\mathcal{B}'} = [\,[D(x^3)]_{\mathcal{C}} \mid [D(x^2)]_{\mathcal{C}} \mid [D(x)]_{\mathcal{C}} \mid [D(1)]_{\mathcal{C}}\,]$$
$$= \begin{bmatrix}0 & 0 & 1 & 0\\0 & 2 & 0 & 0\\3 & 0 & 0 & 0\end{bmatrix}$$

(This shows that the *order* of the vectors in the bases $\mathcal{B}$ and $\mathcal{C}$ affects the matrix of a transformation with respect to these bases.)

(c) First we compute $D(5 - x + 2x^3) = -1 + 6x^2$ directly, getting the coordinate vector

$$[D(5 - x + 2x^3)]_{\mathcal{C}} = [-1 + 6x^2]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix}$$

On the other hand,

$$[5 - x + 2x^3]_{\mathcal{B}} = \begin{bmatrix} 5 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

so

$$A[5 - x + 2x^3]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} = [D(5 - x + 2x^3)]_{\mathcal{C}}$$

which agrees with Theorem 6.26. We leave proof of the general case as an exercise.

Since the linear transformation in Example 6.77 is easy to use directly, there is really no advantage to using the matrix of this transformation to do calculations. However, in other examples—especially large ones—the matrix approach may be simpler, as it is very well-suited to computer implementation. Example 6.78 illustrates the basic idea behind this indirect approach.

Example 6.78

Let $T : \mathscr{P}_2 \to \mathscr{P}_2$ be the linear transformation defined by

$$T(p(x)) = p(2x - 1)$$

(a) Find the matrix of T with respect to $\mathcal{E} = \{1, x, x^2\}$.

(b) Compute $T(3 + 2x - x^2)$ indirectly, using part (a).

Solution (a) We see that

$$T(1) = 1, \quad T(x) = 2x - 1, \quad T(x^2) = (2x - 1)^2 = 1 - 4x + 4x^2$$

so the coordinate vectors are

$$[T(1)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(x)]_{\mathcal{E}} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad [T(x^2)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ -4 \\ 4 \end{bmatrix}$$

Therefore,

$$[T]_{\mathcal{E}} = [\,[T(1)]_{\mathcal{E}} \mid [T(x)]_{\mathcal{E}} \mid [T(x^2)]_{\mathcal{E}}\,] = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

(b) We apply Theorem 6.26 as follows: The coordinate vector of $p(x) = 3 + 2x - x^2$ with respect to $\mathcal{E}$ is

$$[p(x)]_{\mathcal{E}} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$$

Therefore, by Theorem 6.26,

$$\begin{aligned} [T(3 + 2x - x^2)]_{\mathcal{E}} &= [T(p(x))]_{\mathcal{E}} \\ &= [T]_{\mathcal{E}}[p(x)]_{\mathcal{E}} \\ &= \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -4 \end{bmatrix} \end{aligned}$$

It follows that $T(3 + 2x - x^2) = 0 \cdot 1 + 8 \cdot x - 4 \cdot x^2 = 8x - 4x^2$. [Verify this by computing $T(3 + 2x - x^2) = 3 + 2(2x - 1) - (2x - 1)^2$ directly.]

The matrix of a linear transformation can sometimes be used in surprising ways. Example 6.79 shows its application to a traditional calculus problem.

Example 6.79

Let $\mathscr{D}$ be the vector space of all differentiable functions. Consider the subspace W of $\mathscr{D}$ given by $W = \text{span}(e^{3x}, xe^{3x}, x^2e^{3x})$. Since the set $\mathcal{B} = \{e^{3x}, xe^{3x}, x^2e^{3x}\}$ is linearly independent (why?), it is a basis for W.

(a) Show that the differential operator D maps W into itself.

(b) Find the matrix of D with respect to $\mathcal{B}$.

(c) Compute the derivative of $5e^{3x} + 2xe^{3x} - x^2e^{3x}$ indirectly, using Theorem 6.26, and verify it using part (a).

Solution (a) Applying D to a general element of W, we see that

$$D(ae^{3x} + bxe^{3x} + cx^2e^{3x}) = (3a + b)e^{3x} + (3b + 2c)xe^{3x} + 3cx^2e^{3x}$$

(check this), which is again in W.

(b) Using the formula in part (a), we see that

$$D(e^{3x}) = 3e^{3x}, \quad D(xe^{3x}) = e^{3x} + 3xe^{3x}, \quad D(x^2e^{3x}) = 2xe^{3x} + 3x^2e^{3x}$$

so

$$[D(e^{3x})]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \quad [D(xe^{3x})]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \quad [D(x^2e^{3x})]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$$

It follows that

$$[D]_{\mathcal{B}} = [[D(e^{3x})]_{\mathcal{B}} \,\vdots\, [D(xe^{3x})]_{\mathcal{B}} \,\vdots\, [D(x^2e^{3x})]_{\mathcal{B}}] = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

(c) For $f(x) = 5e^{3x} + 2xe^{3x} - x^2e^{3x}$, we see by inspection that

$$[f(x)]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

Hence, by Theorem 6.26, we have

$$[D(f(x))]_{\mathcal{B}} = [D]_{\mathcal{B}}[f(x)]_{\mathcal{B}} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 17 \\ 4 \\ -3 \end{bmatrix}$$

which, in turn, implies that $f'(x) = D(f(x)) = 17e^{3x} + 4xe^{3x} - 3x^2e^{3x}$, in agreement with the formula in part (a).

Remark The point of Example 6.79 is not that this method is easier than direct differentiation. Indeed, once the formula in part (a) has been established, there is little to do. What is significant is that matrix methods can be used at all in what appears, on the surface, to be a calculus problem. We will explore this idea further in Example 6.83.

Example 6.80

Let V be an n-dimensional vector space and let I be the identity transformation on V. What is the matrix of I with respect to bases $\mathcal{B}$ and $\mathcal{C}$ of V if $\mathcal{B} = \mathcal{C}$ (including the order of the basis vectors)? What if $\mathcal{B} \neq \mathcal{C}$?

Solution Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Then $I(\mathbf{v}_1) = \mathbf{v}_1, \ldots, I(\mathbf{v}_n) = \mathbf{v}_n$, so

$$[I(\mathbf{v}_1)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{e}_1, \quad [I(\mathbf{v}_2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{e}_2, \quad \ldots, \quad [I(\mathbf{v}_n)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{e}_n$$

and, if $\mathcal{B} = \mathcal{C}$,

$$\begin{aligned} [I]_{\mathcal{B}} &= [[I(\mathbf{v}_1)]_{\mathcal{B}} \mid [I(\mathbf{v}_2)]_{\mathcal{B}} \mid \cdots \mid [I(\mathbf{v}_n)]_{\mathcal{B}}] \\ &= [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n] \\ &= I_n \end{aligned}$$

the $n \times n$ identity matrix. (This is what you expected, isn't it?)

In the case $\mathcal{B} \neq \mathcal{C}$, we have

$$[I(\mathbf{v}_1)]_{\mathcal{C}} = [\mathbf{v}_1]_{\mathcal{C}}, \quad \ldots, \quad [I(\mathbf{v}_n)]_{\mathcal{C}} = [\mathbf{v}_n]_{\mathcal{C}}$$

so

$$\begin{aligned} [I]_{\mathcal{C}\leftarrow\mathcal{B}} &= [[\mathbf{v}_1]_{\mathcal{C}} \mid \cdots \mid [\mathbf{v}_n]_{\mathcal{C}}] \\ &= P_{\mathcal{C}\leftarrow\mathcal{B}} \end{aligned}$$

the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$.

Matrices of Composite and Inverse Linear Transformations

We now generalize Theorems 3.32 and 3.33 to get a theorem that will allow us to easily find the inverse of a linear transformation between finite-dimensional vector spaces (if it exists).

Theorem 6.27 Let U, V, and W be finite-dimensional vector spaces with bases $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$, respectively. Let $T: U \to V$ and $S: V \to W$ be linear transformations. Then

$$[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}} = [S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}$$

Remarks

- In words, this theorem says, "The matrix of the composite is the product of the matrices."
- Notice how the "inner subscripts" $\mathcal{C}$ must match and appear to cancel each other out, leaving the "outer subscripts" in the form $\mathcal{D} \leftarrow \mathcal{B}$.

Proof We will show that corresponding columns of the matrices $[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}}$ and $[S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ are the same. Let $\mathbf{v}_i$ be the ith basis vector in $\mathcal{B}$. Then the ith column of $[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}}$ is

$$\begin{aligned} [(S \circ T)(\mathbf{v}_i)]_{\mathcal{D}} &= [S(T(\mathbf{v}_i))]_{\mathcal{D}} \\ &= [S]_{\mathcal{D} \leftarrow \mathcal{C}}[T(\mathbf{v}_i)]_{\mathcal{C}} \\ &= [S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}[\mathbf{v}_i]_{\mathcal{B}} \end{aligned}$$

by two applications of Theorem 6.26. But $[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i$ (why?), so

$$[S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}[\mathbf{v}_i]_{\mathcal{B}} = [S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}\mathbf{e}_i$$

is the ith column of the matrix $[S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}$. Therefore, the ith columns of $[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}}$ and $[S]_{\mathcal{D} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ are the same, as we wished to prove.

Example 6.81 Use matrix methods to compute $(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}$ for the linear transformations S and T of Example 6.56.

Solution Recall that $T: \mathbb{R}^2 \to \mathscr{P}_1$ and $S: \mathscr{P}_1 \to \mathscr{P}_2$ are defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x \quad \text{and} \quad S(a + bx) = ax + bx^2$$

Choosing the standard bases $\mathcal{E}$, $\mathcal{E}'$, and $\mathcal{E}''$ for $\mathbb{R}^2$, $\mathscr{P}_1$, and $\mathscr{P}_2$, respectively, we see that

$$[T]_{\mathcal{E}' \leftarrow \mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad [S]_{\mathcal{E}'' \leftarrow \mathcal{E}'} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(Verify these.) By Theorem 6.27, the matrix of $S \circ T$ with respect to $\mathcal{E}$ and $\mathcal{E}''$ is

$$\begin{aligned} [(S \circ T)]_{\mathcal{E}'' \leftarrow \mathcal{E}} &= [S]_{\mathcal{E}'' \leftarrow \mathcal{E}'}[T]_{\mathcal{E}' \leftarrow \mathcal{E}} \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Thus, by Theorem 6.26,

$$\left[(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}\right]_{\mathcal{E}''} = [(S \circ T)]_{\mathcal{E}'' \leftarrow \mathcal{E}} \begin{bmatrix} a \\ b \end{bmatrix}_{\mathcal{E}}$$

$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ a \\ a + b \end{bmatrix}$$

Consequently, $(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix} = ax + (a + b)x^2$, which agrees with the solution to Example 6.56.

In Theorem 6.24, we proved that a linear transformation is invertible if and only if it is one-to-one and onto (i.e., if it is an isomorphism). When the vector spaces involved are finite-dimensional, we can use the matrix methods we have developed to find the inverse of such a linear transformation.

Theorem 6.28 Let $T : V \to W$ be a linear transformation between n-dimensional vector spaces V and W and let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Then T is invertible if and only if the matrix $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ is invertible. In this case,

$$([T]_{\mathcal{C} \leftarrow \mathcal{B}})^{-1} = [T^{-1}]_{\mathcal{B} \leftarrow \mathcal{C}}$$

Proof Observe that the matrices of T and T^{-1} (if T is invertible) are $n \times n$. If T is invertible, then $T^{-1} \circ T = I_V$. Applying Theorem 6.27, we have

$$I_n = [I_V]_{\mathcal{B}} = [T^{-1} \circ T]_{\mathcal{B}}$$
$$= [T^{-1}]_{\mathcal{B} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}$$

This shows that $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ is invertible and that $([T]_{\mathcal{C} \leftarrow \mathcal{B}})^{-1} = [T^{-1}]_{\mathcal{B} \leftarrow \mathcal{C}}$.

Conversely, assume that $A = [T]_{\mathcal{C} \leftarrow \mathcal{B}}$ is invertible. To show that T is invertible, it is enough to show that $\ker(T) = \{\mathbf{0}\}$. (Why?) To this end, let $\mathbf{v}$ be in the kernel of T. Then $T(\mathbf{v}) = \mathbf{0}$, so

$$A[\mathbf{v}]_{\mathcal{B}} = [T]_{\mathcal{C} \leftarrow \mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}} = [\mathbf{0}]_{\mathcal{C}} = \mathbf{0}$$

which means that $[\mathbf{v}]_{\mathcal{B}}$ is in the null space of the invertible matrix A. By the Fundamental Theorem, this implies that $[\mathbf{v}]_{\mathcal{B}} = \mathbf{0}$, which, in turn, implies that $\mathbf{v} = \mathbf{0}$, as required.

Example 6.82 In Example 6.70, the linear transformation $T : \mathbb{R}^2 \to \mathcal{P}_1$ defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

was shown to be one-to-one and onto and hence invertible. Find T^{-1}.

Solution In Example 6.81, we found the matrix of T with respect to the standard bases $\mathcal{E}$ and $\mathcal{E}'$ for $\mathbb{R}^2$ and $\mathcal{P}_1$, respectively, to be

$$[T]_{\mathcal{E}'\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

By Theorem 6.28, it follows that the matrix of T^{-1} with respect to $\mathcal{E}'$ and $\mathcal{E}$ is

$$[T^{-1}]_{\mathcal{E}\leftarrow\mathcal{E}'} = ([T]_{\mathcal{E}'\leftarrow\mathcal{E}})^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

By Theorem 6.26,

$$\begin{aligned} [T^{-1}(a+bx)]_{\mathcal{E}} &= [T^{-1}]_{\mathcal{E}\leftarrow\mathcal{E}'}[a+bx]_{\mathcal{E}'} \\ &= \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} \\ &= \begin{bmatrix} a \\ b-a \end{bmatrix} \end{aligned}$$

This means that

$$T^{-1}(a+bx) = a\mathbf{e}_1 + (b-a)\mathbf{e}_2 = \begin{bmatrix} a \\ b-a \end{bmatrix}$$

(Note that the choice of the standard basis makes this last calculation virtually irrelevant.)

The next example, a continuation of Example 6.79, shows that matrices can be used in certain integration problems in calculus. The specific integral we consider is usually evaluated in a calculus course by means of two applications of integration by parts. Contrast this approach with our method.

Example 6.83

Show that the differential operator, restricted to the subspace $W = \text{span}(e^{3x}, xe^{3x}, x^2e^{3x})$ of $\mathcal{D}$, is invertible, and use this fact to find the integral

$$\int x^2e^{3x}\,dx$$

Solution In Example 6.79, we found the matrix of D with respect to the basis $\mathcal{B} = \{e^{3x}, xe^{3x}, x^2e^{3x}\}$ of W to be

$$[D]_{\mathcal{B}} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

By Theorem 6.28, therefore, D is invertible on W, and the matrix of D^{-1} is

$$[D^{-1}]_{\mathcal{B}} = ([D]_{\mathcal{B}})^{-1} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{9} & \frac{2}{27} \\ 0 & \frac{1}{3} & -\frac{2}{9} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Since integration is *antidifferentiation*, this is the matrix corresponding to integration on W. We want to integrate the function x^2e^{3x} whose coordinate vector is

$$[x^2e^{3x}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Consequently, by Theorem 6.26,

$$\begin{aligned} \left[\int x^2e^{3x}\,dx\right]_{\mathcal{B}} &= [D^{-1}(x^2e^{3x})]_{\mathcal{B}} \\ &= [D^{-1}]_{\mathcal{B}}[x^2e^{3x}]_{\mathcal{B}} \\ &= \begin{bmatrix} \frac{1}{3} & -\frac{1}{9} & \frac{2}{27} \\ 0 & \frac{1}{3} & -\frac{2}{9} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{27} \\ -\frac{2}{9} \\ \frac{1}{3} \end{bmatrix} \end{aligned}$$

It follows that

$$\int x^2e^{3x}\,dx = \tfrac{2}{27}e^{3x} - \tfrac{2}{9}xe^{3x} + \tfrac{1}{3}x^2e^{3x}$$

(To be fully correct, we need to add a constant of integration. It does not show up here because we are working with *linear* transformations, which must send zero vectors to zero vectors, forcing the constant of integration to be zero as well.)

Warning In general, differentiation is *not* an invertible transformation. (See Exercise 22.) What the preceding example shows is that, suitably restricted, it sometimes is. Exercises 27–30 explore this idea further.

Change of Basis and Similarity

Suppose $T: V \to V$ is a linear transformation and $\mathcal{B}$ and $\mathcal{C}$ are two different bases for V. It is natural to wonder how, if at all, the matrices $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$ are related. It turns out that the answer to this question is quite satisfying and relates to some questions we first considered in Chapter 4.

Figure 6.12 suggests one way to address this problem. Chasing the arrows around the diagram from the upper left-hand corner to the lower right-hand corner in two different, but equivalent, ways shows that $I \circ T = T \circ I$, something we already knew, since both are equal to T. However, if the "upper" version of T is with respect to the

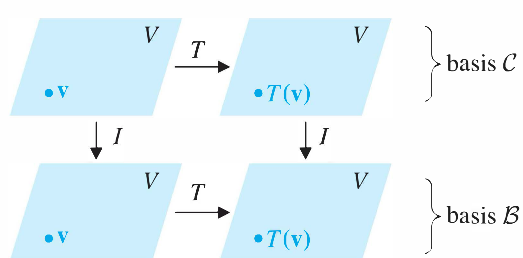


Figure 6.12
$I \circ T = T \circ I$

basis $\mathcal{C}$ and the "lower" version is with respect to $\mathcal{B}$, then $T = I \circ T = T \circ I$ is with respect to $\mathcal{C}$ in its domain and with respect to $\mathcal{B}$ in its codomain. Thus, the matrix of T in this case is $[T]_{\mathcal{B}\leftarrow\mathcal{C}}$. But

$$[T]_{\mathcal{B}\leftarrow\mathcal{C}} = [I \circ T]_{\mathcal{B}\leftarrow\mathcal{C}} = [I]_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{C}}$$

and

$$[T]_{\mathcal{B}\leftarrow\mathcal{C}} = [T \circ I]_{\mathcal{B}\leftarrow\mathcal{C}} = [T]_{\mathcal{B}\leftarrow\mathcal{B}}[I]_{\mathcal{B}\leftarrow\mathcal{C}}$$

Therefore, $[I]_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{C}} = [T]_{\mathcal{B}\leftarrow\mathcal{B}}[I]_{\mathcal{B}\leftarrow\mathcal{C}}$.

From Example 6.80, we know that $[I]_{\mathcal{B}\leftarrow\mathcal{C}} = P_{\mathcal{B}\leftarrow\mathcal{C}}$, the (invertible) change-of-basis matrix from $\mathcal{C}$ to $\mathcal{B}$. If we denote this matrix by P, then we also have

$$P^{-1} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1} = P_{\mathcal{C}\leftarrow\mathcal{B}}$$

With this notation,

$$P[T]_{\mathcal{C}\leftarrow\mathcal{C}} = [T]_{\mathcal{B}\leftarrow\mathcal{B}}P$$

so

$$[T]_{\mathcal{C}\leftarrow\mathcal{C}} = P^{-1}[T]_{\mathcal{B}\leftarrow\mathcal{B}}P \quad \text{or} \quad [T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P$$

Thus, the matrices $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$ are similar, in the terminology of Section 4.4.

We summarize the foregoing discussion as a theorem.

Theorem 6.29 Let V be a finite-dimensional vector space with bases $\mathcal{B}$ and $\mathcal{C}$ and let $T: V \to V$ be a linear transformation. Then

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P$$

where P is the change-of-basis matrix from $\mathcal{C}$ to $\mathcal{B}$.

Remark As an aid in remembering that P must be the change-of-basis matrix from $\mathcal{C}$ to $\mathcal{B}$, and not $\mathcal{B}$ to $\mathcal{C}$, it is instructive to look at what Theorem 6.29 says when written in full detail. As shown below, the "inner subscripts" must be the same (all $\mathcal{B}$s) and must appear to cancel, leaving the "outer subscripts," which are both $\mathcal{C}$s.

$$[T]_{\mathcal{C}\leftarrow\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[T]_{\mathcal{B}\leftarrow\mathcal{B}}P_{\mathcal{B}\leftarrow\mathcal{C}}$$

Same Same

Same

Theorem 6.29 is often used when we are trying to find a basis with respect to which the matrix of a linear transformation is particularly simple. For example, we can ask whether there is a basis $\mathcal{C}$ of V such that the matrix $[T]_{\mathcal{C}}$ of $T: V \to V$ is a diagonal matrix. Example 6.84 illustrates this application.

Example 6.84 Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 3y \\ 2x + 2y \end{bmatrix}$$

If possible, find a basis $\mathcal{C}$ for $\mathbb{R}^2$ such that the matrix of T with respect to $\mathcal{C}$ is diagonal.

Solution The matrix of T with respect to the standard basis $\mathcal{E}$ is

$$[T]_{\mathcal{E}} = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

This matrix is diagonalizable, as we saw in Example 4.24. Indeed, if

$$P = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$$

then $P^{-1}[T]_{\mathcal{E}}P = D$. If we let $\mathcal{C}$ be the basis of $\mathbb{R}^2$ consisting of the columns of P, then P is the change-of-basis matrix $P_{\mathcal{E}\leftarrow\mathcal{C}}$ from $\mathcal{C}$ to $\mathcal{E}$. By Theorem 6.29,

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{E}}P = D$$

so the matrix of T with respect to the basis $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$ is diagonal.

Remarks

- It is easy to check that the solution above is correct by computing $[T]_{\mathcal{C}}$ directly. We find that

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 3 \\ -2 \end{bmatrix} \quad \text{and} \quad T\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Thus, the coordinate vectors that form the columns of $[T]_{\mathcal{C}}$ are

$$\left[T\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \quad \text{and} \quad \left[T\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right]_{\mathcal{C}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

in agreement with our solution above.

- The general procedure for a problem like Example 6.84 is to take the standard matrix $[T]_{\mathcal{E}}$ and determine whether it is diagonalizable by finding bases for its eigenspaces, as in Chapter 4. The solution then proceeds exactly as in the preceding example.

Example 6.84 motivates the following definition.

Definition Let V be a finite-dimensional vector space and let $T : V \to V$ be a linear transformation. Then T is called ***diagonalizable*** if there is a basis $\mathcal{C}$ for V such that the matrix $[T]_{\mathcal{C}}$ is a diagonal matrix.

It is not hard to show that if $\mathcal{B}$ is *any* basis for V, then T is diagonalizable if and only if the matrix $[T]_{\mathcal{B}}$ is diagonalizable. This is essentially what we did, for a special case, in the last example. You are asked to prove this result in general in Exercise 42.

Sometimes it is easiest to write down the matrix of a linear transformation with respect to a "nonstandard" basis. We can then reverse the process of Example 6.84 to find the standard matrix. We illustrate this idea by revisiting Example 3.59.

Example 6.85

Let ℓ be the line through the origin in $\mathbb{R}^2$ with direction vector $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$. Find the standard matrix of the projection onto ℓ.

Solution Let T denote the projection. There is no harm in assuming that $\mathbf{d}$ is a unit vector (i.e., $d_1^2 + d_2^2 = 1$), since any nonzero multiple of $\mathbf{d}$ can serve as a direction vector for ℓ. Let $\mathbf{d}' = \begin{bmatrix} -d_2 \\ d_1 \end{bmatrix}$ so that $\mathbf{d}$ and $\mathbf{d}'$ are orthogonal. Since $\mathbf{d}'$ is also a unit vector, the set $\mathcal{D} = \{\mathbf{d}, \mathbf{d}'\}$ is an orthonormal basis for $\mathbb{R}^2$.

As Figure 6.13 shows, $T(\mathbf{d}) = \mathbf{d}$ and $T(\mathbf{d}') = \mathbf{0}$. Therefore,

$$[T(\mathbf{d})]_{\mathcal{D}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad [T(\mathbf{d}')]_{\mathcal{D}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

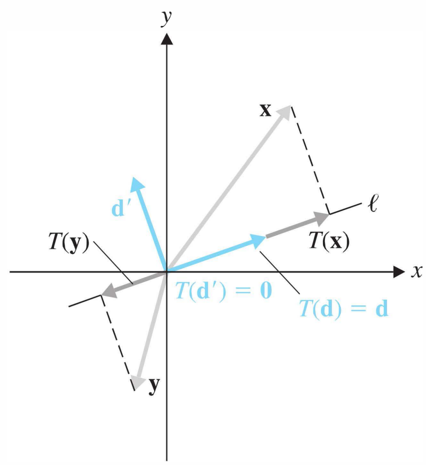


Figure 6.13
Projection onto ℓ

so

$$[T]_{\mathcal{D}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The change-of-basis matrix from $\mathcal{D}$ to the standard basis $\mathcal{E}$ is

$$P_{\mathcal{E}\leftarrow\mathcal{D}} = \begin{bmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{bmatrix}$$

so the change-of-basis matrix from $\mathcal{E}$ to $\mathcal{D}$ is

$$P_{\mathcal{D}\leftarrow\mathcal{E}} = (P_{\mathcal{E}\leftarrow\mathcal{D}})^{-1} = \begin{bmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{bmatrix}^{-1} = \begin{bmatrix} d_1 & d_2 \\ -d_2 & d_1 \end{bmatrix}$$

By Theorem 6.29, then, the standard matrix of T is

$$\begin{aligned} [T]_{\mathcal{E}} &= P_{\mathcal{E}\leftarrow\mathcal{D}}[T]_{\mathcal{D}}P_{\mathcal{D}\leftarrow\mathcal{E}} \\ &= \begin{bmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} d_1 & d_2 \\ -d_2 & d_1 \end{bmatrix} \\ &= \begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix} \end{aligned}$$

which agrees with part (b) of Example 3.59.

Example 6.86

Let $T : \mathscr{P}_2 \to \mathscr{P}_2$ be the linear transformation defined by

$$T(p(x)) = p(2x - 1)$$

(a) Find the matrix of T with respect to the basis $\mathcal{B} = \{1 + x, 1 - x, x^2\}$ of $\mathscr{P}_2$.

(b) Show that T is diagonalizable and find a basis $\mathcal{C}$ for $\mathscr{P}_2$ such that $[T]_{\mathcal{C}}$ is a diagonal matrix.

Solution (a) In Example 6.78, we found that the matrix of T with respect to the standard basis $\mathcal{E} = \{1, x, x^2\}$ is

$$[T]_{\mathcal{E}} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

The change-of-basis matrix from $\mathcal{B}$ to $\mathcal{E}$ is

$$P = P_{\mathcal{E} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It follows that the matrix of T with respect to $\mathcal{B}$ is

$$\begin{aligned} [T]_{\mathcal{B}} &= P^{-1}[T]_{\mathcal{E}}P \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -\frac{3}{2} \\ -1 & 2 & \frac{5}{2} \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$

(Check this.)

(b) The eigenvalues of $[T]_{\mathcal{E}}$ are 1, 2, and 4 (why?), so we know from Theorem 4.25 that $[T]_{\mathcal{E}}$ is diagonalizable. Eigenvectors corresponding to these eigenvalues are

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

respectively. Therefore, setting

$$P = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

we have $P^{-1}[T]_{\mathcal{E}}P = D$. Furthermore, P is the change-of-basis matrix from a basis $\mathcal{C}$ to $\mathcal{E}$, and the columns of P are thus the coordinate vectors of $\mathcal{C}$ in terms of $\mathcal{E}$. It follows that

$$\mathcal{C} = \{1, -1 + x, 1 - 2x + x^2\}$$

and $[T]_{\mathcal{C}} = D$.

The preceding ideas can be generalized to relate the matrices $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ and $[T]_{\mathcal{C}'\leftarrow\mathcal{B}'}$ of a linear transformation $T: V \to W$, where $\mathcal{B}$ and $\mathcal{B}'$ are bases for V and $\mathcal{C}$ and $\mathcal{C}'$ are bases for W. (See Exercise 44.)

We conclude this section by revisiting the Fundamental Theorem of Invertible Matrices and incorporating some results from this chapter.

Theorem 6.30 **The Fundamental Theorem of Invertible Matrices: Version 4**

Let A be an $n \times n$ matrix and let $T: V \to W$ be a linear transformation whose matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ with respect to bases $\mathcal{B}$ and $\mathcal{C}$ of V and W, respectively, is A. The following statements are equivalent:

a. A is invertible.
b. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
d. The reduced row echelon form of A is I_n.
e. A is a product of elementary matrices.
f. $\text{rank}(A) = n$
g. $\text{nullity}(A) = 0$
h. The column vectors of A are linearly independent.
i. The column vectors of A span $\mathbb{R}^n$.
j. The column vectors of A form a basis for $\mathbb{R}^n$.
k. The row vectors of A are linearly independent.
l. The row vectors of A span $\mathbb{R}^n$.
m. The row vectors of A form a basis for $\mathbb{R}^n$.
n. $\det A \neq 0$
o. 0 is not an eigenvalue of A.
p. T is invertible.
q. T is one-to-one.
r. T is onto.
s. $\ker(T) = \{\mathbf{0}\}$
t. $\text{range}(T) = W$

Proof The equivalence (q) $\Leftrightarrow$ (s) is Theorem 6.20, and (r) $\Leftrightarrow$ (t) is the definition of onto. Since A is $n \times n$, we must have $\dim V = \dim W = n$. From Theorems 6.21 and 6.24, we get (p) $\Leftrightarrow$ (q) $\Leftrightarrow$ (r). Finally, we connect the last five statements to the others by Theorem 6.28, which implies that (a) $\Leftrightarrow$ (p).

Exercises 6.6

In Exercises 1–12, find the matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ of the linear transformation $T: V \to W$ with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ of V and W, respectively. Verify Theorem 6.26 for the vector $\mathbf{v}$ by computing $T(\mathbf{v})$ directly and using the theorem.

1. $T: \mathcal{P}_1 \to \mathcal{P}_1$ defined by $T(a + bx) = b - ax$, $\mathcal{B} = \mathcal{C} = \{1, x\}$, $\mathbf{v} = p(x) = 4 + 2x$

2. $T: \mathcal{P}_1 \to \mathcal{P}_1$ defined by $T(a + bx) = b - ax$, $\mathcal{B} = \{1 + x, 1 - x\}$, $\mathcal{C} = \{1, x\}$, $\mathbf{v} = p(x) = 4 + 2x$

3. $T: \mathcal{P}_2 \to \mathcal{P}_2$ defined by $T(p(x)) = p(x + 2)$, $\mathcal{B} = \{1, x, x^2\}$, $\mathcal{C} = \{1, x + 2, (x + 2)^2\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

4. $T: \mathcal{P}_2 \to \mathcal{P}_2$ defined by $T(p(x)) = p(x+2)$, $\mathcal{B} = \{1, x+2, (x+2)^2\}$, $\mathcal{C} = \{1, x, x^2\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

5. $T: \mathcal{P}_2 \to \mathbb{R}^2$ defined by $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$, $\mathcal{B} = \{1, x, x^2\}$, $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

6. $T: \mathcal{P}_2 \to \mathbb{R}^2$ defined by $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$, $\mathcal{B} = \{x^2, x, 1\}$, $\mathcal{C} = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

7. $T: \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+2b \\ -a \\ b \end{bmatrix}, \quad \mathcal{B} = \left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix}\right\},$$

$$\mathcal{C} = \left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}, \quad \mathbf{v} = \begin{bmatrix} -7 \\ 7 \end{bmatrix}$$

8. Repeat Exercise 7 with $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

9. $T: M_{22} \to M_{22}$ defined by $T(A) = A^T$, $\mathcal{B} = \mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, $\mathbf{v} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

10. Repeat Exercise 9 with $\mathcal{B} = \{E_{22}, E_{21}, E_{12}, E_{11}\}$ and $\mathcal{C} = \{E_{12}, E_{21}, E_{22}, E_{11}\}$.

11. $T: M_{22} \to M_{22}$ defined by $T(A) = AB - BA$, where $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, $\mathcal{B} = \mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, $\mathbf{v} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

12. $T: M_{22} \to M_{22}$ defined by $T(A) = A - A^T$, $\mathcal{B} = \mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, $\mathbf{v} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

13. Consider the subspace W of $\mathcal{D}$, given by $W = \text{span}(\sin x, \cos x)$.

(a) Show that the differential operator D maps W into itself.

(b) Find the matrix of D with respect to $\mathcal{B} = \{\sin x, \cos x\}$.

(c) Compute the derivative of $f(x) = 3\sin x - 5\cos x$ indirectly, using Theorem 6.26, and verify that it agrees with $f'(x)$ as computed directly.

14. Consider the subspace W of $\mathcal{D}$, given by $W = \text{span}(e^{2x}, e^{-2x})$.

(a) Show that the differential operator D maps W into itself.

(b) Find the matrix of D with respect to $\mathcal{B} = \{e^{2x}, e^{-2x}\}$.

(c) Compute the derivative of $f(x) = e^{2x} - 3e^{-2x}$ indirectly, using Theorem 6.26, and verify that it agrees with $f'(x)$ as computed directly.

15. Consider the subspace W of $\mathcal{D}$, given by $W = \text{span}(e^{2x}, e^{2x}\cos x, e^{2x}\sin x)$.

(a) Find the matrix of D with respect to $\mathcal{B} = \{e^{2x}, e^{2x}\cos x, e^{2x}\sin x\}$.

(b) Compute the derivative of $f(x) = 3e^{2x} - e^{2x}\cos x + 2e^{2x}\sin x$ indirectly, using Theorem 6.26, and verify that it agrees with $f'(x)$ as computed directly.

16. Consider the subspace W of $\mathcal{D}$, given by $W = \text{span}(\cos x, \sin x, x\cos x, x\sin x)$.

(a) Find the matrix of D with respect to $\mathcal{B} = \{\cos x, \sin x, x\cos x, x\sin x\}$.

(b) Compute the derivative of $f(x) = \cos x + 2x\cos x$ indirectly, using Theorem 6.26, and verify that it agrees with $f'(x)$ as computed directly.

In Exercises 17 and 18, $T: U \to V$ and $S: V \to W$ are linear transformations and $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ are bases for U, V, and W, respectively. Compute $[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}}$ in two ways: (a) by finding $S \circ T$ directly and then computing its matrix and (b) by finding the matrices of S and T separately and using Theorem 6.27.

17. $T: \mathcal{P}_1 \to \mathbb{R}^2$ defined by $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$, $S: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $S\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a-2b \\ 2a-b \end{bmatrix}$, $\mathcal{B} = \{1, x\}$, $\mathcal{C} = \mathcal{D} = \{\mathbf{e}_1, \mathbf{e}_2\}$

18. $T: \mathcal{P}_1 \to \mathcal{P}_2$ defined by $T(p(x)) = p(x+1)$, $S: \mathcal{P}_2 \to \mathcal{P}_2$ defined by $S(p(x)) = p(x+1)$, $\mathcal{B} = \{1, x\}$, $\mathcal{C} = \mathcal{D} = \{1, x, x^2\}$

In Exercises 19–26, determine whether the linear transformation T is invertible by considering its matrix with respect to the standard bases. If T is invertible, use Theorem 6.28 and the method of Example 6.82 to find T^{-1}.

19. T in Exercise 1

20. T in Exercise 5

21. T in Exercise 3

22. $T: \mathcal{P}_2 \to \mathcal{P}_2$ defined by $T(p(x)) = p'(x)$

23. $T: \mathcal{P}_2 \to \mathcal{P}_2$ defined by $T(p(x)) = p(x) + p'(x)$

24. $T : M_{22} \to M_{22}$ defined by $T(A) = AB$, where

$$B = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$$

25. T in Exercise 11 **26.** T in Exercise 12

In Exercises 27–30, use the method of Example 6.83 to evaluate the given integral.

27. $\int (\sin x - 3\cos x)\, dx$. (See Exercise 13.)

28. $\int 5e^{-2x}\, dx$. (See Exercise 14.)

29. $\int (e^{2x}\cos x - 2e^{2x}\sin x)\, dx$. (See Exercise 15.)

30. $\int (x\cos x + x\sin x)\, dx$. (See Exercise 16.)

In Exercises 31–36, a linear transformation $T : V \to V$ is given. If possible, find a basis $\mathcal{C}$ for V such that the matrix $[T]_{\mathcal{C}}$ of T with respect to $\mathcal{C}$ is diagonal.

31. $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -4b \\ a + 5b \end{bmatrix}$

32. $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a - b \\ a + b \end{bmatrix}$

33. $T : \mathcal{P}_1 \to \mathcal{P}_1$ defined by $T(a + bx) = (4a + 2b) + (a + 3b)x$

34. $T : \mathcal{P}_2 \to \mathcal{P}_2$ defined by $T(p(x)) = p(x + 1)$

35. $T : \mathcal{P}_1 \to \mathcal{P}_1$ defined by $T(p(x)) = p(x) + xp'(x)$

36. $T : \mathcal{P}_2 \to \mathcal{P}_2$ defined by $T(p(x)) = p(3x + 2)$

37. Let ℓ be the line through the origin in $\mathbb{R}^2$ with direction vector $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$. Use the method of Example 6.85 to find the standard matrix of a reflection in ℓ.

38. Let W be the plane in $\mathbb{R}^3$ with equation $x - y + 2z = 0$. Use the method of Example 6.85 to find the standard matrix of an orthogonal projection onto W. Verify that your answer is correct by using it to compute the orthogonal projection of $\mathbf{v}$ onto W, where

$$\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$$

Compare your answer with Example 5.11. [*Hint:* Find an orthogonal decomposition of $\mathbb{R}^3$ as $\mathbb{R}^3 = W + W^{\perp}$ using an orthogonal basis for W. See Example 5.3.]

39. Let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces and let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Show that the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ is unique. That is, if A is a matrix such that $A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$ for all $\mathbf{v}$ in V, then $A = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$. [*Hint:* Find values of $\mathbf{v}$ that will show this, one column at a time.]

In Exercises 40–45, let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces V and W. Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively, and let $A = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$.

40. Show that nullity(T) = nullity(A).

41. Show that rank(T) = rank(A).

42. If $V = W$ and $\mathcal{B} = \mathcal{C}$, show that T is diagonalizable if and only if A is diagonalizable.

43. Use the results of this section to give a matrix-based proof of the Rank Theorem (Theorem 6.19).

44. If $\mathcal{B}'$ and $\mathcal{C}'$ are also bases for V and W, respectively, what is the relationship between $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ and $[T]_{\mathcal{C}'\leftarrow\mathcal{B}'}$? Prove your assertion.

45. If $\dim V = n$ and $\dim W = m$, prove that $\mathcal{L}(V, W) \cong M_{mn}$. (See the exercises for Section 6.4.) [*Hint:* Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Show that the mapping $\varphi(T) = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$, for T in $\mathcal{L}(V, W)$, defines a linear transformation $\varphi : \mathcal{L}(V, W) \to M_{mn}$ that is an isomorphism.]

46. If V is a vector space, then the ***dual space*** of V is the vector space $V^* = \mathcal{L}(V, \mathbb{R})$. Prove that if V is finite-dimensional, then $V^* \cong V$.

Exploration

Tilings, Lattices, and the Crystallographic Restriction

Repeating patterns are frequently found in nature and in art. The molecular structure of crystals often exhibits repetition, as do the tilings and mosaics found in the artwork of many cultures. *Tiling* (or *tessellation*) is covering of a plane by shapes that do not overlap and leave no gaps. The Dutch artist M. C. Escher (1898–1972) produced many works in which he explored the possibility of tiling a plane using fanciful shapes (Figure 6.14).

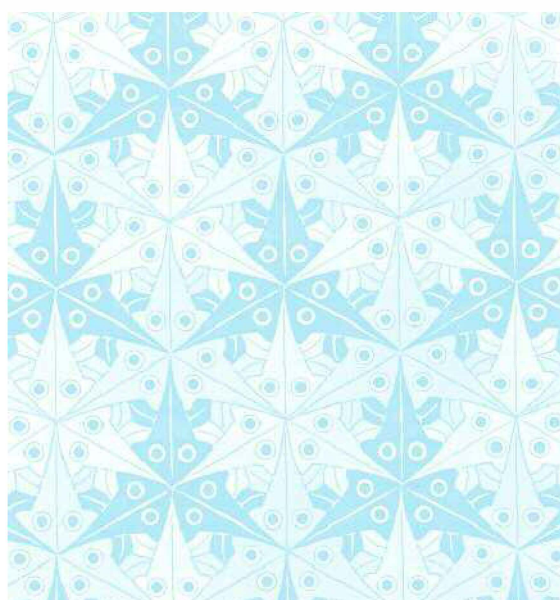


Figure 6.14
M. C. Escher's "Symmetry Drawing E103"

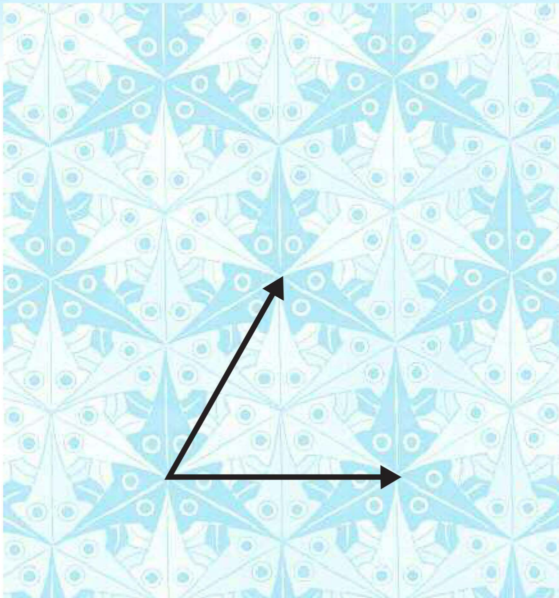


Figure 6.15
Invariance under translation
M. C. Escher's "Symmetry Drawing E103"

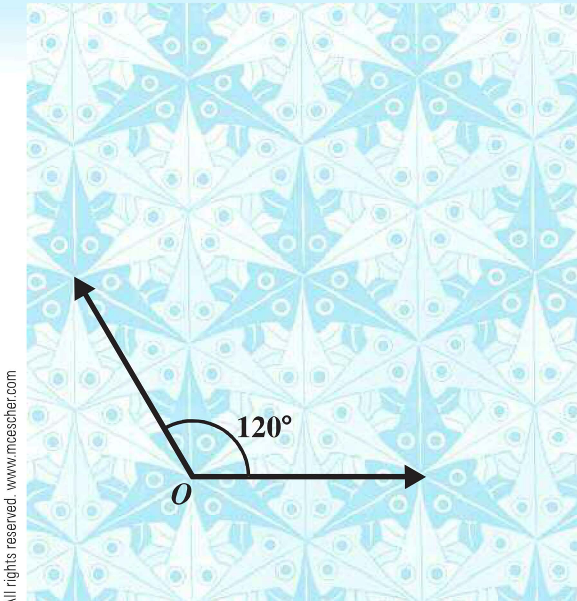


Figure 6.17
Rotational symmetry
M. C. Escher's "Symmetry Drawing E103"

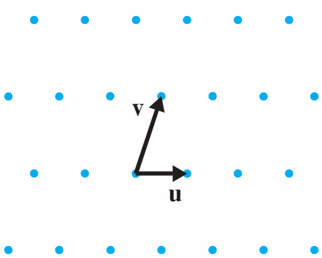


Figure 6.16
A lattice

In this exploration, we will be interested in patterns such as those in Figure 6.14, which we assume to be infinite and repeating in all directions of the plane. Such a pattern has the property that it can be shifted (or *translated*) in at least two directions (corresponding to two linearly independent vectors) so that it appears not to have been moved at all. We say that the pattern is *invariant* under translations and has ***translational symmetry*** in these directions. For example, the pattern in Figure 6.14 has translational symmetry in the directions shown in Figure 6.15.

If a pattern has translational symmetry in two directions, it has translational symmetry in infinitely many directions.

1. Let the two vectors shown in Figure 6.15 be denoted by $\mathbf{u}$ and $\mathbf{v}$. Show that the pattern in Figure 6.14 is invariant under translation by any *integer* linear combination of $\mathbf{u}$ and $\mathbf{v}$—that is, by any vector of the form $a\mathbf{u} + b\mathbf{v}$, where a and b are integers.

For any two linearly independent vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, the set of points determined by all integer linear combinations of $\mathbf{u}$ and $\mathbf{v}$ is called a ***lattice.*** Figure 6.16 shows an example of a lattice.

2. Draw the lattice corresponding to the vectors $\mathbf{u}$ and $\mathbf{v}$ of Figure 6.15.

Figure 6.14 also exhibits ***rotational symmetry.*** That is, it is possible to rotate the entire pattern about some point and have it appear unchanged. We say that it is *invariant* under such a rotation. For example, the pattern of Figure 6.14 is invariant under a rotation of 120° about the point O, as shown in Figure 6.17. We call O a ***center*** of rotational symmetry (or a ***rotation center***).

Note that if a pattern is based on an underlying lattice, then any symmetries of the pattern must also be possessed by the lattice.

3. Explain why, if a point O is a rotation center through an angle θ, then it is a rotation center through every integer multiple of θ. Deduce that if $0 < \theta \leq 360°$, then $360/\theta$ must be an integer. (If $360/\theta = n$, we say the pattern or lattice has ***n-fold*** rotational symmetry.)

4. What is the smallest positive angle of rotational symmetry for the lattice in Problem 2? Does the pattern in Figure 6.14 also have rotational symmetry through this angle?

5. Take various values of θ such that $0 < \theta \leq 360°$ and $360/\theta$ is an integer. Try to draw a lattice that has rotational symmetry through the angle θ. In particular, can you draw a lattice with eight-fold rotational symmetry?

We will show that values of θ that are possible angles of rotational symmetry for a lattice are severely restricted. The technique we will use is to consider rotation transformations in terms of different bases. Accordingly, let R_θ denote a rotation about the origin through an angle θ and let $\mathcal{E}$ be the standard basis for $\mathbb{R}^2$. Then the standard matrix of R_θ is

$$[R_\theta]_{\mathcal{E}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

6. Referring to Problems 2 and 4, take the origin to be at the tails of **u** and **v**.

(a) What is the actual (i.e., numerical) value of $[R_\theta]_{\mathcal{E}}$ in this case?

(b) Let $\mathcal{B}$ be the basis $\{\mathbf{u}, \mathbf{v}\}$. Compute the matrix $[R_\theta]_{\mathcal{B}}$.

7. In general, let **u** and **v** be any two linearly independent vectors in $\mathbb{R}^2$ and suppose that the lattice determined by **u** and **v** is invariant under a rotation through an angle θ. If $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$, show that the matrix of R_θ with respect to $\mathcal{B}$ must have the form

$$[R_\theta]_{\mathcal{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

where a, b, c, and d are *integers.*

8. In the terminology and notation of Problem 7, show that $2\cos\theta$ must be an integer. [*Hint:* Use Exercise 35 in Section 4.4 and Theorem 6.29.]

9. Using Problem 8, make a list of all possible values of θ, with $0 < \theta \leq 360°$, that can be angles of rotational symmetry of a lattice. Record the corresponding values of n, where $n = 360/\theta$, to show that a lattice can have n-fold rotational symmetry if and only if $n = 1, 2, 3, 4$, or 6. This result, known as the ***crystallographic restriction,*** was first proved by W. Barlow in 1894.

10. In the library or on the Internet, see whether you can find an Escher tiling for each of the five possible types of rotational symmetry—that is, where the *smallest* angle of rotational symmetry of the pattern is one of those specified by the crystallographic restriction.

6.7 Applications

Homogeneous Linear Differential Equations

In Exercises 69–72 in Section 4.6, we showed that if $y = y(t)$ is a twice-differentiable function that satisfies the differential equation

$$y'' + ay' + by = 0 \tag{1}$$

then y is of the form

$$y = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

if λ_1 and λ_2 are *distinct* roots of the associated characteristic equation $\lambda^2 + a\lambda + b = 0$. (The case where $\lambda_1 = \lambda_2$ was left unresolved.) Example 6.12 and Exercise 20 in this section show that the set of solutions to Equation (1) forms a subspace of $\mathscr{F}$, the vector space of functions. In this section, we pursue these ideas further, paying particular attention to the role played by vector spaces, bases, and dimension.

To set the stage, we consider a simpler class of examples. A differential equation of the form

$$y' + ay = 0 \tag{2}$$

is called a ***first-order, homogeneous, linear differential equation.*** ("First-order" refers to the fact that the highest derivative that is involved is a first derivative, and "homogeneous" means that the right-hand side is zero. Do you see why the equation is "linear"?) A ***solution*** to Equation (2) is a differentiable function $y = y(t)$ that satisfies Equation (2) for all values of t.

It is easy to check that one solution to Equation (2) is $y = e^{-at}$. (Do it.) However, we would like to describe *all* solutions—and this is where vector spaces come in. We have the following theorem.

Theorem 6.31 The set S of all solutions to $y' + ay = 0$ is a subspace of $\mathscr{F}$.

Proof Since the zero function certainly satisfies Equation (2), S is nonempty. Let x and y be two differentiable functions of t that are in S and let c be a scalar. Then

$$x' + ax = 0 \quad \text{and} \quad y' + ay = 0$$

so, using rules for differentiation, we have

$$(x + y)' + a(x + y) = x' + y' + ax + ay = (x' + ax) + (y' + ay) = 0 + 0 = 0$$

and

$$(cy)' + a(cy) = cy' + c(ay) = c(y' + ay) = c \cdot 0 = 0$$

Hence, $x + y$ and cy are also in S, so S is a subspace of $\mathscr{F}$.

Now we will show that S is a one-dimensional subspace of $\mathscr{F}$ and that $\{e^{-at}\}$ is a basis. To this end, let $x = x(t)$ be in S. Then, for all t,

$$x'(t) + ax(t) = 0 \quad \text{or} \quad x'(t) = -ax(t)$$

Define a new function $z(t) = x(t)e^{at}$. Then, by the Chain Rule for differentiation,

$$\begin{aligned} z'(t) &= x(t)ae^{at} + x'(t)e^{at} \\ &= ax(t)e^{at} - ax(t)e^{at} \\ &= 0 \end{aligned}$$

Since z' is identically zero, z must be a constant function—say, $z(t) = k$. But this means that

$$x(t)e^{at} = z(t) = k \quad \text{for all } t$$

so $x(t) = ke^{-at}$. Therefore, all solutions to Equation (2) are scalar multiples of the single solution $y = e^{-at}$. We have proved the following theorem.

Theorem 6.32 If S is the solution space of $y' + ay = 0$, then $\dim S = 1$ and $\{e^{-at}\}$ is a basis for S.

One model for population growth assumes that the growth rate of the population is proportional to the size of the population. This model works well if there are few restrictions (such as limited space, food, or the like) on growth. If the size of the population at time t is $p(t)$, then the growth rate, or rate of change of the population, is its derivative $p'(t)$. Our assumption that the growth rate of the population is proportional to its size can be written as

$$p'(t) = kp(t)$$

where k is the proportionality constant. Thus, p satisfies the differential equation $p' - kp = 0$, so, by Theorem 6.32,

$$p(t) = ce^{kt}$$

for some scalar c. The constants c and k are determined using experimental data.

Example 6.87 The bacterium *Escherichia coli* (or *E. coli*, for short) is commonly found in the intestines of humans and other mammals. It poses severe health risks if it escapes into the environment. Under laboratory conditions, each cell of the bacterium divides into two every 20 minutes. If we start with a single *E. coli* cell, how many will there be after 1 day?

E. coli is mentioned in Michael Crichton's novel *The Andromeda Strain* (New York: Dell, 1969), although the "villain" in that novel was supposedly an alien virus. In real life, *E. coli* contaminated the town water supply of Walkerton, Ontario, in 2000, resulting in seven deaths and causing hundreds of people to become seriously ill.

Solution We do not need to use differential equations to solve this problem, but we will, in order to illustrate the basic method.

To determine c and k, we use the data given in the statement of the problem. If we take 1 unit of time to be 20 minutes, then we are given that $p(0) = 1$ and $p(1) = 2$. Therefore,

$$c = c \cdot 1 = ce^{k \cdot 0} = 1 \quad \text{and} \quad 2 = ce^{k \cdot 1} = e^k$$

It follows that $k = \ln 2$, so

$$p(t) = e^{t \ln 2} = e^{\ln 2^t} = 2^t$$

After 1 day, $t = 72$, so the number of bacteria cells will be $p(72) = 2^{72} \approx 4.72 \times 10^{21}$ (see Figure 6.18).

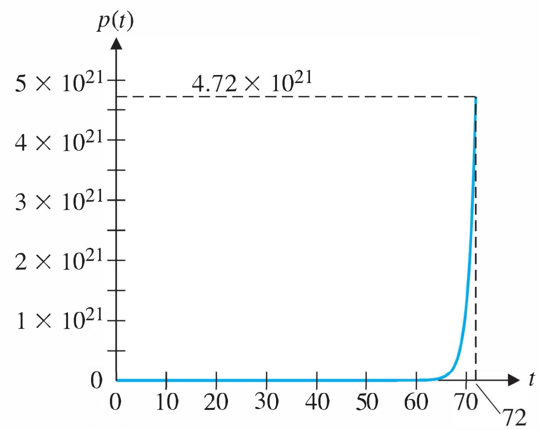


Figure 6.18
Exponential growth

Radioactive substances decay by emitting radiation. If $m(t)$ denotes the mass of the substance at time t, then the rate of decay is $m'(t)$. Physicists have found that the rate of decay of a substance is proportional to its mass; that is,

$$m'(t) = km(t) \quad \text{or} \quad m' - km = 0$$

where k is a negative constant. Applying Theorem 6.32, we have

$$m(t) = ce^{kt}$$

for some constant c. The time required for half of a radioactive substance to decay is called its ***half-life.***

Example 6.88

After 5.5 days, a 100 mg sample of radon-222 decayed to 37 mg.

(a) Find a formula for $m(t)$, the mass remaining after t days.

(b) What is the half-life of radon-222?

(c) When will only 10 mg remain?

Solution (a) From $m(t) = ce^{kt}$, we have

$$100 = m(0) = ce^{k \cdot 0} = c \cdot 1 = c$$

so

$$m(t) = 100e^{kt}$$

With time measured in days, we are given that $m(5.5) = 37$. Therefore,

$$100e^{5.5k} = 37$$

so

$$e^{5.5k} = 0.37$$

Solving for k, we find

$$5.5k = \ln(0.37)$$

so

$$k = \frac{\ln(0.37)}{5.5} \approx -0.18$$

Therefore, $m(t) = 100e^{-0.18t}$.

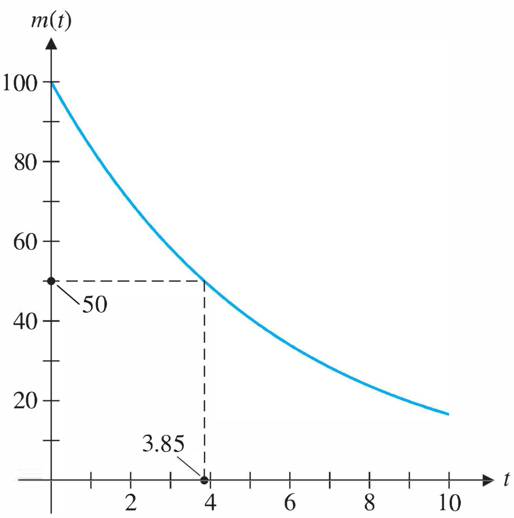


Figure 6.19
Radioactive decay

(b) To find the half-life of radon-222, we need the value of t for which $m(t) = 50$. Solving this equation, we find

$$100e^{-0.18t} = 50$$

so

$$e^{-0.18t} = 0.50$$

Hence,

$$-0.18t = \ln(\tfrac{1}{2}) = -\ln 2$$

and

$$t = \frac{\ln 2}{0.18} \approx 3.85$$

Thus, radon-222 has a half-life of approximately 3.85 days. (See Figure 6.19.)

(c) We need to determine the value of t such that $m(t) = 10$. That is, we must solve the equation

$$100e^{-0.18t} = 10 \quad \text{or} \quad e^{-0.18t} = 0.1$$

Taking the natural logarithm of both sides yields $-0.18t = \ln 0.1$. Thus,

$$t = \frac{\ln 0.1}{-0.18} \approx 12.79$$

so 10 mg of the sample will remain after approximately 12.79 days.

See *Linear Algebra* by S. H. Friedberg, A. J. Insel, and L. E. Spence (Englewood Cliffs, NJ: Prentice-Hall, 1979).

The solution set S of the second-order differential equation $y'' + ay' + by = 0$ is also a subspace of $\mathscr{F}$ (Exercise 20), and it turns out that the dimension of S is 2. Part (a) of Theorem 6.33, which extends Theorem 6.32, is implied by Theorem 4.40. Our approach here is to use the power of vector spaces; doing so allows us to obtain part (b) of Theorem 6.33 as well, a result that we could not obtain with our previous methods.

Theorem 6.33 Let S be the solution space of

$$y'' + ay' + by = 0$$

and let λ_1 and λ_2 be the roots of the characteristic equation $\lambda^2 + a\lambda + b = 0$.

a. If $\lambda_1 \neq \lambda_2$, then $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ is a basis for S.
b. If $\lambda_1 = \lambda_2$, then $\{e^{\lambda_1 t}, te^{\lambda_1 t}\}$ is a basis for S.

Remarks

- Observe that what the theorem says, in other words, is that the solutions of $y'' + ay' + by = 0$ are of the form

$$y = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

in the first case and

$$y = c_1 e^{\lambda_1 t} + c_2 t e^{\lambda_1 t}$$

in the second case.

- Compare Theorem 6.33 with Theorem 4.38. Linear differential equations and linear recurrence relations have much in common. Although the former belong to *continuous* mathematics and the latter to *discrete* mathematics, there are many parallels.

Proof (a) We first show that $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ is contained in S. Let λ be any root of the characteristic equation and let $f(t) = e^{\lambda t}$. Then

$$f'(t) = \lambda e^{\lambda t} \quad \text{and} \quad f''(t) = \lambda^2 e^{\lambda t}$$

from which it follows that

$$\begin{aligned} f'' + af' + bf &= \lambda^2 e^{\lambda t} + a\lambda e^{\lambda t} + be^{\lambda t} \\ &= (\lambda^2 + a\lambda + b)e^{\lambda t} \\ &= 0 \cdot e^{\lambda t} = 0 \end{aligned}$$

Therefore, f is in S. But, since λ_1 and λ_2 are roots of the characteristic equation, this means that $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ are in S.

The set $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ is also linearly independent, since if

$$c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} = 0$$

then, setting $t = 0$, we have

$$c_1 + c_2 = 0 \quad \text{or} \quad c_2 = -c_1$$

Next, we set $t = 1$ to obtain

$$c_1 e^{\lambda_1} - c_1 e^{\lambda_2} = 0 \quad \text{or} \quad c_1(e^{\lambda_1} - e^{\lambda_2}) = 0$$

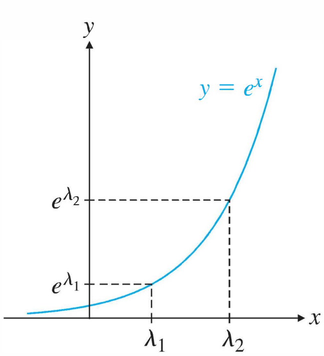


Figure 6.20

But $e^{\lambda_1} - e^{\lambda_2} \neq 0$, since $e^{\lambda_1} - e^{\lambda_2} = 0$ implies that $e^{\lambda_1} = e^{\lambda_2}$, which is clearly impossible if $\lambda_1 \neq \lambda_2$. (See Figure 6.20.) We deduce that $c_1 = 0$ and, hence, $c_2 = 0$, so $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ is linearly independent.

Since $\dim S = 2$, $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ must be a basis for S.

(b) You are asked to prove this property in Exercise 21.

Example 6.89

Find all solutions of $y'' - 5y' + 6y = 0$.

Solution The characteristic equation is $\lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3) = 0$. Thus, the roots are 2 and 3, so $\{e^{2t}, e^{3t}\}$ is a basis for the solution space. It follows that the solutions to the given equation are of the form

$$y = c_1e^{2t} + c_2e^{3t}$$

The constants c_1 and c_2 can be determined if additional equations, called ***boundary conditions,*** are specified.

Example 6.90

Find the solution of $y'' + 6y' + 9y = 0$ that satisfies $y(0) = 1$, $y'(0) = 0$.

Solution The characteristic equation is $\lambda^2 + 6\lambda + 9 = (\lambda + 3)^2 = 0$, so -3 is a repeated root. Therefore, $\{e^{-3t}, te^{-3t}\}$ is a basis for the solution space, and the general solution is of the form

$$y = c_1e^{-3t} + c_2te^{-3t}$$

The first boundary condition gives

$$1 = y(0) = c_1e^{-3\cdot 0} + 0 = c_1$$

so $y = e^{-3t} + c_2te^{-3t}$. Differentiating, we have

$$y' = -3e^{-3t} + c_2(-3te^{-3t} + e^{-3t})$$

so the second boundary condition gives

$$0 = y'(0) = -3e^{-3\cdot 0} + c_2(0 + e^{-3\cdot 0}) = -3 + c_2$$

or

$$c_2 = 3$$

Therefore, the required solution is

$$y = e^{-3t} + 3te^{-3t} = (1 + 3t)e^{-3t}$$

a + bi

Theorem 6.33 includes the case in which the roots of the characteristic equation are complex. If $\lambda = p + qi$ is a complex root of the equation $\lambda^2 + a\lambda + b = 0$, then so is its conjugate $\bar{\lambda} = p - qi$. (See Appendices C and D.) By Theorem 6.33(a), the solution space S of the differential equation $y'' + ay' + by = 0$ has $\{e^{\lambda t}, e^{\bar{\lambda}t}\}$ as a basis. Now

$$e^{\lambda t} = e^{(p+qi)t} = e^{pt}e^{i(qt)} = e^{pt}(\cos qt + i\sin qt)$$

and

$$e^{\bar{\lambda}t} = e^{(p-qi)t} = e^{pt}e^{i(-qt)} = e^{pt}(\cos qt - i\sin qt)$$

so

$$e^{pt}\cos qt = \frac{e^{\lambda t} + e^{\bar{\lambda}t}}{2} \quad \text{and} \quad e^{pt}\sin qt = \frac{e^{\lambda t} - e^{\bar{\lambda}t}}{2i}$$

It follows that $\{e^{pt}\cos qt, e^{pt}\sin qt\}$ is contained in $\text{span}(e^{\lambda t}, e^{\bar{\lambda}t}) = S$. Since $e^{pt}\cos qt$ and $e^{pt}\sin qt$ are linearly independent (see Exercise 22) and $\dim S = 2$, $\{e^{pt}\cos qt, e^{pt}\sin qt\}$ is also a basis for S. Thus, when its characteristic equation has a complex root $p + qi$, the differential equation $y'' + ay' + by = 0$ has solutions of the form

$$y = c_1e^{pt}\cos qt + c_2e^{pt}\sin qt$$

a + bi

Example 6.91

Find all solutions of $y'' - 2y' + 4 = 0$.

Solution The characteristic equation is $\lambda^2 - 2\lambda + 4 = 0$ with roots $1 \pm i\sqrt{3}$. The foregoing discussion tells us that the general solution to the given differential equation is

$$y = c_1 e^t \cos \sqrt{3}t + c_2 e^t \sin \sqrt{3}t$$

a + bi

Example 6.92

A mass is attached to the end of a vertical spring (Figure 6.21). If the mass is pulled downward and released, it will oscillate up and down. Two laws of physics govern this situation. The first, ***Hooke's law,*** states that if the spring is stretched (or compressed) x units, the force F needed to restore it to its original position is proportional to x:

$$F = -kx$$

where k is a positive constant (called the spring constant). ***Newton's Second Law of Motion*** states that force equals mass times acceleration. Since $x = x(t)$ represents distance, or displacement, of the spring at time t, x' gives its velocity and x'' its acceleration. Thus, we have

$$mx'' = -kx \quad \text{or} \quad x'' + \left(\frac{k}{m}\right)x = 0$$

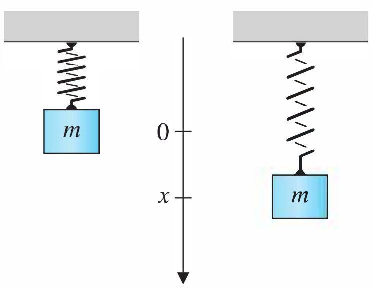


Figure 6.21

Since both k and m are positive, so is $K = k/m$, and our differential equation has the form $x'' + Kx = 0$, where K is positive.

The characteristic equation is $\lambda^2 + K = 0$ with roots $\pm i\sqrt{K}$. Therefore, the general solution to the differential equation of the oscillating spring is

$$x = c_1 \cos \sqrt{K}t + c_2 \sin \sqrt{K}t$$

Suppose the spring is at rest ($x = 0$) at time $t = 0$ seconds and is stretched as far as possible, to a length of 20 cm, before it is released. Then

$$0 = x(0) = c_1 \cos 0 + c_2 \sin 0 = c_1$$

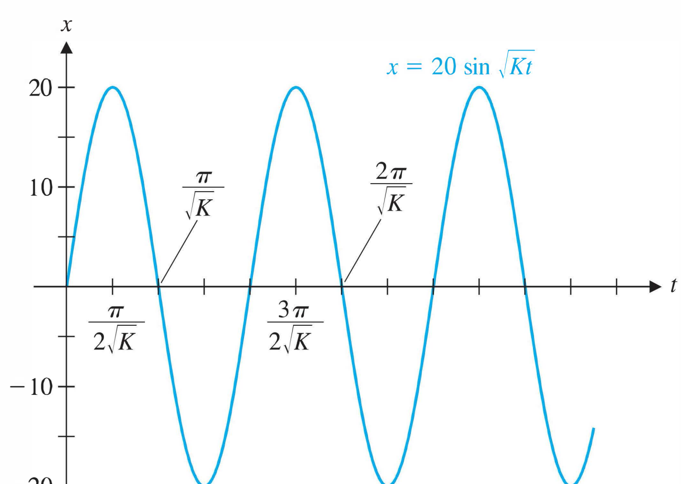


Figure 6.22

so $x = c_2 \sin \sqrt{K}t$. Since the maximum value of the sine function is 1, we must have $c_2 = 20$ (occurring for the first time when $t = \pi/2\sqrt{K}$), giving us the solution

$$x = 20 \sin \sqrt{K}t$$

(See Figure 6.22.)

Of course, this is an idealized solution, since it neglects any form of resistance and predicts that the spring will oscillate forever. It is possible to take damping effects (such as friction) into account, but this simple model has served to introduce an important application of differential equations and the techniques we have developed.

Exercises 6.7

Homogeneous Linear Differential Equations

In Exercises 1–12, find the solution of the differential equation that satisfies the given boundary condition(s).

1. $y' - 3y = 0, y(1) = 2$

2. $x' + x = 0, x(1) = 1$

3. $y'' - 7y' + 12y = 0, y(0) = y(1) = 1$

4. $x'' + x' - 12x = 0, x(0) = 0, x'(0) = 1$

5. $f'' - f' - f = 0, f(0) = 0, f(1) = 1$

6. $g'' - 2g = 0, g(0) = 1, g(1) = 0$

7. $y'' - 2y' + y = 0, y(0) = y(1) = 1$

8. $x'' + 4x' + 4x = 0, x(0) = 1, x'(0) = 1$

9. $y'' - k^2y = 0, k \neq 0, y(0) = y'(0) = 1$

10. $y'' - 2ky' + k^2y = 0, k \neq 0, y(0) = 1, y(1) = 0$

11. $f'' - 2f' + 5f = 0, f(0) = 1, f(\pi/4) = 0$

12. $h'' - 4h' + 5h = 0, h(0) = 0, h'(0) = -1$

13. A strain of bacteria has a growth rate that is proportional to the size of the population. Initially, there are 100 bacteria; after 3 hours, there are 1600.

(a) If $p(t)$ denotes the number of bacteria after t hours, find a formula for $p(t)$.

(b) How long does it take for the population to double?

(c) When will the population reach one million?

CAS **14.** Table 6.2 gives the population of the United States at 10-year intervals for the years 1900–2000.

(a) Assuming an exponential growth model, use the data for 1900 and 1910 to find a formula for $p(t)$, the population in year t. [*Hint:* Let $t = 0$ be 1900 and let $t = 1$ be 1910.] How accurately does your formula calculate the U.S. population in 2000?

(b) Repeat part (a), but use the data for the years 1970 and 1980 to solve for $p(t)$. Does this approach give a better approximation for the year 2000?

(c) What can you conclude about U.S. population growth?

Table 6.2

Year	Population (in millions)
1900	76
1910	92
1920	106
1930	123
1940	131
1950	150
1960	179
1970	203
1980	227
1990	250
2000	281

Source: U.S. Bureau of the Census

15. The half-life of radium-226 is 1590 years. Suppose we start with a sample of radium-226 whose mass is 50 mg.

(a) Find a formula for the mass $m(t)$ remaining after t years and use this formula to predict the mass remaining after 1000 years.

(b) When will only 10 mg remain?

16. *Radiocarbon dating* is a method used by scientists to estimate the age of ancient objects that were once living matter, such as bone, leather, wood, or paper.

All of these contain carbon, a proportion of which is carbon-14, a radioactive isotope that is continuously being formed in the upper atmosphere. Since living organisms take up radioactive carbon along with other carbon atoms, the ratio between the two forms remains constant. However, when an organism dies, the carbon-14 in its cells decays and is not replaced. Carbon-14 has a known half-life of 5730 years, so by measuring the concentration of carbon-14 in an object, scientists can determine its approximate age.

One of the most successful applications of radiocarbon dating has been to determine the age of the Stonehenge monument in England (Figure 6.23). Samples taken from the remains of wooden posts were found to have a concentration of carbon-14 that was 45% of that found in living material. What is the estimated age of these posts?

Figure 6.23
Stonehenge

17. A mass is attached to a spring, as in Example 6.92. At time $t = 0$ second, the spring is stretched to a length of 10 cm below its position at rest. The spring is released, and its length 10 seconds later is observed to be 5 cm. Find a formula for the length of the spring at time t seconds.

18. A 50 g mass is attached to a spring, as in Example 6.92. If the period of oscillation is 10 seconds, find the spring constant.

19. A pendulum consists of a mass, called a *bob*, that is affixed to the end of a string of length L (see Figure 6.24). When the bob is moved from its rest position and released, it swings back and forth. The time it takes the pendulum to swing from its farthest right position to its farthest left position and back to its next farthest right position is called the *period* of the pendulum.

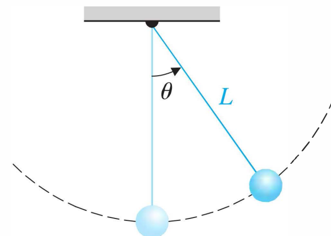


Figure 6.24

Let $\theta = \theta(t)$ be the angle of the pendulum from the vertical. It can be shown that if there is no resistance, then when θ is small it satisfies the differential equation

$$\theta'' + \frac{g}{L}\theta = 0$$

where g is the constant of acceleration due to gravity, approximately 9.7 m/s^2. Suppose that $L = 1$ m and that the pendulum is at rest (i.e., $\theta = 0$) at time $t = 0$ second. The bob is then drawn to the right at an angle of θ_1 radians and released.

(a) Find the period of the pendulum.

(b) Does the period depend on the angle θ_1 at which the pendulum is released? This question was posed and answered by Galileo in 1638. [Galileo Galilei (1564–1642) studied medicine as a student at the University of Pisa, but his real interest was always mathematics. In 1592, Galileo was appointed professor of mathematics at the University of Padua in Venice, where he taught primarily geometry and astronomy. He was the first to use a telescope to look at the stars and planets, and in so doing, he produced experimental data in support of the Copernican view that the planets revolve around the sun and not the earth. For this, Galileo was summoned before the Inquisition, placed under house arrest, and forbidden to publish his results. While under house arrest, he was able to write up his research on falling objects and pendulums. His notes were smuggled out of Italy and published as *Discourses on Two New Sciences* in 1638.]

20. Show that the solution set S of the second-order differential equation $y'' + ay' + by = 0$ is a subspace of $\mathscr{F}$.

21. Prove Theorem 6.33(b).

22. Show that $e^{pt}\cos qt$ and $e^{pt}\sin qt$ are linearly independent.

Chapter Review

Key Definitions and Concepts

basis, 446
Basis Theorem, 453
change-of-basis matrix, 465
composition of linear transformations, 477
coordinate vector, 449
diagonalizable linear transformation, 509
dimension, 453
Fundamental Theorem of Invertible Matrices, 512
identity transformation, 474
invertible linear transformation, 478
isomorphism, 493
kernel of a linear transformation, 482
linear combination of vectors, 433
linear transformation, 472
linearly dependent vectors, 443, 446
linearly independent vectors, 443, 446
matrix of a linear transformation, 498
nullity of a linear transformation, 484
one-to-one, 488
onto, 488
range of a linear transformation, 482
rank of a linear transformation, 484
Rank Theorem, 486
span of a set of vectors, 438
standard basis, 447
subspace, 434
trivial subspace, 437
vector, 429
vector space, 429
zero subspace, 437
zero transformation, 474

Review Questions

1. Mark each of the following statements true or false:

(a) If $V = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, then every spanning set for V contains at least n vectors.

(b) If $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a linearly independent set of vectors, then so is $\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{u} + \mathbf{w}\}$.

(c) M_{22} has a basis consisting of invertible matrices.

(d) M_{22} has a basis consisting of matrices whose trace is zero.

(e) The transformation $T : \mathbb{R}^n \to \mathbb{R}$ defined by $T(\mathbf{x}) = \|\mathbf{x}\|$ is a linear transformation.

(f) If $T : V \to W$ is a linear transformation and dim $V \neq$ dim W, then T cannot be both one-to-one and onto.

(g) If $T : V \to W$ is a linear transformation and $\ker(T) = V$, then $W = \{\mathbf{0}\}$.

(h) If $T : M_{33} \to \mathscr{P}_4$ is a linear transformation and $\text{nullity}(T) = 4$, then T is onto.

(i) The vector space $V = \{p(x) \text{ in } \mathscr{P}_4 : p(1) = 0\}$ is isomorphic to $\mathscr{P}_3$.

(j) If $I : V \to V$ is the identity transformation, then the matrix $[I]_{\mathcal{C}\leftarrow\mathcal{B}}$ is the identity matrix for any bases $\mathcal{B}$ and $\mathcal{C}$ of V.

In Questions 2–5, determine whether W is a subspace of V.

2. $V = \mathbb{R}^2$, $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x^2 + 3y^2 = 0 \right\}$

3. $V = M_{22}$, $W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a + b = c + d = a + c = b + d \right\}$

4. $V = \mathscr{P}_3$, $W = \{p(x) \text{ in } \mathscr{P}_3 : x^3p(1/x) = p(x)\}$

5. $V = \mathscr{F}$, $W = \{f \text{ in } \mathscr{F} : f(x + \pi) = f(x) \text{ for all } x\}$

6. Determine whether $\{1, \cos 2x, 3\sin^2 x\}$ is linearly dependent or independent.

7. Let A and B be nonzero $n \times n$ matrices such that A is symmetric and B is skew-symmetric. Prove that $\{A, B\}$ is linearly independent.

In Questions 8 and 9, find a basis for W and state the dimension of W.

8. $W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a + d = b + c \right\}$

9. $W = \{p(x) \text{ in } \mathscr{P}_5 : p(-x) = p(x)\}$

10. Find the change-of-basis matrices $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$ with respect to the bases $\mathcal{B} = \{1, 1 + x, 1 + x + x^2\}$ and $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\}$ of $\mathscr{P}_2$.

In Questions 11–13, determine whether T is a linear transformation.

11. $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathbf{x}) = \mathbf{y}\mathbf{x}^T\mathbf{y}$, where $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

12. $T : M_{nn} \to M_{nn}$ defined by $T(A) = A^T A$

13. $T : \mathcal{P}_n \to \mathcal{P}_n$ defined by $T(p(x)) = p(2x - 1)$

14. If $T: \mathcal{P}_2 \to M_{22}$ is a linear transformation such that $T(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $T(1 + x) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $T(1 + x + x^2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, find $T(5 - 3x + 2x^2)$.

15. Find the nullity of the linear transformation $T : M_{nn} \to \mathbb{R}$ defined by $T(A) = \text{tr}(A)$.

16. Let W be the vector space of upper triangular 2×2 matrices.

(a) Find a linear transformation $T : M_{22} \to M_{22}$ such that $\ker(T) = W$.

(b) Find a linear transformation $T : M_{22} \to M_{22}$ such that $\text{range}(T) = W$.

17. Find the matrix $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ of the linear transformation T in Question 14 with respect to the standard bases $\mathcal{B} = \{1, x, x^2\}$ of $\mathcal{P}_2$ and $\mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ of M_{22}.

18. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in a vector space V with the property that every vector in V can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in exactly one way. Prove that S is a basis for V.

19. If $T : U \to V$ and $S : V \to W$ are linear transformations such that $\text{range}(T) \subseteq \ker(S)$, what can be deduced about $S \circ T$?

20. Let $T : V \to V$ be a linear transformation, and let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V such that $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is also a basis for V. Prove that T is invertible.