CHAPTER

Linear Transformations

5

TRANSFORMING SPACE

Although a vector can be used to indicate a particular type of movement, actual vectors themselves are essentially static, unchanging objects. For example, if we represent the edges of a particular image on a computer screen by vectors, then these vectors are fixed in place. However, when we want to move or alter the image in some way, such as rotating it about a point on the screen, we need a function to calculate the new position for each of the original vectors.

This suggests that we need another "tool" in our arsenal: functions that move a given set of vectors in a prescribed "linear" manner. Such functions are called linear transformations. Just as we saw in Chapter 4 that general vector spaces are abstract generalizations of $\mathbb{R}^n$, we will find in this chapter that linear transformations are the corresponding abstract generalization of matrix multiplication.

In this chapter, we study functions that map the vectors in one vector space to those in another. We concentrate on a special class of these functions, known as linear transformations. The formal definition of a linear transformation is introduced in Section 5.1 along with several of its fundamental properties. In Section 5.2, we show that the effect of any linear transformation is equivalent to multiplication by a corresponding matrix. In Section 5.3, we examine an important relationship between the dimensions of the domain and the range of a linear transformation, known as the Dimension Theorem. In Section 5.4, we introduce two special types of linear transformations: one-to-one and onto. In Section 5.5, these two types of linear transformations are combined to form isomorphisms, which are used to establish that all n-dimensional vector spaces are in some sense equivalent. Finally, in Section 5.6, we return to the topic of eigenvalues and eigenvectors to study them in the context of linear transformations.

Elementary Linear Algebra

5.1 INTRODUCTION TO LINEAR TRANSFORMATIONS

In this section, we introduce linear transformations and examine their elementary properties.

Functions

If you are not familiar with the terms *domain*, *codomain*, *range*, *image*, and *pre-image* in the context of functions, read Appendix B before proceeding. The following example illustrates some of these terms:

Example 1

Let $f\colon \mathcal{M}_{23} \to \mathcal{M}_{22}$ be given by

$$f\left(\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\right) = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}.$$

Then f is a function that maps one vector space to another. The domain of f is $\mathcal{M}_{23}$, the codomain of f is $\mathcal{M}_{22}$, and the range of f is the set of all 2×2 matrices with second row entries equal to zero. The image of $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ under f is $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$. The matrix $\begin{bmatrix} 1 & 2 & 10 \\ 11 & 12 & 13 \end{bmatrix}$ is one of the pre-images of $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ under f. Also, the image under f of the set S of all matrices of the form $\begin{bmatrix} 7 & * & * \\ * & * & * \end{bmatrix}$ (where "$*$" represents any real number) is the set $f(S)$ containing all matrices of the form $\begin{bmatrix} 7 & * \\ 0 & 0 \end{bmatrix}$. Finally, the pre-image under f of the set T of all matrices of the form $\begin{bmatrix} a & a+2 \\ 0 & 0 \end{bmatrix}$ is the set $f^{-1}(T)$ consisting of all matrices of the form $\begin{bmatrix} a & a+2 & * \\ * & * & * \end{bmatrix}$. ∎

Linear Transformations

Definition Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, and let $f\colon \mathcal{V} \to \mathcal{W}$ be a function from $\mathcal{V}$ to $\mathcal{W}$. (That is, for each vector $\mathbf{v} \in \mathcal{V}$, $f(\mathbf{v})$ denotes exactly one vector of $\mathcal{W}$.) Then f is a **linear transformation** if and only if both of the following are true:

(1) $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$, for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$

(2) $f(c\mathbf{v}) = cf(\mathbf{v})$, for all $c \in \mathbb{R}$ and all $\mathbf{v} \in \mathcal{V}$.

Properties (1) and (2) insist that the operations of addition and scalar multiplication give the same result on vectors whether the operations are performed before

f is applied (in $\mathcal{V}$) or after f is applied (in $\mathcal{W}$). Thus, a linear transformation is a function between vector spaces that "preserves" the operations that give structure to the spaces.

To determine whether a given function f from a vector space $\mathcal{V}$ to a vector space $\mathcal{W}$ is a linear transformation, we need only verify properties (1) and (2) in the definition, as in the next three examples.

Example 2

Consider the mapping $f\colon \mathcal{M}_{mn} \to \mathcal{M}_{nm}$, given by $f(\mathbf{A}) = \mathbf{A}^T$ for any $m \times n$ matrix $\mathbf{A}$. We will show that f is a linear transformation.

(1) We must show that $f(\mathbf{A}_1 + \mathbf{A}_2) = f(\mathbf{A}_1) + f(\mathbf{A}_2)$, for matrices $\mathbf{A}_1, \mathbf{A}_2 \in \mathcal{M}_{mn}$. However, $f(\mathbf{A}_1 + \mathbf{A}_2) = (\mathbf{A}_1 + \mathbf{A}_2)^T = \mathbf{A}_1^T + \mathbf{A}_2^T$ (by part (2) of Theorem 1.12) $= f(\mathbf{A}_1) + f(\mathbf{A}_2)$.

(2) We must show that $f(c\mathbf{A}) = cf(\mathbf{A})$, for all $c \in \mathbb{R}$ and for all $\mathbf{A} \in \mathcal{M}_{mn}$. However, $f(c\mathbf{A}) = (c\mathbf{A})^T = c(\mathbf{A}^T)$ (by part (3) of Theorem 1.12) $= cf(\mathbf{A})$.

Hence, f is a linear transformation. ■

Example 3

Consider the function $g : \mathcal{P}_n \to \mathcal{P}_{n-1}$ given by $g(\mathbf{p}) = \mathbf{p}'$, the derivative of $\mathbf{p}$. We will show that g is a linear transformation.

(1) We must show that $g(\mathbf{p}_1 + \mathbf{p}_2) = g(\mathbf{p}_1) + g(\mathbf{p}_2)$, for all $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}_n$. Now, $g(\mathbf{p}_1 + \mathbf{p}_2) = \left(\mathbf{p}_1 + \mathbf{p}_2\right)'$. From calculus we know that the derivative of a sum is the sum of the derivatives, so $\left(\mathbf{p}_1 + \mathbf{p}_2\right)' = \mathbf{p}_1' + \mathbf{p}_2' = g(\mathbf{p}_1) + g(\mathbf{p}_2)$.

(2) We must show that $g(c\mathbf{p}) = cg(\mathbf{p})$, for all $c \in \mathbb{R}$ and $\mathbf{p} \in \mathcal{P}_n$. Now, $g(c\mathbf{p}) = \left(c\mathbf{p}\right)'$. Again, from calculus we know that the derivative of a constant times a function is equal to the constant times the derivative of the function, so $(c\mathbf{p})' = c(\mathbf{p}') = cg(\mathbf{p})$.

Hence, g is a linear transformation. ■

Example 4

Let $\mathcal{V}$ be a finite dimensional vector space, and let B be an ordered basis for $\mathcal{V}$. Then every element $\mathbf{v} \in \mathcal{V}$ has its coordinatization $[\mathbf{v}]_B$ with respect to B. Consider the mapping $f\colon \mathcal{V} \to \mathbb{R}^n$ given by $f(\mathbf{v}) = [\mathbf{v}]_B$. We will show that f is a linear transformation.

Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$. By Theorem 4.20, $[\mathbf{v}_1 + \mathbf{v}_2]_B = [\mathbf{v}_1]_B + [\mathbf{v}_2]_B$. Hence,

$$f(\mathbf{v}_1 + \mathbf{v}_2) = [\mathbf{v}_1 + \mathbf{v}_2]_B = [\mathbf{v}_1]_B + [\mathbf{v}_2]_B = f(\mathbf{v}_1) + f(\mathbf{v}_2).$$

Next, let $c \in \mathbb{R}$ and $\mathbf{v} \in \mathcal{V}$. Again by Theorem 4.20, $[c\mathbf{v}]_B = c\,[\mathbf{v}]_B$. Hence,

$$f(c\mathbf{v}) = [c\mathbf{v}]_B = c\,[\mathbf{v}]_B = cf(\mathbf{v}).$$

Thus, f is a linear transformation from $\mathcal{V}$ to $\mathbb{R}^n$. ■

Not every function between vector spaces is a linear transformation. For example, consider the function $h:\mathbb{R}^2 \to \mathbb{R}^2$ given by $h([x,y]) = [x+1, y-2] = [x,y] + [1,-2]$. In this case, h merely adds $[1,-2]$ to each vector $[x,y]$ (see Figure 5.1). This type of mapping is called a **translation**. However, h is not a linear transformation. To show that it is not, we have to produce a counterexample to verify that either property (1) or property (2) of the definition fails. Property (1) fails, since $h([1,2]+[3,4]) = h([4,6]) = [5,4]$, while $h([1,2]) + h([3,4]) = [2,0] + [4,2] = [6,2]$.

In general, when given a function f between vector spaces, we do not always know right away whether f is a linear transformation. If we suspect that either property (1) or (2) does not hold for f, then we look for a counterexample.

Linear Operators and Some Geometric Examples

An important type of linear transformation is one that maps a vector space to itself.

Definition Let $\mathcal{V}$ be a vector space. A **linear operator** on $\mathcal{V}$ is a linear transformation whose domain and codomain are both $\mathcal{V}$.

Example 5

If $\mathcal{V}$ is any vector space, then the mapping $i\colon \mathcal{V} \to \mathcal{V}$ given by $i(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in \mathcal{V}$ is a linear operator, known as the **identity linear operator**. Also, the constant mapping $z\colon \mathcal{V} \to \mathcal{V}$ given by $z(\mathbf{v}) = \mathbf{0}_{\mathcal{V}}$ is a linear operator known as the **zero linear operator** (see Exercise 2). ■

The next few examples exhibit important geometric operators. In these examples, assume that all vectors begin at the origin.

Example 6

Reflections: Consider the mapping $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by $f([a_1,a_2,a_3]) = [a_1,a_2,-a_3]$. This mapping "reflects" the vector $[a_1,a_2,a_3]$ through the xy-plane, which acts like a "mirror" (see

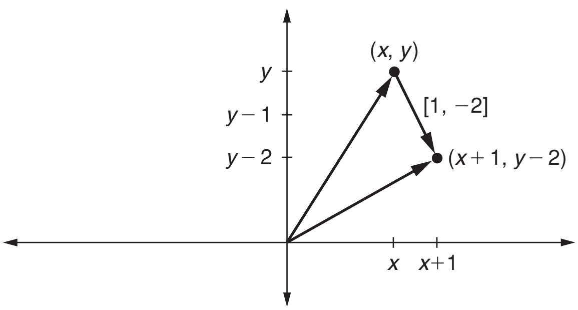

FIGURE 5.1

A translation in $\mathbb{R}^2$

Figure 5.2). Now, since

$$\begin{aligned} f([a_1,a_2,a_3]+[b_1,b_2,b_3]) &= f([a_1+b_1,a_2+b_2,a_3+b_3]) \\ &= [a_1+b_1,a_2+b_2,-(a_3+b_3)] \\ &= [a_1,a_2,-a_3]+[b_1,b_2,-b_3] \\ &= f([a_1,a_2,a_3])+f([b_1,b_2,b_3]), \quad \text{and} \end{aligned}$$

$$f(c[a_1,a_2,a_3]) = f([ca_1,ca_2,ca_3]) = [ca_1,ca_2,-ca_3] = c[a_1,a_2,-a_3] = cf([a_1,a_2,a_3]),$$

we see that f is a linear operator. Similarly, reflection through the xz-plane or the yz-plane is also a linear operator on $\mathbb{R}^3$ (see Exercise 4).

Example 7

Contractions and Dilations: Consider the mapping $g: \mathbb{R}^n \to \mathbb{R}^n$ given by scalar multiplication by k, where $k \in \mathbb{R}$; that is, $g(\mathbf{v}) = k\mathbf{v}$, for $\mathbf{v} \in \mathbb{R}^n$. The function g is a linear operator (see Exercise 3). If $|k| > 1$, g represents a **dilation** (lengthening) of the vectors in $\mathbb{R}^n$; if $|k| < 1$, g represents a **contraction** (shrinking).

Example 8

Projections: Consider the mapping $h: \mathbb{R}^3 \to \mathbb{R}^3$ given by $h([a_1,a_2,a_3]) = [a_1,a_2,0]$. This mapping takes each vector in $\mathbb{R}^3$ to a corresponding vector in the xy-plane (see Figure 5.3). Similarly,

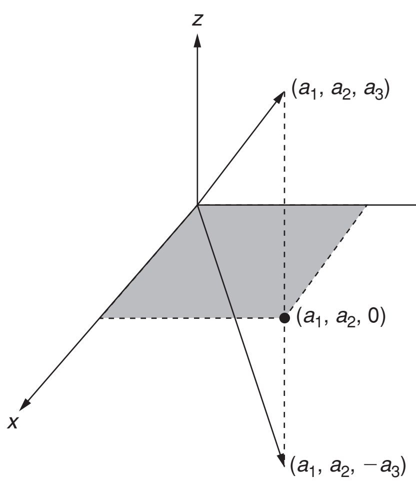

FIGURE 5.2

Reflection in $\mathbb{R}^3$ through the xy-plane

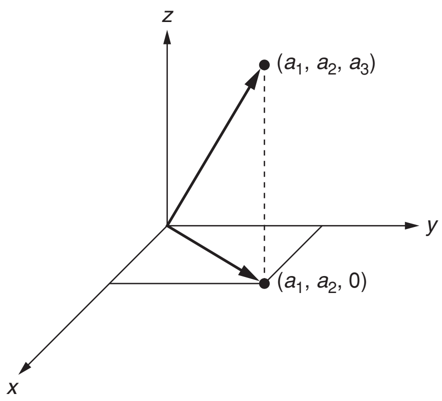

FIGURE 5.3

Projection of $[a_1, a_2, a_3]$ to the xy-plane

consider the mapping $j: \mathbb{R}^4 \to \mathbb{R}^4$ given by $j([a_1, a_2, a_3, a_4]) = [0, a_2, 0, a_4]$. This mapping takes each vector in $\mathbb{R}^4$ to a corresponding vector whose first and third coordinates are zero. The functions h and j are both linear operators (see Exercise 5). Such mappings, where at least one of the coordinates is "zeroed out," are examples of **projection mappings**. You can verify that all such mappings are linear operators. (Other types of projection mappings are illustrated in Exercises 6 and 7.)

■

Example 9

Rotations: Let θ be a fixed angle in $\mathbb{R}^2$, and let $l: \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$l\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}.$$

In Exercise 9 you are asked to show that l rotates $[x, y]$ counterclockwise through the angle θ (see Figure 5.4).

Now, let $\mathbf{v}_1 = [x_1, y_1]$ and $\mathbf{v}_2 = [x_2, y_2]$ be two vectors in $\mathbb{R}^2$. Then,

$$\begin{aligned} l(\mathbf{v}_1 + \mathbf{v}_2) &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}(\mathbf{v}_1 + \mathbf{v}_2) \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\mathbf{v}_1 + \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\mathbf{v}_2 \\ &= l(\mathbf{v}_1) + l(\mathbf{v}_2). \end{aligned}$$

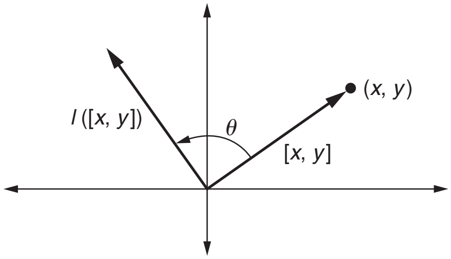

FIGURE 5.4

Counterclockwise rotation of $[x, y]$ through an angle θ in $\mathbb{R}^2$

Similarly, $l(c\mathbf{v}) = cl(\mathbf{v})$, for any $c \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^2$. Hence, l is a linear operator. ∎

Beware! Not all geometric operations are linear operators. Recall that the translation function is not a linear operator!

Multiplication Transformation

The linear operator in Example 9 is actually a special case of the next example, which shows that multiplication by an $m \times n$ matrix is always a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Example 10

Let $\mathbf{A}$ be a given $m \times n$ matrix. We show that the function $f: \mathbb{R}^n \to \mathbb{R}^m$ defined by $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, for all $\mathbf{x} \in \mathbb{R}^n$, is a linear transformation. Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$. Then $f(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{A}\mathbf{x}_1 + \mathbf{A}\mathbf{x}_2 = f(\mathbf{x}_1) + f(\mathbf{x}_2)$. Also, let $\mathbf{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then, $f(c\mathbf{x}) = \mathbf{A}(c\mathbf{x}) = c(\mathbf{A}\mathbf{x}) = cf(\mathbf{x})$. ∎

For a specific example of the multiplication transformation, consider the matrix $\mathbf{A} = \begin{bmatrix} -1 & 4 & 2 \\ 5 & 6 & -3 \end{bmatrix}$. The mapping given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -1 & 4 & 2 \\ 5 & 6 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_1 + 4x_2 + 2x_3 \\ 5x_1 + 6x_2 - 3x_3 \end{bmatrix}$$

is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$. In the next section, we will show that the converse of the result in Example 10 also holds; every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is equivalent to multiplication by an appropriate $m \times n$ matrix.

Elementary Properties of Linear Transformations

We now prove some basic properties of linear transformations. From here on, we usually use italicized capital letters, such as "L," to represent linear transformations.

Theorem 5.1 Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, and let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Let $\mathbf{0}_{\mathcal{V}}$ be the zero vector in $\mathcal{V}$ and $\mathbf{0}_{\mathcal{W}}$ be the zero vector in $\mathcal{W}$. Then

(1) $L(\mathbf{0}_{\mathcal{V}}) = \mathbf{0}_{\mathcal{W}}$

(2) $L(-\mathbf{v}) = -L(\mathbf{v})$, for all $\mathbf{v} \in \mathcal{V}$

(3) $L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) = a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2) + \cdots + a_nL(\mathbf{v}_n)$, for all $a_1, \ldots, a_n \in \mathbb{R}$, and $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{V}$, for $n \geq 2$.

Proof.

Part (1):

$$\begin{aligned} L(\mathbf{0}_{\mathcal{V}}) &= L(0\mathbf{0}_{\mathcal{V}}) && \text{part (2) of Theorem 4.1, in } \mathcal{V} \\ &= 0L(\mathbf{0}_{\mathcal{V}}) && \text{property (2) of linear transformation} \\ &= \mathbf{0}_{\mathcal{W}} && \text{part (2) of Theorem 4.1, in } \mathcal{W} \end{aligned}$$

Part (2):

$$\begin{aligned} L(-\mathbf{v}) &= L(-1\mathbf{v}) && \text{part (3) of Theorem 4.1, in } \mathcal{V} \\ &= -1(L(\mathbf{v})) && \text{property (2) of linear transformation} \\ &= -L(\mathbf{v}) && \text{part (3) of Theorem 4.1, in } \mathcal{W} \end{aligned}$$

Part (3): (Abridged) This part is proved by induction. We prove the Base Step $(n = 2)$ here and leave the Inductive Step as Exercise 29. For the Base Step, we must show that $L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) = a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2)$. But,

$$\begin{aligned} L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) &= L(a_1\mathbf{v}_1) + L(a_2\mathbf{v}_2) && \text{property (1) of linear transformation} \\ &= a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2) && \text{property (2) of linear transformation.} \end{aligned}$$ □

The next theorem asserts that the composition $L_2 \circ L_1$ of linear transformations L_1 and L_2 is again a linear transformation (see Appendix B for a review of composition of functions).

Theorem 5.2 Let $\mathcal{V}_1, \mathcal{V}_2$, and $\mathcal{V}_3$ be vector spaces. Let $L_1: \mathcal{V}_1 \to \mathcal{V}_2$ and $L_2: \mathcal{V}_2 \to \mathcal{V}_3$ be linear transformations. Then $L_2 \circ L_1: \mathcal{V}_1 \to \mathcal{V}_3$ given by $(L_2 \circ L_1)(\mathbf{v}) = L_2(L_1(\mathbf{v}))$, for all $\mathbf{v} \in \mathcal{V}_1$, is a linear transformation.

Proof. (Abridged) To show that $L_2 \circ L_1$ is a linear transformation, we must show that for all $c \in \mathbb{R}$ and $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$,

$$(L_2 \circ L_1)(\mathbf{v}_1 + \mathbf{v}_2) = (L_2 \circ L_1)(\mathbf{v}_1) + (L_2 \circ L_1)(\mathbf{v}_2)$$
$$\text{and} \quad (L_2 \circ L_1)(c\mathbf{v}) = c(L_2 \circ L_1)(\mathbf{v}).$$

The first property holds since

$$\begin{aligned}
(L_2 \circ L_1)(\mathbf{v}_1 + \mathbf{v}_2) &= L_2(L_1(\mathbf{v}_1 + \mathbf{v}_2)) \\
&= L_2(L_1(\mathbf{v}_1) + L_1(\mathbf{v}_2)) && \text{because } L_1 \text{ is a linear transformation} \\
&= L_2(L_1(\mathbf{v}_1)) + L_2(L_1(\mathbf{v}_2)) && \text{because } L_2 \text{ is a linear transformation} \\
&= (L_2 \circ L_1)(\mathbf{v}_1) + (L_2 \circ L_1)(\mathbf{v}_2).
\end{aligned}$$

We leave the proof of the second property as Exercise 33. □

Example 11

Let L_1 represent the rotation of vectors in $\mathbb{R}^2$ through a fixed angle θ (as in Example 9), and let L_2 represent the reflection of vectors in $\mathbb{R}^2$ through the x-axis. That is, if $\mathbf{v} = [v_1, v_2]$, then

$$L_1(\mathbf{v}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{and} \quad L_2(\mathbf{v}) = \begin{bmatrix} v_1 \\ -v_2 \end{bmatrix}.$$

Because L_1 and L_2 are both linear transformations, Theorem 5.2 asserts that

$$L_2(L_1(\mathbf{v})) = L_2\left(\begin{bmatrix} v_1\cos\theta - v_2\sin\theta \\ v_1\sin\theta + v_2\cos\theta \end{bmatrix}\right) = \begin{bmatrix} v_1\cos\theta - v_2\sin\theta \\ -v_1\sin\theta - v_2\cos\theta \end{bmatrix}$$

is also a linear transformation. $L_2 \circ L_1$ represents a rotation of $\mathbf{v}$ through θ followed by a reflection through the x-axis. ■

Theorem 5.2 generalizes naturally to more than two linear transformations. That is, if $L_1, L_2, \ldots, L_k$ are linear transformations and the composition $L_k \circ \cdots \circ L_2 \circ L_1$ makes sense, then $L_k \circ \cdots \circ L_2 \circ L_1$ is also a linear transformation.

Linear Transformations and Subspaces

The final theorem of this section assures us that, under a linear transformation $L: \mathcal{V} \to \mathcal{W}$, subspaces of $\mathcal{V}$ "correspond" to subspaces of $\mathcal{W}$, and vice versa.

Theorem 5.3 Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation.

(1) If $\mathcal{V}'$ is a subspace of $\mathcal{V}$, then $L(\mathcal{V}') = \{L(\mathbf{v}) \mid \mathbf{v} \in \mathcal{V}'\}$, the image of $\mathcal{V}'$ in $\mathcal{W}$, is a subspace of $\mathcal{W}$. In particular, the range of L is a subspace of $\mathcal{W}$.

(2) If $\mathcal{W}'$ is a subspace of $\mathcal{W}$, then $L^{-1}(\mathcal{W}') = \{\mathbf{v} \mid L(\mathbf{v}) \in \mathcal{W}'\}$, the pre-image of $\mathcal{W}'$ in $\mathcal{V}$, is a subspace of $\mathcal{V}$.

We prove part (1) and leave part (2) as Exercise 31.

Proof. Part (1): Suppose that $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and that $\mathcal{V}'$ is a subspace of $\mathcal{V}$. Now, $L(\mathcal{V}')$, the image of $\mathcal{V}'$ in $\mathcal{W}$ (see Figure 5.5), is certainly nonempty (why?). Hence, to show that $L(\mathcal{V}')$ is a subspace of $\mathcal{W}$, we must prove that $L(\mathcal{V}')$ is closed under addition and scalar multiplication.

First, suppose that $\mathbf{w}_1, \mathbf{w}_2 \in L(\mathcal{V}')$. Then, by definition of $L(\mathcal{V}')$, we have $\mathbf{w}_1 = L(\mathbf{v}_1)$ and $\mathbf{w}_2 = L(\mathbf{v}_2)$, for some $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}'$. Then, $\mathbf{w}_1 + \mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1 + \mathbf{v}_2)$ because L is a linear transformation. However, since $\mathcal{V}'$ is a subspace of $\mathcal{V}$, $(\mathbf{v}_1 + \mathbf{v}_2) \in \mathcal{V}'$. Thus, $(\mathbf{w}_1 + \mathbf{w}_2)$ is the image of $(\mathbf{v}_1 + \mathbf{v}_2) \in \mathcal{V}'$, and so $(\mathbf{w}_1 + \mathbf{w}_2) \in L(\mathcal{V}')$. Hence, $L(\mathcal{V}')$ is closed under addition.

Next, suppose that $c \in \mathbb{R}$ and $\mathbf{w} \in L(\mathcal{V}')$. By definition of $L(\mathcal{V}')$, $\mathbf{w} = L(\mathbf{v})$, for some $\mathbf{v} \in \mathcal{V}'$. Then, $c\mathbf{w} = cL(\mathbf{v}) = L(c\mathbf{v})$ since L is a linear transformation. Now, $c\mathbf{v} \in \mathcal{V}'$, because $\mathcal{V}'$ is a subspace of $\mathcal{V}$. Thus, $c\mathbf{w}$ is the image of $c\mathbf{v} \in \mathcal{V}'$, and so $c\mathbf{w} \in L(\mathcal{V}')$. Hence, $L(\mathcal{V}')$ is closed under scalar multiplication. $\square$

Example 12

Let $L: \mathcal{M}_{22} \to \mathbb{R}^3$, where $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = [b, 0, c]$. L is a linear transformation (verify!). By Theorem 5.3, the range of any linear transformation is a subspace of the codomain. Hence, the range of $L = \{[b, 0, c] \mid b, c \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$.

Also, consider the subspace $\mathcal{U}_2 = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \middle| \, a, b, d \in \mathbb{R} \right\}$ of $\mathcal{M}_{22}$. Then the image of $\mathcal{U}_2$ under L is $\{[b, 0, 0] \mid b \in \mathbb{R}\}$. This image is a subspace of $\mathbb{R}^3$, as Theorem 5.3 asserts. Finally, consider the subspace $\mathcal{W} = \{[b, e, 2b] \mid b, e \in \mathbb{R}\}$ of $\mathbb{R}^3$. The pre-image of $\mathcal{W}$ consists of all

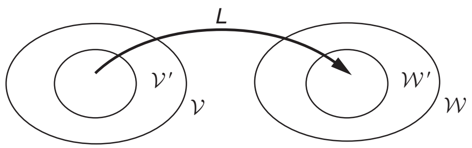

FIGURE 5.5

Subspaces of $\mathcal{V}$ correspond to subspaces of $\mathcal{W}$ under a linear transformation $L: \mathcal{V} \to \mathcal{W}$

matrices in $\mathcal{M}_{22}$ of the form $\begin{bmatrix} a & b \\ 2b & d \end{bmatrix}$. Notice that this pre-image is a subspace of $\mathcal{M}_{22}$, as claimed by Theorem 5.3. ■

New Vocabulary

codomain (of a linear transformation)
composition of linear transformations
contraction (mapping)
dilation (mapping)
domain (of a linear transformation)
identity linear operator
image (of a vector in the domain)
linear operator
linear transformation
pre-image (of a vector in the codomain)
projection (mapping)
range (of a linear transformation)
reflection (mapping)
rotation (mapping)
shear (mapping)
translation (mapping)
zero linear operator

Highlights

- A linear transformation is a function from one vector space to another that preserves the operations of addition and scalar multiplication. That is, under a linear transformation, the image of a linear combination of vectors is the linear combination of the images of the vectors having the same coefficients.
- A linear operator is a linear transformation from a vector space to itself.
- A nontrivial translation of the plane ($\mathbb{R}^2$) or of space ($\mathbb{R}^3$) is never a linear operator, but all of the following are linear operators: contraction (of $\mathbb{R}^n$), dilation (of $\mathbb{R}^n$), reflection of space through the xy-plane (or xz-plane or yz-plane), rotation of the plane about the origin through a given angle θ, projection (of $\mathbb{R}^n$) in which one or more of the coordinates are zeroed out.
- Multiplication of vectors in $\mathbb{R}^n$ on the left by a fixed $m \times n$ matrix $\mathbf{A}$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.
- Multiplying a vector on the left by the matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is equivalent to rotating the vector counterclockwise about the origin through the angle θ.
- Linear transformations always map the zero vector of the domain to the zero vector of the codomain.
- A composition of linear transformations is a linear transformation.
- Under a linear transformation, subspaces of the domain map to subspaces of the codomain, and the pre-image of a subspace of the codomain is a subspace of the domain.

EXERCISES FOR SECTION 5.1

1. Determine which of the following functions are linear transformations. Prove that your answers are correct. Which are linear operators?

  ⋆(a) $f: \mathbb{R}^2 \to \mathbb{R}^2$ given by $f([x,y]) = [3x - 4y, -x + 2y]$

  ⋆(b) $h: \mathbb{R}^4 \to \mathbb{R}^4$ given by $h([x_1,x_2,x_3,x_4]) = [x_1 + 2, x_2 - 1, x_3, -3]$

  (c) $k: \mathbb{R}^3 \to \mathbb{R}^3$ given by $k([x_1,x_2,x_3]) = [x_2, x_3, x_1]$

  ⋆(d) $l: \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $l\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a - 2c + d & 3b - c \\ -4a & b + c - 3d \end{bmatrix}$

  (e) $n: \mathcal{M}_{22} \to \mathbb{R}$ given by $n\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$

  ⋆(f) $r: \mathcal{P}_3 \to \mathcal{P}_2$ given by $r(ax^3 + bx^2 + cx + d) = (\sqrt[3]{a})x^2 - b^2x + c$

  (g) $s: \mathbb{R}^3 \to \mathbb{R}^3$ given by $s([x_1,x_2,x_3]) = [\cos x_1, \sin x_2, e^{x_3}]$

  ⋆(h) $t: \mathcal{P}_3 \to \mathbb{R}$ given by $t(a_3x^3 + a_2x^2 + a_1x + a_0) = a_3 + a_2 + a_1 + a_0$

  (i) $u: \mathbb{R}^4 \to \mathbb{R}$ given by $u\left([x_1,x_2,x_3,x_4]\right) = |x_2|$

  ⋆(j) $v: \mathcal{P}_2 \to \mathbb{R}$ given by $v\left(ax^2 + bx + c\right) = abc$

  ⋆(k) $g: \mathcal{M}_{32} \to \mathcal{P}_4$ given by $g\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}\right) = a_{11}x^4 - a_{21}x^2 + a_{31}$

  ⋆(l) $e: \mathbb{R}^2 \to \mathbb{R}$ given by $e([x,y]) = \sqrt{x^2 + y^2}$

2. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces.

  (a) Show that the identity mapping $i: \mathcal{V} \to \mathcal{V}$ given by $i(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v} \in \mathcal{V}$, is a linear operator.

  (b) Show that the zero mapping $z: \mathcal{V} \to \mathcal{W}$ given by $z(\mathbf{v}) = \mathbf{0}_{\mathcal{W}}$, for all $\mathbf{v} \in \mathcal{V}$, is a linear transformation.

3. Let k be a fixed scalar in $\mathbb{R}$. Show that the mapping $f: \mathbb{R}^n \to \mathbb{R}^n$ given by $f([x_1,x_2,\ldots,x_n]) = k[x_1,x_2,\ldots,x_n]$ is a linear operator.

4. (a) Show that $f: \mathbb{R}^3 \to \mathbb{R}^3$ given by $f([x,y,z]) = [-x,y,z]$ (reflection of a vector through the yz-plane) is a linear operator.

  (b) What mapping from $\mathbb{R}^3$ to $\mathbb{R}^3$ would reflect a vector through the xz-plane? Is it a linear operator? Why or why not?

  (c) What mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ would reflect a vector through the y-axis? through the x-axis? Are these linear operators? Why or why not?

5. Show that the projection mappings $h: \mathbb{R}^3 \to \mathbb{R}^3$ given by $h([a_1,a_2,a_3]) = [a_1,a_2,0]$ and $j: \mathbb{R}^4 \to \mathbb{R}^4$ given by $j([a_1,a_2,a_3,a_4]) = [0,a_2,0,a_4]$ are linear operators.

6. The mapping $f: \mathbb{R}^n \to \mathbb{R}$ given by $f([x_1, x_2, \ldots, x_i, \ldots, x_n]) = x_i$ is another type of projection mapping. Show that f is a linear transformation.

7. Let $\mathbf{x}$ be a fixed nonzero vector in $\mathbb{R}^3$. Show that the mapping $g: \mathbb{R}^3 \to \mathbb{R}^3$ given by $g(\mathbf{y}) = \mathbf{proj_x y}$ is a linear operator.

8. Let $\mathbf{x}$ be a fixed vector in $\mathbb{R}^n$. Prove that $L: \mathbb{R}^n \to \mathbb{R}$ given by $L(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ is a linear transformation.

9. Let θ be a fixed angle in the xy-plane. Show that the linear operator $L: \mathbb{R}^2 \to \mathbb{R}^2$ given by $L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ rotates the vector $[x, y]$ counterclockwise through the angle θ in the plane. (Hint: Consider the vector $[x', y']$, obtained by rotating $[x, y]$ counterclockwise through the angle θ. Let $r = \sqrt{x^2 + y^2}$. Then $x = r\cos\alpha$ and $y = r\sin\alpha$, where α is the angle shown in Figure 5.6. Notice that $x' = r(\cos(\theta + \alpha))$ and $y' = r(\sin(\theta + \alpha))$. Then show that $L([x, y]) = [x', y']$.)

10. (a) Explain why the mapping $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

is a linear operator.

(b) Show that the mapping L in part (a) rotates every vector in $\mathbb{R}^3$ about the z-axis through an angle of θ (as measured relative to the xy-plane).

★(c) What matrix should be multiplied times $[x, y, z]$ to create the linear operator that rotates $\mathbb{R}^3$ about the y-axis through an angle θ (relative to the xz-plane)? (Hint: When looking down from the positive y-axis toward

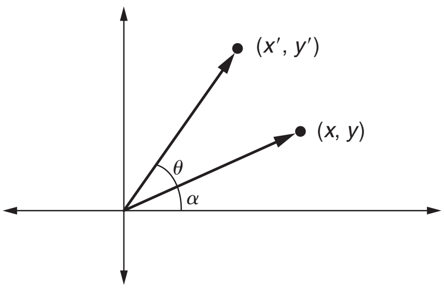

FIGURE 5.6

The vectors $[x, y]$ and $[x', y']$

the xz-plane in a right-handed system, the positive z-axis rotates 90° counterclockwise into the positive x-axis.)

11. **Shears**: Let $f_1, f_2: \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$f_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + ky \\ y \end{bmatrix}$$

and

$$f_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ kx + y \end{bmatrix}.$$

The mapping f_1 is called a **shear in the x-direction with factor k**; f_2 is called a **shear in the y-direction with factor k**. The effect of these functions (for $k > 1$) on the vector $[1, 1]$ is shown in Figure 5.7. Show that f_1 and f_2 are linear operators directly, without using Example 10.

12. Let f: $\mathcal{M}_{nn} \to \mathbb{R}$ be given by $f(\mathbf{A}) = \text{trace}(\mathbf{A})$. (The trace is defined in Exercise 14 of Section 1.4.) Prove that f is a linear transformation.

13. Show that the mappings $g, h: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $g(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T$ and $h(\mathbf{A}) = \mathbf{A} - \mathbf{A}^T$ are linear operators on $\mathcal{M}_{nn}$.

14. **(a)** Show that if $\mathbf{p} \in \mathcal{P}_n$, then the (indefinite integral) function $f: \mathcal{P}_n \to \mathcal{P}_{n+1}$, where $f(\mathbf{p})$ is the vector $\int \mathbf{p}(x)\,dx$ with zero constant term, is a linear transformation.

(b) Show that if $\mathbf{p} \in \mathcal{P}_n$, then the (definite integral) function $g: \mathcal{P}_n \to \mathbb{R}$ given by $g(\mathbf{p}) = \int_a^b \mathbf{p}\,dx$ is a linear transformation, for any fixed $a, b \in \mathbb{R}$.

15. Let $\mathcal{V}$ be the vector space of all functions f from $\mathbb{R}$ to $\mathbb{R}$ that are infinitely differentiable (that is, for which $f^{(n)}$, the nth derivative of f, exists for every

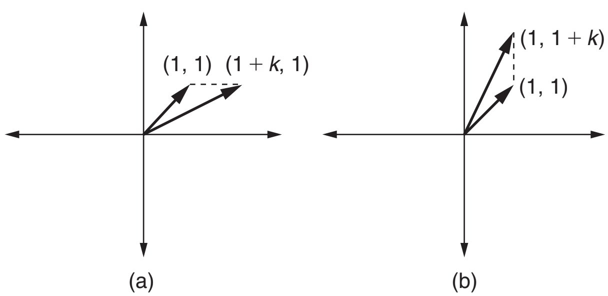

FIGURE 5.7

(a) Shear in the x-direction; (b) shear in the y-direction (both for $k > 0$)

integer $n \geq 1$). Use induction and Theorem 5.2 to show that for any given integer $k \geq 1$, $L: \mathcal{V} \to \mathcal{V}$ given by $L(f) = f^{(k)}$ is a linear operator.

16. Consider the function $f: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $f(\mathbf{A}) = \mathbf{BA}$, where $\mathbf{B}$ is some fixed $n \times n$ matrix. Show that f is a linear operator.

17. Let $\mathbf{B}$ be a fixed nonsingular matrix in $\mathcal{M}_{nn}$. Show that the mapping $f: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $f(\mathbf{A}) = \mathbf{B}^{-1}\mathbf{AB}$ is a linear operator.

18. Let a be a fixed real number.

(a) Let $L: \mathcal{P}_n \to \mathbb{R}$ be given by $L(\mathbf{p}(x)) = \mathbf{p}(a)$. (That is, L evaluates polynomials in $\mathcal{P}_n$ at $x = a$.) Show that L is a linear transformation.

(b) Let $L: \mathcal{P}_n \to \mathcal{P}_n$ be given by $L(\mathbf{p}(x)) = \mathbf{p}(x + a)$. (For example, when a is positive, L shifts the graph of $\mathbf{p}(x)$ to the *left* by a units.) Prove that L is a linear operator.

19. Let $\mathbf{A}$ be a fixed matrix in $\mathcal{M}_{nn}$. Define $f: \mathcal{P}_n \to \mathcal{M}_{nn}$ by

$$\begin{aligned} f(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) \\ = a_n\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{I}_n. \end{aligned}$$

Show that f is a linear transformation.

20. Let $\mathcal{V}$ be the unusual vector space from Example 7 in Section 4.1. Show that $L: \mathcal{V} \to \mathbb{R}$ given by $L(x) = \ln(x)$ is a linear transformation.

21. Let $\mathcal{V}$ be a vector space, and let $\mathbf{x} \neq \mathbf{0}$ be a fixed vector in $\mathcal{V}$. Prove that the translation function $f: \mathcal{V} \to \mathcal{V}$ given by $f(\mathbf{v}) = \mathbf{v} + \mathbf{x}$ is not a linear transformation.

22. Show that if $\mathbf{A}$ is a fixed matrix in $\mathcal{M}_{mn}$ and $\mathbf{y} \neq \mathbf{0}$ is a fixed vector in $\mathbb{R}^m$, then the mapping $f: \mathbb{R}^n \to \mathbb{R}^m$ given by $f(\mathbf{x}) = \mathbf{Ax} + \mathbf{y}$ is not a linear transformation by showing that part (1) of Theorem 5.1 fails for f.

23. Prove that $f: \mathcal{M}_{33} \to \mathbb{R}$ given by $f(\mathbf{A}) = |\mathbf{A}|$ is not a linear transformation. (A similar result is true for $\mathcal{M}_{nn}$, for $n > 1$.)

24. Suppose $L_1: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $L_2: \mathcal{V} \to \mathcal{W}$ is defined by $L_2(\mathbf{v}) = L_1(2\mathbf{v})$. Show that L_2 is a linear transformation.

25. Suppose $L: \mathbb{R}^3 \to \mathbb{R}^3$ is a linear operator and $L([1,0,0]) = [-2,1,0]$, $L([0,1,0]) = [3,-2,1]$, and $L([0,0,1]) = [0,-1,3]$. Find $L([-3,2,4])$. Give a formula for $L([x,y,z])$, for any $[x,y,z] \in \mathbb{R}^3$.

★26. Suppose $L: \mathbb{R}^2 \to \mathbb{R}^2$ is a linear operator and $L(\mathbf{i} + \mathbf{j}) = \mathbf{i} - 3\mathbf{j}$ and $L(-2\mathbf{i} + 3\mathbf{j}) = -4\mathbf{i} + 2\mathbf{j}$. Express $L(\mathbf{i})$ and $L(\mathbf{j})$ as linear combinations of $\mathbf{i}$ and $\mathbf{j}$.

27. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Show that $L(\mathbf{x} - \mathbf{y}) = L(\mathbf{x}) - L(\mathbf{y})$, for all vectors $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

28. Part (3) of Theorem 5.1 assures us that if $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then $L(a\mathbf{v}_1 + b\mathbf{v}_2) = aL(\mathbf{v}_1) + bL(\mathbf{v}_2)$, for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ and all $a, b \in \mathbb{R}$. Prove that the converse of this statement is true. (Hint: Consider two cases: first $a = b = 1$ and then $b = 0$.)

▶**29.** Finish the proof of part (3) of Theorem 5.1 by doing the Inductive Step.

30. **(a)** Suppose that $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation. Show that if $\{L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_n)\}$ is a linearly independent set of n distinct vectors in $\mathcal{W}$, for some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{V}$, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a linearly independent set in $\mathcal{V}$.

⋆**(b)** Find a counterexample to the converse of part (a).

▶**31.** Finish the proof of Theorem 5.3 by showing that if $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\mathcal{W}'$ is a subspace of $\mathcal{W}$ with pre-image $L^{-1}(\mathcal{W}')$, then $L^{-1}(\mathcal{W}')$ is a subspace of $\mathcal{V}$.

32. Show that every linear operator $L: \mathbb{R} \to \mathbb{R}$ has the form $L(\mathbf{x}) = c\mathbf{x}$, for some $c \in \mathbb{R}$.

33. Finish the proof of Theorem 5.2 by proving property (2) of a linear transformation for $L_2 \circ L_1$.

34. Let $L_1, L_2: \mathcal{V} \to \mathcal{W}$ be linear transformations. Define $(L_1 \oplus L_2): \mathcal{V} \to \mathcal{W}$ by $(L_1 \oplus L_2)(\mathbf{v}) = L_1(\mathbf{v}) + L_2(\mathbf{v})$ (where the latter addition takes place in $\mathcal{W}$). Also define $(c \odot L_1): \mathcal{V} \to \mathcal{W}$ by $(c \odot L_1)(\mathbf{v}) = c\,(L_1(\mathbf{v}))$ (where the latter scalar multiplication takes place in $\mathcal{W}$).

(a) Show that $(L_1 \oplus L_2)$ and $(c \odot L_1)$ are linear transformations.

(b) Use the results in part (a) above and part (b) of Exercise 2 to show that the set of all linear transformations from $\mathcal{V}$ to $\mathcal{W}$ is a vector space under the operations $\oplus$ and $\odot$.

35. Let $L: \mathbb{R}^2 \to \mathbb{R}^2$ be a nonzero linear operator. Show that L maps a line to either a line or a point.

⋆**36.** True or False:

(a) If $L: \mathcal{V} \to \mathcal{W}$ is a function between vector spaces for which $L(c\mathbf{v}) = cL(\mathbf{v})$, then L is a linear transformation.

(b) If $\mathcal{V}$ is an n-dimensional vector space with ordered basis B, then $L: \mathcal{V} \to \mathbb{R}^n$ given by $L(\mathbf{v}) = [\mathbf{v}]_B$ is a linear transformation.

(c) The function $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by $L([x, y, z]) = [x + 1, y - 2, z + 3]$ is a linear operator.

(d) If $\mathbf{A}$ is a 4×3 matrix, then $L(\mathbf{v}) = \mathbf{A}\mathbf{v}$ is a linear transformation from $\mathbb{R}^4$ to $\mathbb{R}^3$.

(e) A linear transformation from $\mathcal{V}$ to $\mathcal{W}$ always maps $\mathbf{0}_{\mathcal{V}}$ to $\mathbf{0}_{\mathcal{W}}$.

(f) If $M_1: \mathcal{V} \to \mathcal{W}$ and $M_2: \mathcal{W} \to \mathcal{X}$ are linear transformations, then $M_1 \circ M_2$ is a well-defined linear transformation.

(g) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then the image of any subspace of $\mathcal{V}$ is a subspace of $\mathcal{W}$.

(h) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then the pre-image of $\{\mathbf{0}_{\mathcal{W}}\}$ is a subspace of $\mathcal{V}$.

5.2 THE MATRIX OF A LINEAR TRANSFORMATION

In this section, we show that the behavior of any linear transformation $L: \mathcal{V} \to \mathcal{W}$ is determined by its effect on a basis for $\mathcal{V}$. In particular, when $\mathcal{V}$ and $\mathcal{W}$ are finite dimensional and ordered bases for $\mathcal{V}$ and $\mathcal{W}$ are chosen, we can obtain a matrix corresponding to L that is useful in computing images under L. Finally, we investigate how the matrix for L changes as the bases for $\mathcal{V}$ and $\mathcal{W}$ change.

A Linear Transformation Is Determined by Its Action on a Basis

If the action of a linear transformation $L: \mathcal{V} \to \mathcal{W}$ on a basis for $\mathcal{V}$ is known, then the action of L can be computed for all elements of $\mathcal{V}$, as we see in the next example.

Example 1

You can quickly verify that

$$B = \big([0,4,0,1], [-2,5,0,2], [-3,5,1,1], [-1,2,0,1]\big)$$

is an ordered basis for $\mathbb{R}^4$. Now suppose that $L: \mathbb{R}^4 \to \mathbb{R}^3$ is a linear transformation for which

$$L\big([0,4,0,1]\big) = [3,1,2], \qquad L([-2,5,0,2]) = [2,-1,1],$$

$$L([-3,5,1,1]) = [-4,3,0], \quad \text{and} \quad L([-1,2,0,1]) = [6,1,-1].$$

We can use the values of L on B to compute L for other vectors in $\mathbb{R}^4$. For example, let $\mathbf{v} = [-4,14,1,4]$. By using row reduction, we see that $[\mathbf{v}]_B = [2,-1,1,3]$ (verify!). So,

$$\begin{aligned} L(\mathbf{v}) &= L\big(2[0,4,0,1] - 1[-2,5,0,2] + 1[-3,5,1,1] + 3[-1,2,0,1]\big) \\ &= 2L\big([0,4,0,1]\big) - 1L([-2,5,0,2]) + 1L([-3,5,1,1]) \\ &\quad + 3L([-1,2,0,1]) \\ &= 2[3,1,2] - [2,-1,1] + [-4,3,0] + 3[6,1,-1] \\ &= [18,9,0]. \end{aligned}$$

In general, if $\mathbf{v} \in \mathbb{R}^4$ and $[\mathbf{v}]_B = [k_1, k_2, k_3, k_4]$, then

$$\begin{aligned} L(\mathbf{v}) &= k_1[3,1,2] + k_2[2,-1,1] + k_3[-4,3,0] + k_4[6,1,-1] \\ &= [3k_1 + 2k_2 - 4k_3 + 6k_4,\ k_1 - k_2 + 3k_3 + k_4,\ 2k_1 + k_2 - k_4]. \end{aligned}$$

Thus, we have derived a general formula for L from its effect on the basis B. ■

Example 1 illustrates the next theorem.

> **Theorem 5.4** Let $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ be an ordered basis for a vector space $\mathcal{V}$. Let $\mathcal{W}$ be a vector space, and let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ be any n vectors in $\mathcal{W}$. Then there is a unique linear transformation $L\colon \mathcal{V} \to \mathcal{W}$ such that $L(\mathbf{v}_1) = \mathbf{w}_1$, $L(\mathbf{v}_2) = \mathbf{w}_2, \ldots, L(\mathbf{v}_n) = \mathbf{w}_n$.

Proof. (Abridged) Let $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ be an ordered basis for $\mathcal{V}$, and let $\mathbf{v} \in \mathcal{V}$. Then $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, for some unique a_i's in $\mathbb{R}$. Let $\mathbf{w}_1, \ldots, \mathbf{w}_n$ be any vectors in $\mathcal{W}$. Define $L\colon \mathcal{V} \to \mathcal{W}$ by $L(\mathbf{v}) = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n$. Notice that $L(\mathbf{v})$ is well defined since the a_i's are unique.

To show that L is a linear transformation, we must prove that $L(\mathbf{x}_1 + \mathbf{x}_2) = L(\mathbf{x}_1) + L(\mathbf{x}_2)$ and $L(c\mathbf{x}_1) = cL(\mathbf{x}_1)$, for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{V}$ and all $c \in \mathbb{R}$. Suppose that $\mathbf{x}_1 = d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n$ and $\mathbf{x}_2 = e_1\mathbf{v}_1 + \cdots + e_n\mathbf{v}_n$. Then, by definition of L, $L(\mathbf{x}_1) = d_1\mathbf{w}_1 + \cdots + d_n\mathbf{w}_n$ and $L(\mathbf{x}_2) = e_1\mathbf{w}_n + \cdots + e_n\mathbf{w}_n$. However,

$$\mathbf{x}_1 + \mathbf{x}_2 = (d_1 + e_1)\mathbf{v}_1 + \cdots + (d_n + e_n)\mathbf{v}_n,$$

$$\text{so,} \quad L(\mathbf{x}_1 + \mathbf{x}_2) = (d_1 + e_1)\mathbf{w}_1 + \cdots + (d_n + e_n)\mathbf{w}_n,$$

again by definition of L. Hence, $L(\mathbf{x}_1) + L(\mathbf{x}_2) = L(\mathbf{x}_1 + \mathbf{x}_2)$.

Similarly, suppose $\mathbf{x} \in \mathcal{V}$, and $\mathbf{x} = t_1\mathbf{v}_1 + \cdots + t_n\mathbf{v}_n$. Then, $c\mathbf{x} = ct_1\mathbf{v}_1 + \cdots + ct_n\mathbf{v}_n$, and so $L(c\mathbf{x}) = ct_1\mathbf{w}_1 + \cdots + ct_n\mathbf{w}_n = cL(\mathbf{x})$. Hence, L is a linear transformation.

Finally, the proof of the uniqueness assertion is straightforward and is left as Exercise 25. □

The Matrix of a Linear Transformation

Our next goal is to show that every linear transformation on a finite dimensional vector space can be expressed as a matrix multiplication. This will allow us to solve problems involving linear transformations by performing matrix multiplications, which can easily be done by computer. As we will see, the matrix for a linear transformation is determined by the ordered bases B and C chosen for the domain and codomain, respectively. Our goal is to find a matrix that takes the B-coordinates of a vector in the domain to the C-coordinates of its image vector in the codomain.

Recall the linear transformation $L\colon \mathbb{R}^4 \to \mathbb{R}^3$ with the ordered basis B for $\mathbb{R}^4$ from Example 1. For $\mathbf{v} \in \mathbb{R}^4$, we let $[\mathbf{v}]_B = [k_1, k_2, k_3, k_4]$, and obtained the following formula for L:

$$L(\mathbf{v}) = [3k_1 + 2k_2 - 4k_3 + 6k_4,\ k_1 - k_2 + 3k_3 + k_4,\ 2k_1 + k_2 - k_4].$$

Now, to keep matters simple, we select the standard basis $C = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for the codomain $\mathbb{R}^3$, so that the C-coordinates of vectors in the codomain are the same as the vectors themselves. (That is, $L(\mathbf{v}) = [L(\mathbf{v})]_C$, since C is the standard basis.) Then this formula for L takes the B-coordinates of each vector in the domain to the C-coordinates of its image vector in the codomain. Now, notice that if

$$\mathbf{A}_{BC} = \begin{bmatrix} 3 & 2 & -4 & 6 \\ 1 & -1 & 3 & 1 \\ 2 & 1 & 0 & -1 \end{bmatrix}, \text{ then } \mathbf{A}_{BC}\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} = \begin{bmatrix} 3k_1 + 2k_2 - 4k_3 + 6k_4 \\ k_1 - k_2 + 3k_3 + k_4 \\ 2k_1 + k_2 - k_4 \end{bmatrix}.$$

Hence, the matrix $\mathbf{A}$ contains all of the information needed for carrying out the linear transformation L with respect to the chosen bases B and C.

A similar process can be used for any linear transformation between finite dimensional vector spaces.

Theorem 5.5 Let $\mathcal{V}$ and $\mathcal{W}$ be nontrivial vector spaces, with $\dim(\mathcal{V}) = n$ and $\dim(\mathcal{W}) = m$. Let $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ and $C = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m)$ be ordered bases for $\mathcal{V}$ and $\mathcal{W}$, respectively. Let $L\colon \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then there is a unique $m \times n$ matrix $\mathbf{A}_{BC}$ such that $\mathbf{A}_{BC}[\mathbf{v}]_B = [L(\mathbf{v})]_C$, for all $\mathbf{v} \in \mathcal{V}$. (That is, $\mathbf{A}_{BC}$ times the coordinatization of $\mathbf{v}$ with respect to B gives the coordinatization of $L(\mathbf{v})$ with respect to C.)

Furthermore, for $1 \le i \le n$, the ith column of $\mathbf{A}_{BC} = [L(\mathbf{v}_i)]_C$.

Theorem 5.5 asserts that once ordered bases for $\mathcal{V}$ and $\mathcal{W}$ have been selected, *each linear transformation $L\colon \mathcal{V} \to \mathcal{W}$ is equivalent to multiplication by a unique corresponding matrix*. The matrix $\mathbf{A}_{BC}$ in this theorem is known as the **matrix of the linear transformation L with respect to the ordered bases** B (for $\mathcal{V}$) **and** C (for $\mathcal{W}$). Theorem 5.5 also says that the matrix $\mathbf{A}_{BC}$ is computed as follows: find the image of each domain basis element $\mathbf{v}_i$ in turn, and then express these images in C-coordinates to get the respective columns of $\mathbf{A}_{BC}$.

The subscripts B and C on $\mathbf{A}$ are sometimes omitted when the bases being used are clear from context. Beware! If different ordered bases are chosen for $\mathcal{V}$ or $\mathcal{W}$, the matrix for the linear transformation will probably change.

Proof. Consider the $m \times n$ matrix $\mathbf{A}_{BC}$ whose ith column equals $[L(\mathbf{v}_i)]_C$, for $1 \le i \le n$. Let $\mathbf{v} \in \mathcal{V}$. We first prove that $\mathbf{A}_{BC}[\mathbf{v}]_B = [L(\mathbf{v})]_C$.

Suppose that $[\mathbf{v}]_B = [k_1, k_2, \ldots, k_n]$. Then $\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n$, and $L(\mathbf{v}) = k_1 L(\mathbf{v}_1) + k_2 L(\mathbf{v}_2) + \cdots + k_n L(\mathbf{v}_n)$, by Theorem 5.1. Hence,

$$\begin{aligned} [L(\mathbf{v})]_C &= [k_1 L(\mathbf{v}_1) + k_2 L(\mathbf{v}_2) + \cdots + k_n L(\mathbf{v}_n)]_C \\ &= k_1 [L(\mathbf{v}_1)]_C + k_2 [L(\mathbf{v}_2)]_C + \cdots + k_n [L(\mathbf{v}_n)]_C \quad \text{by Theorem 4.19} \end{aligned}$$

$$= k_1(\text{1st column of } \mathbf{A}_{BC}) + k_2(\text{2nd column of } \mathbf{A}_{BC})$$
$$+ \cdots + k_n(n\text{th column of } \mathbf{A}_{BC})$$
$$= \mathbf{A}_{BC}\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \mathbf{A}_{BC}[\mathbf{v}]_B.$$

To complete the proof, we need to establish the uniqueness of $\mathbf{A}_{BC}$. Suppose that $\mathbf{H}$ is an $m \times n$ matrix such that $\mathbf{H}[\mathbf{v}]_B = [L(\mathbf{v})]_C$ for all $\mathbf{v} \in \mathcal{V}$. We will show that $\mathbf{H} = \mathbf{A}_{BC}$. It is enough to show that the ith column of $\mathbf{H}$ equals the ith column of $\mathbf{A}_{BC}$, for $1 \leq i \leq n$. Consider the ith vector, $\mathbf{v}_i$, of the ordered basis B for $\mathcal{V}$. Since $[\mathbf{v}_i]_B = \mathbf{e}_i$, we have ith column of $\mathbf{H} = \mathbf{He}_i = \mathbf{H}[\mathbf{v}_i]_B = [L(\mathbf{v}_i)]_C$, and this is the ith column of $\mathbf{A}_{BC}$. □

Notice that in the special case where the codomain $\mathcal{W}$ is $\mathbb{R}^m$, and the basis C for $\mathcal{W}$ is the standard basis, Theorem 5.5 asserts that the ith column of $\mathbf{A}_{BC}$ is simply $L(\mathbf{v}_i)$ itself (why?).

Example 2

Table 5.1 lists the matrices corresponding to some geometric linear operators on $\mathbb{R}^3$, with respect to the standard basis. The columns of each matrix are quickly calculated using Theorem 5.5, since we simply find the images $L(\mathbf{e}_1), L(\mathbf{e}_2)$, and $L(\mathbf{e}_3)$ of the domain basis elements $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$. (Each image is equal to its coordinatization in the codomain since we are using the standard basis for the codomain as well.)

Once the matrix for each transformation is calculated, we can easily find the image of any vector using matrix multiplication. For example, to find the effect of the reflection L_1 in Table 5.1 on the vector $[3, -4, 2]$, we simply multiply by the matrix for L_1 to get

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

Example 3

We will find the matrix for the linear transformation $L: \mathcal{P}_3 \to \mathbb{R}^3$ given by

$$L(a_3x^3 + a_2x^2 + a_1x + a_0) = [a_0 + a_1, 2a_2, a_3 - a_0]$$

with respect to the standard ordered bases $B = (x^3, x^2, x, 1)$ for $\mathcal{P}_3$ and $C = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for $\mathbb{R}^3$. We first need to find $L(\mathbf{v})$, for each $\mathbf{v} \in B$. By definition of L, we have

$$L(x^3) = [0,0,1], \quad L(x^2) = [0,2,0], \quad L(x) = [1,0,0], \text{ and } \quad L(1) = [1,0,-1].$$

Table 5.1 Matrices for several geometric linear operators on $\mathbb{R}^3$

Transformation	Formula	Matrix
Reflection (through xy-plane)	$L_1\left(\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\right)=\begin{bmatrix}a_1\\a_2\\-a_3\end{bmatrix}$	$\begin{array}{ccc} L_1(\mathbf{e}_1) & L_1(\mathbf{e}_2) & L_1(\mathbf{e}_3) \end{array}$ $\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}$
Contraction or dilation	$L_2\left(\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\right)=\begin{bmatrix}ca_1\\ca_2\\ca_3\end{bmatrix}$, for $c\in\mathbb{R}$	$\begin{array}{ccc} L_2(\mathbf{e}_1) & L_2(\mathbf{e}_2) & L_2(\mathbf{e}_3) \end{array}$ $\begin{bmatrix}c&0&0\\0&c&0\\0&0&c\end{bmatrix}$
Projection (onto xy-plane)	$L_3\left(\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\right)=\begin{bmatrix}a_1\\a_2\\0\end{bmatrix}$	$\begin{array}{ccc} L_3(\mathbf{e}_1) & L_3(\mathbf{e}_2) & L_3(\mathbf{e}_3) \end{array}$ $\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$
Rotation (about z-axis through angle θ) (relative to the xy-plane)	$L_4\left(\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\right)=\begin{bmatrix}a_1\cos\theta-a_2\sin\theta\\a_1\sin\theta+a_2\cos\theta\\a_3\end{bmatrix}$	$\begin{array}{ccc} L_4(\mathbf{e}_1) & L_4(\mathbf{e}_2) & L_4(\mathbf{e}_3) \end{array}$ $\begin{bmatrix}\cos\theta&-\sin\theta&0\\\sin\theta&\cos\theta&0\\0&0&1\end{bmatrix}$
Shear (in the z-direction with factor k) (analog of Exercise 11 in Section 5.1)	$L_5\left(\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\right)=\begin{bmatrix}a_1+ka_3\\a_2+ka_3\\a_3\end{bmatrix}$	$\begin{array}{ccc} L_5(\mathbf{e}_1) & L_5(\mathbf{e}_2) & L_5(\mathbf{e}_3) \end{array}$ $\begin{bmatrix}1&0&k\\0&1&k\\0&0&1\end{bmatrix}$

Since we are using the standard basis C for $\mathbb{R}^3$, each of these images in $\mathbb{R}^3$ is its own C-coordinatization. Then by Theorem 5.5, the matrix $\mathbf{A}_{BC}$ for L is the matrix whose columns are these images; that is,

$$\mathbf{A}_{BC}=\begin{array}{c}\begin{array}{cccc} L(x^3) & L(x^2) & L(x) & L(1)\end{array}\\ \begin{bmatrix}0&0&1&1\\0&2&0&0\\1&0&0&-1\end{bmatrix}\end{array}.$$

We will compute $L(5x^3-x^2+3x+2)$ using this matrix. Now, $[5x^3-x^2+3x+2]_B=[5,-1,3,2]$. Hence, multiplication by $\mathbf{A}_{BC}$ gives

$$\left[L(5x^3-x^2+3x+2)\right]_C=\begin{bmatrix}0&0&1&1\\0&2&0&0\\1&0&0&-1\end{bmatrix}\begin{bmatrix}5\\-1\\3\\2\end{bmatrix}=\begin{bmatrix}5\\-2\\3\end{bmatrix}.$$

Since C is the standard basis for $\mathbb{R}^3$, we have $L(5x^3-x^2+3x+2)=[5,-2,3]$, which can be quickly verified to be the correct answer. ∎

Example 4

We will find the matrix for the same linear transformation $L: \mathcal{P}_3 \to \mathbb{R}^3$ of Example 3 with respect to the different ordered bases

$$D = (x^3 + x^2,\ x^2 + x,\ x + 1,\ 1)$$
$$\text{and} \quad E = ([-2, 1, -3], [1, -3, 0], [3, -6, 2]).$$

You should verify that D and E are bases for $\mathcal{P}_3$ and $\mathbb{R}^3$, respectively.

We first need to find $L(\mathbf{v})$, for each $\mathbf{v} \in D$. By definition of L, we have $L(x^3 + x^2) = [0, 2, 1]$, $L(x^2 + x) = [1, 2, 0]$, $L(x + 1) = [2, 0, -1]$, and $L(1) = [1, 0, -1]$. Now we must find the coordinatization of each of these images in terms of the basis E for $\mathbb{R}^3$. Since we must solve for the coordinates of many vectors, it is quicker to use the transition matrix $\mathbf{Q}$ from the standard basis C for $\mathbb{R}^3$ to the basis E. From Theorem 4.22, $\mathbf{Q}$ is the inverse of the matrix whose columns are the vectors in E; that is,

$$\mathbf{Q} = \begin{bmatrix} -2 & 1 & 3 \\ 1 & -3 & -6 \\ -3 & 0 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix}.$$

Now, multiplying $\mathbf{Q}$ by each of the images, we get

$$\left[L(x^3 + x^2)\right]_E = \mathbf{Q}\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \qquad \left[L(x^2 + x)\right]_E = \mathbf{Q}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -10 \\ 26 \\ -15 \end{bmatrix},$$

$$[L(x + 1)]_E = \mathbf{Q}\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -15 \\ 41 \\ -23 \end{bmatrix}, \quad \text{and} \quad [L(1)]_E = \mathbf{Q}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -9 \\ 25 \\ -14 \end{bmatrix}.$$

By Theorem 5.5, the matrix $\mathbf{A}_{DE}$ for L is the matrix whose columns are these products.

$$\mathbf{A}_{DE} = \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix}$$

We will compute $L(5x^3 - x^2 + 3x + 2)$ using this matrix. We must first find the representation for $5x^3 - x^2 + 3x + 2$ in terms of the basis D. Solving $5x^3 - x^2 + 3x + 2 = a(x^3 + x^2) + b(x^2 + x) + c(x + 1) + d(1)$ for $a, b, c,$ and d, we get the unique solution $a = 5$, $b = -6$, $c = 9$, and $d = -7$ (verify!). Hence, $\left[5x^3 - x^2 + 3x + 2\right]_D = [5, -6, 9, -7]$. Then

$$\left[L(5x^3 - x^2 + 3x + 2)\right]_E = \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix}\begin{bmatrix} 5 \\ -6 \\ 9 \\ -7 \end{bmatrix} = \begin{bmatrix} -17 \\ 43 \\ -24 \end{bmatrix}.$$

This answer represents a coordinate vector in terms of the basis E, and so

$$L(5x^3 - x^2 + 3x + 2) = -17\begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix} + 43\begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix} - 24\begin{bmatrix} 3 \\ -6 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix},$$

which agrees with the answer in Example 3. ■

Finding the New Matrix for a Linear Transformation after a Change of Basis

The next theorem indicates precisely how the matrix for a linear transformation changes when we alter the bases for the domain and codomain.

> **Theorem 5.6** Let $\mathcal{V}$ and $\mathcal{W}$ be two nontrivial finite dimensional vector spaces with ordered bases B and C, respectively. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation with matrix $\mathbf{A}_{BC}$ with respect to bases B and C. Suppose that D and E are other ordered bases for $\mathcal{V}$ and $\mathcal{W}$, respectively. Let $\mathbf{P}$ be the transition matrix from B to D, and let $\mathbf{Q}$ be the transition matrix from C to E. Then the matrix $\mathbf{A}_{DE}$ for L with respect to bases D and E is given by $\mathbf{A}_{DE} = \mathbf{Q}\mathbf{A}_{BC}\mathbf{P}^{-1}$.

The situation in Theorem 5.6 is summarized in Figure 5.8.

Proof. For all $\mathbf{v} \in \mathcal{V}$,

$$\begin{aligned}
& \mathbf{A}_{BC}[\mathbf{v}]_B = [L(\mathbf{v})]_C && \text{by Theorem 5.5} \\
\Rightarrow\ & \mathbf{Q}\mathbf{A}_{BC}[\mathbf{v}]_B = \mathbf{Q}[L(\mathbf{v})]_C && \\
\Rightarrow\ & \mathbf{Q}\mathbf{A}_{BC}[\mathbf{v}]_B = [L(\mathbf{v})]_E && \text{because } \mathbf{Q} \text{ is the transition matrix from } C \text{ to } E \\
\Rightarrow\ & \mathbf{Q}\mathbf{A}_{BC}\mathbf{P}^{-1}[\mathbf{v}]_D = [L(\mathbf{v})]_E. && \text{because } \mathbf{P}^{-1} \text{ is the transition matrix from } D \text{ to } B
\end{aligned}$$

However, $\mathbf{A}_{DE}$ is the *unique* matrix such that $\mathbf{A}_{DE}[\mathbf{v}]_D = [L(\mathbf{v})]_E$, for all $\mathbf{v} \in \mathcal{V}$. Hence, $\mathbf{A}_{DE} = \mathbf{Q}\mathbf{A}_{BC}\mathbf{P}^{-1}$. □

$$\begin{tikzcd}[column sep=huge, row sep=huge]
{[\mathbf{v}]_B} \arrow[r, "\mathbf{A}_{BC}", "\text{(Matrix for } L \text{ using } B,\ C)"'] \arrow[d, "\text{Transition matrix } \mathbf{P}"'] & {[L(\mathbf{v})]_C} \arrow[d, "\text{Transition matrix } \mathbf{Q}"] \\
{[\mathbf{v}]_D} \arrow[r, "\mathbf{A}_{DE}", "\text{(Matrix for } L \text{ using } D,\ E)"'] & {[L(\mathbf{v})]_E}
\end{tikzcd}$$

FIGURE 5.8

Relationship between matrices $\mathbf{A}_{BC}$ and $\mathbf{A}_{DE}$ for a linear transformation under a change of basis

Theorem 5.6 gives us an alternate method for finding the matrix of a linear transformation with respect to one pair of bases when the matrix for another pair of bases is known.

Example 5

Recall the linear transformation $L\colon \mathcal{P}_3 \to \mathbb{R}^3$ from Examples 3 and 4, given by

$$L(a_3x^3 + a_2x^2 + a_1x + a_0) = [a_0 + a_1, 2a_2, a_3 - a_0].$$

Example 3 shows that the matrix for L using the standard bases B (for $\mathcal{P}_3$) and C (for $\mathbb{R}^3$) is

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

Also, in Example 4, we computed directly to find the matrix $\mathbf{A}_{DE}$ for the ordered bases $D = (x^3 + x^2, x^2 + x, x + 1, 1)$ for $\mathcal{P}_3$ and $E = ([-2, 1, -3], [1, -3, 0], [3, -6, 2])$ for $\mathbb{R}^3$. Instead, we now use Theorem 5.6 to calculate $\mathbf{A}_{DE}$. Recall from Example 4 that the transition matrix $\mathbf{Q}$ from bases C to E is

$$\mathbf{Q} = \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix}.$$

Also, the transition matrix $\mathbf{P}^{-1}$ from bases D to B is

$$\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \qquad \text{(Verify!)}$$

Hence,

$$\mathbf{A}_{DE} = \mathbf{Q}\mathbf{A}_{BC}\mathbf{P}^{-1} = \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix},$$

which agrees with the result obtained for $\mathbf{A}_{DE}$ in Example 4.

Linear Operators and Similarity

Suppose L is a linear operator on a finite dimensional vector space $\mathcal{V}$. If B is a basis for $\mathcal{V}$, then there is some matrix $\mathbf{A}_{BB}$ for L with respect to B. Also, if C is another basis for $\mathcal{V}$, then there is some matrix $\mathbf{A}_{CC}$ for L with respect to C. Let $\mathbf{P}$ be the transition matrix from B to C (see Figure 5.9). Notice that by Theorem 5.6 we have $\mathbf{A}_{BB} = \mathbf{P}^{-1}\mathbf{A}_{CC}\mathbf{P}$, and so, by the definition of similar matrices, $\mathbf{A}_{BB}$ and $\mathbf{A}_{CC}$ are similar. This argument shows that any two matrices for the same linear operator with respect to different bases are similar. In fact, the converse is also true (see Exercise 20).

Example 6

Consider the linear operator $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ whose matrix with respect to the standard basis B for $\mathbb{R}^3$ is

$$\mathbf{A}_{BB} = \frac{1}{7}\begin{bmatrix} 6 & 3 & -2 \\ 3 & -2 & 6 \\ -2 & 6 & 3 \end{bmatrix}.$$

We will use eigenvectors to find another basis D for $\mathbb{R}^3$ so that with respect to D, L has a much simpler matrix representation. Now, $p_{\mathbf{A}_{BB}}(x) = |x\mathbf{I}_3 - \mathbf{A}_{BB}| = x^3 - x^2 - x + 1 = (x-1)^2(x+1)$ (verify!).

By row reducing $(1\mathbf{I}_3 - \mathbf{A}_{BB})$ and $(-1\mathbf{I}_3 - \mathbf{A}_{BB})$ we find the basis $\{[3,1,0],[-2,0,1]\}$ for the eigenspace E_1 for $\mathbf{A}_{BB}$ and the basis $\{[1,-3,2]\}$ for the eigenspace E_{-1} for $\mathbf{A}_{BB}$. (Again, verify!) A quick check verifies that $D = \{[3,1,0],[-2,0,1],[1,-3,2]\}$ is a basis for $\mathbb{R}^3$ consisting of eigenvectors for $\mathbf{A}_{BB}$.

Next, recall that $\mathbf{A}_{DD}$ is similar to $\mathbf{A}_{BB}$. In particular, from the remarks right before this example, $\mathbf{A}_{DD} = \mathbf{P}^{-1}\mathbf{A}_{BB}\mathbf{P}$, where $\mathbf{P}$ is the transition matrix from D to B. Now, the matrix whose columns are the vectors in D is the transition matrix from D to the standard basis B. Thus,

$$\mathbf{P} = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 2 \end{bmatrix}, \quad \text{with} \quad \mathbf{P}^{-1} = \frac{1}{14}\begin{bmatrix} 3 & 5 & 6 \\ -2 & 6 & 10 \\ 1 & -3 & 2 \end{bmatrix}.$$

$$\begin{tikzcd}[column sep=huge, row sep=huge]
{[\mathbf{v}]_B} \arrow[r, "\mathbf{A}_{BB}", "\text{(Matrix for } L \text{ using } B)"'] & {[L(\mathbf{v})]_B} \arrow[d, "\text{Transition matrix } \mathbf{P}"] \\
{[\mathbf{v}]_C} \arrow[u, "\text{Transition matrix } \mathbf{P}^{-1}"] \arrow[r, "\mathbf{A}_{CC}", "\text{(Matrix for } L \text{ using } C)"'] & {[L(\mathbf{v})]_C}
\end{tikzcd}$$

FIGURE 5.9

Relationship between matrices $\mathbf{A}_{BB}$ and $\mathbf{A}_{CC}$ for a linear operator under a change of basis as the transition matrix from B to D. Then,

$$\mathbf{A}_{DD} = \mathbf{P}^{-1}\mathbf{A}_{BB}\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

a diagonal matrix with the eigenvalues 1 and -1 on the main diagonal.

Written in this form, the operator L is more comprehensible. Compare $\mathbf{A}_{DD}$ to the matrix for a reflection through the xy-plane given in Table 5.1. Now, because D is not the standard basis for $\mathbb{R}^3$, L is *not* a reflection through the xy-plane. But we can show that L is a reflection of all vectors in $\mathbb{R}^3$ through the plane formed by the two basis vectors for E_1 (that is, the plane is the eigenspace E_1 itself). By the uniqueness assertion in Theorem 5.4, it is enough to show that L acts as a reflection through the plane E_1 for each of the three basis vectors of D.

Since $[3,1,0]$ and $[-2,0,1]$ are in the plane E_1, we need to show that L "reflects" these vectors to themselves. But this is true since $L([3,1,0]) = 1[3,1,0] = [3,1,0]$, and similarly for $[-2,0,1]$. Finally, notice that $[1,-3,2]$ is orthogonal to the plane E_1 (since it is orthogonal to both $[3,1,0]$ and $[-2,0,1]$). Therefore, we need to show that L "reflects" this vector to its opposite. But, $L([1,-3,2]) = -1[1,-3,2] = -[1,-3,2]$, and we are done. Hence, L is a reflection through the plane E_1. ■

Because the matrix $\mathbf{A}_{DD}$ in Example 6 is diagonal, it is easy to see that $p_{\mathbf{A}_{DD}}(x) = (x-1)^2(x+1)$. In Exercise 6 of Section 3.4, you were asked to prove that similar matrices have the same characteristic polynomial. Therefore, $p_{\mathbf{A}_{BB}}(x)$ also equals $(x-1)^2(x+1)$.

Matrix for the Composition of Linear Transformations

Our final theorem for this section shows how to find the corresponding matrix for the composition of linear transformations. The proof is left as Exercise 15.

Theorem 5.7 Let $\mathcal{V}_1, \mathcal{V}_2$, and $\mathcal{V}_3$ be nontrivial finite dimensional vector spaces with ordered bases B, C, and D, respectively. Let $L_1\colon \mathcal{V}_1 \to \mathcal{V}_2$ be a linear transformation with matrix $\mathbf{A}_{BC}$ with respect to bases B and C, and let $L_2\colon \mathcal{V}_2 \to \mathcal{V}_3$ be a linear transformation with matrix $\mathbf{A}_{CD}$ with respect to bases C and D. Then the matrix $\mathbf{A}_{BD}$ for the composite linear transformation $L_2 \circ L_1\colon \mathcal{V}_1 \to \mathcal{V}_3$ with respect to bases B and D is the product $\mathbf{A}_{CD}\mathbf{A}_{BC}$.

Theorem 5.7 can be generalized to compositions of several linear transformations, as in the next example.

Example 7

Let $L_1, L_2, \ldots, L_5$ be the geometric linear operators on $\mathbb{R}^3$ given in Table 5.1. Let $\mathbf{A}_1, \ldots, \mathbf{A}_5$ be the matrices for these operators using the standard basis for $\mathbb{R}^3$. Then, the matrix for the composition $L_4 \circ L_5$ is

$$\mathbf{A}_4\mathbf{A}_5 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & k\cos\theta - k\sin\theta \\ \sin\theta & \cos\theta & k\sin\theta + k\cos\theta \\ 0 & 0 & 1 \end{bmatrix}.$$

Similarly, the matrix for the composition $L_2 \circ L_3 \circ L_1 \circ L_5$ is

$$\mathbf{A}_2\mathbf{A}_3\mathbf{A}_1\mathbf{A}_5 = \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c & 0 & kc \\ 0 & c & kc \\ 0 & 0 & 0 \end{bmatrix}.$$

■

- ♦ **Supplemental Material:** You have now covered the prerequisites for Section 7.3, "Complex Vector Spaces."
- ♦ **Application:** You have now covered the prerequisites for Section 8.8, "Computer Graphics."

New Vocabulary

matrix for a linear transformation

Highlights

- ■ A linear transformation between finite dimensional vector spaces is uniquely determined once the images of an ordered basis for the domain are specified. (More specifically, let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, with $\dim(\mathcal{V}) = n$. Let $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ be an ordered basis for $\mathcal{V}$, and let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ be any n (not necessarily distinct) vectors in $\mathcal{W}$. Then there is a unique linear transformation $L: \mathcal{V} \to \mathcal{W}$ such that $L(\mathbf{v}_i) = \mathbf{w}_i$, for $1 \le i \le n$.)
- ■ Every linear transformation between (nontrivial) finite dimensional vector spaces has a unique matrix $\mathbf{A}_{BC}$ with respect to the ordered bases B and C chosen for the domain and codomain, respectively. (More specifically, let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation, with $\dim(\mathcal{V}) = n$, $\dim(\mathcal{W}) = m$. Let $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ and $C = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m)$ be ordered bases for $\mathcal{V}$ and $\mathcal{W}$, respectively. Then there is a unique $m \times n$ matrix $\mathbf{A}_{BC}$ such that $\mathbf{A}_{BC}[\mathbf{v}]_B = [L(\mathbf{v})]_C$, for all $\mathbf{v} \in \mathcal{V}$.)
- ■ If $\mathbf{A}_{BC}$ is the matrix for a linear transformation with respect to the ordered bases B and C chosen for the domain and codomain, respectively, then the ith column of $\mathbf{A}_{BC}$ is the C-coordinatization of the image of the ith vector in B. That is, the ith column of $\mathbf{A}_{BC}$ equals $[L(\mathbf{v}_i)]_C$.
- ■ After a change of bases for the domain and codomain, the new matrix for a given linear transformation can be found using the original matrix and the transition

matrices between bases. (More specifically, let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation between (nontrivial) finite dimensional vector spaces with ordered bases B and C, respectively, and with matrix $\mathbf{A}_{BC}$ in terms of bases B and C. If D and E are other ordered bases for $\mathcal{V}$ and $\mathcal{W}$, respectively, and $\mathbf{P}$ is the transition matrix from B to D, and $\mathbf{Q}$ is the transition matrix from C to E, then the matrix $\mathbf{A}_{DE}$ for L in terms of bases D and E is $\mathbf{A}_{DE} = \mathbf{Q}\mathbf{A}_{BC}\mathbf{P}^{-1}$.)

- Matrices for several useful geometric operators on $\mathbb{R}^3$ are given in Table 5.1.
- The matrix for a linear operator (on a finite dimensional vector space) after a change of basis is similar to the original matrix.
- The matrix for the composition of linear transformations (using the same ordered bases) is the product of the matrices for the individual linear transformations in reverse order. (More specifically, if $L_1: \mathcal{V}_1 \to \mathcal{V}_2$ is a linear transformation with matrix $\mathbf{A}_{BC}$ with respect to ordered bases B and C, and $L_2: \mathcal{V}_2 \to \mathcal{V}_3$ is a linear transformation with matrix $\mathbf{A}_{CD}$ with respect to bases C and D, then the matrix $\mathbf{A}_{BD}$ for $L_2 \circ L_1: \mathcal{V}_1 \to \mathcal{V}_3$ with respect to bases B and D is given by $\mathbf{A}_{BD} = \mathbf{A}_{CD}\mathbf{A}_{BC}$.)

EXERCISES FOR SECTION 5.2

1. Verify that the correct matrix is given for each of the geometric linear operators in Table 5.1.

2. For each of the following linear transformations $L: \mathcal{V} \to \mathcal{W}$, find the matrix for L with respect to the standard bases for $\mathcal{V}$ and $\mathcal{W}$.

★**(a)** $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by $L([x,y,z]) = [-6x + 4y - z, -2x + 3y - 5z, 3x - y + 7z]$

(b) $L: \mathbb{R}^4 \to \mathbb{R}^2$ given by $L([x,y,z,w]) = [3x - 5y + z - 2w, 5x + y - 2z + 8w]$

★**(c)** $L: \mathcal{P}_3 \to \mathbb{R}^3$ given by $L(ax^3 + bx^2 + cx + d) = [4a - b + 3c + 3d, a + 3b - c + 5d, -2a - 7b + 5c - d]$

(d) $L: \mathcal{P}_3 \to \mathcal{M}_{22}$ given by

$$L(ax^3 + bx^2 + cx + d) = \begin{bmatrix} -3a - 2c & -b + 4d \\ 4b - c + 3d & -6a - b + 2d \end{bmatrix}$$

3. For each of the following linear transformations $L: \mathcal{V} \to \mathcal{W}$, find the matrix $\mathbf{A}_{BC}$ for L with respect to the given bases B for $\mathcal{V}$ and C for $\mathcal{W}$ using the method of Theorem 5.5:

★**(a)** $L: \mathbb{R}^3 \to \mathbb{R}^2$ given by $L([x,y,z]) = [-2x + 3z, x + 2y - z]$ with $B = ([1,-3,2],[-4,13,-3],[2,-3,20])$ and $C = ([-2,-1],[5,3])$

(b) $L: \mathbb{R}^2 \to \mathbb{R}^3$ given by $L([x,y]) = [13x - 9y, \ -x - 2y, -11x + 6y]$ with $B = ([2,3],[-3,-4])$ and $C = ([-1,2,2],[-4,1,3],[1,-1,-1])$

★(c) $L: \mathbb{R}^2 \to \mathcal{P}_2$ given by $L([a,b]) = (-a+5b)x^2 + (3a-b)x + 2b$ with $B = ([5,3],[3,2])$ and $C = (3x^2 - 2x, -2x^2 + 2x - 1, x^2 - x + 1)$

(d) $L: \mathcal{M}_{22} \to \mathbb{R}^3$ given by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = [a - c + 2d, 2a + b - d, -2c + d]$

with $B = \left(\begin{bmatrix} 2 & 5 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} -2 & -2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -3 & -4 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} -1 & -3 \\ 0 & 1 \end{bmatrix}\right)$ and

$C = ([7,0,-3],[2,-1,-2],[-2,0,1])$

★(e) $L: \mathcal{P}_2 \to \mathcal{M}_{23}$ given by

$$L(ax^2 + bx + c) = \begin{bmatrix} -a & 2b + c & 3a - c \\ a + b & c & -2a + b - c \end{bmatrix}$$

with $B = (-5x^2 - x - 1, -6x^2 + 3x + 1, 2x + 1)$ and $C =$
$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$

4. In each case, find the matrix $\mathbf{A}_{DE}$ for the given linear transformation $L: \mathcal{V} \to \mathcal{W}$ with respect to the given bases D and E by first finding the matrix for L with respect to the standard bases B and C for $\mathcal{V}$ and $\mathcal{W}$, respectively, and then using the method of Theorem 5.6.

★(a) $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by $L([a,b,c]) = [-2a + b, -b - c, \ a + 3c]$ with $D = ([15,-6,4],[2,0,1],[3,-1,1])$ and $E = ([1,-3,1],[0,3,-1],[2,-2,1])$

★(b) $L: \mathcal{M}_{22} \to \mathbb{R}^2$ given by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = [6a - b + 3c - 2d, -2a + 3b - c + 4d]$$

with

$$D = \left(\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) \quad \text{and}$$

$E = ([-2,5],[-1,2])$

(c) $L: \mathcal{M}_{22} \to \mathcal{P}_2$ given by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (b - c)x^2 + (3a - d)x + (4a - 2c + d)$$

with

$$D = \left(\begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 0 & 1 \end{bmatrix}\right) \quad \text{and}$$

$$E = (2x - 1, -5x^2 + 3x - 1, x^2 - 2x + 1)$$

5. Verify that the same matrix is obtained for L in Exercise 3(d) by first finding the matrix for L with respect to the standard bases and then using the method of Theorem 5.6.

6. In each case, find the matrix $\mathbf{A}_{BB}$ for each of the given linear operators $L: \mathcal{V} \to \mathcal{V}$ with respect to the given basis B by using the method of Theorem 5.5. Then, check your answer by calculating the matrix for L using the standard basis and applying the method of Theorem 5.6.

★**(a)** L: $\mathbb{R}^2 \to \mathbb{R}^2$ given by $L([x,y]) = [2x - y,\ x - 3y]$ with $B = ([4,-1], [-7,2])$

★**(b)** L: $\mathcal{P}_2 \to \mathcal{P}_2$ given by $L(ax^2 + bx + c) = (b - 2c)x^2 + (2a + c)x + (a - b - c)$ with $B = (2x^2 + 2x - 1, x, -3x^2 - 2x + 1)$

(c) L: $\mathcal{M}_{22} \to \mathcal{M}_{22}$ given by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a - c + d & a - b \\ -3b - 2d & -a - 2c + 3d \end{bmatrix}$$

with

$$B = \left(\begin{bmatrix} -2 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\right)$$

7. ★**(a)** Let L: $\mathcal{P}_3 \to \mathcal{P}_2$ be given by $L(\mathbf{p}) = \mathbf{p}'$, for $\mathbf{p} \in \mathcal{P}_3$. Find the matrix for L with respect to the standard bases for $\mathcal{P}_3$ and $\mathcal{P}_2$. Use this matrix to calculate $L(4x^3 - 5x^2 + 6x - 7)$ by matrix multiplication.

(b) Let L: $\mathcal{P}_2 \to \mathcal{P}_3$ be the indefinite integral linear transformation; that is, $L(\mathbf{p})$ is the vector $\int \mathbf{p}(x)\, dx$ with zero constant term. Find the matrix for L with respect to the standard bases for $\mathcal{P}_2$ and $\mathcal{P}_3$. Use this matrix to calculate $L(2x^2 - x + 5)$ by matrix multiplication.

8. Let $L: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear operator that performs a counterclockwise rotation through an angle of $\frac{\pi}{6}$ radians (30°).

★**(a)** Find the matrix for L with respect to the standard basis for $\mathbb{R}^2$.

(b) Find the matrix for L with respect to the basis $B = ([4,-3],[3,-2])$.

9. Let $L: \mathcal{M}_{23} \to \mathcal{M}_{32}$ be given by $L(\mathbf{A}) = \mathbf{A}^T$.

(a) Find the matrix for L with respect to the standard bases.

★**(b)** Find the matrix for L with respect to the bases
$$B = \left(\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix}, \begin{bmatrix}0&1&-1\\0&0&0\end{bmatrix}, \begin{bmatrix}0&1&0\\0&0&0\end{bmatrix}, \begin{bmatrix}0&0&0\\-1&0&0\end{bmatrix}, \begin{bmatrix}0&0&0\\0&-1&-1\end{bmatrix},\right.$$
$$\left.\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}\right)$$ for $\mathcal{M}_{23}$, and
$$C = \left(\begin{bmatrix}1&1\\0&0\\0&0\end{bmatrix}, \begin{bmatrix}1&-1\\0&0\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&-1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\0&0\\1&1\end{bmatrix}, \begin{bmatrix}0&0\\0&0\\1&-1\end{bmatrix}\right)$$
for $\mathcal{M}_{32}$.

★**10.** Let B be a basis for $\mathcal{V}_1$, C be a basis for $\mathcal{V}_2$, and D be a basis for $\mathcal{V}_3$. Suppose $L_1: \mathcal{V}_1 \to \mathcal{V}_2$ and $L_2: \mathcal{V}_2 \to \mathcal{V}_3$ are represented, respectively, by the matrices

$$\mathbf{A}_{BC} = \begin{bmatrix}-2&3&-1\\4&0&-2\end{bmatrix} \quad \text{and} \quad \mathbf{A}_{CD} = \begin{bmatrix}4&-1\\2&0\\-1&-3\end{bmatrix}.$$

Find the matrix $\mathbf{A}_{BD}$ representing the composition $L_2 \circ L_1: \mathcal{V}_1 \to \mathcal{V}_3$.

11. Let $L_1: \mathbb{R}^3 \to \mathbb{R}^4$ be given by $L_1([x,y,z]) = [x-y-z, 2y+3z, x+3y, -2x+z]$, and let $L_2: \mathbb{R}^4 \to \mathbb{R}^2$ be given by $L_2([x,y,z,w]) = [2y-2z+3w, x-z+w]$.

(a) Find the matrices for L_1 and L_2 with respect to the standard bases in each case.

(b) Find the matrix for $L_2 \circ L_1$ with respect to the standard bases for $\mathbb{R}^3$ and $\mathbb{R}^2$ using Theorem 5.7.

(c) Check your answer to part (b) by computing $(L_2 \circ L_1)([x,y,z])$ and finding the matrix for $L_2 \circ L_1$ directly from this result.

12. Let $\mathbf{A} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$, the matrix representing the counterclockwise rotation of $\mathbb{R}^2$ about the origin through an angle θ.

(a) Use Theorem 5.7 to show that

$$\mathbf{A}^2 = \begin{bmatrix}\cos 2\theta & -\sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{bmatrix}.$$

(b) Generalize the result of part (a) to show that for any integer $n \geq 1$,

$$\mathbf{A}^n = \begin{bmatrix}\cos n\theta & -\sin n\theta\\ \sin n\theta & \cos n\theta\end{bmatrix}.$$

13. Let $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ be an ordered basis for a vector space $\mathcal{V}$. Find the matrix with respect to B for each of the following linear operators $L: \mathcal{V} \to \mathcal{V}$:

★**(a)** $L(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v} \in \mathcal{V}$ (identity linear operator)

(b) $L(\mathbf{v}) = \mathbf{0}$, for all $\mathbf{v} \in \mathcal{V}$ (zero linear operator)

★(c) $L(\mathbf{v}) = c\mathbf{v}$, for all $\mathbf{v} \in \mathcal{V}$, and for some fixed $c \in \mathbb{R}$ (scalar linear operator)

(d) $L: \mathcal{V} \to \mathcal{V}$ given by $L(\mathbf{v}_1) = \mathbf{v}_2$, $L(\mathbf{v}_2) = \mathbf{v}_3, \ldots, L(\mathbf{v}_{n-1}) = \mathbf{v}_n$, $L(\mathbf{v}_n) = \mathbf{v}_1$ (forward replacement of basis vectors)

★(e) $L: \mathcal{V} \to \mathcal{V}$ given by $L(\mathbf{v}_1) = \mathbf{v}_n$, $L(\mathbf{v}_2) = \mathbf{v}_1, \ldots, L(\mathbf{v}_{n-1}) = \mathbf{v}_{n-2}$, $L(\mathbf{v}_n) = \mathbf{v}_{n-1}$ (reverse replacement of basis vectors)

14. Let $L: \mathbb{R}^n \to \mathbb{R}$ be a linear transformation. Prove that there is a vector $\mathbf{x}$ in $\mathbb{R}^n$ such that $L(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{y} \in \mathbb{R}^n$.

►15. Prove Theorem 5.7.

16. Let L: $\mathbb{R}^3 \to \mathbb{R}^3$ be given by $L([x,y,z]) = [-4y - 13z, -6x + 5y + 6z, 2x - 2y - 3z]$.

(a) What is the matrix for L with respect to the standard basis for $\mathbb{R}^3$?

(b) What is the matrix for L with respect to the basis

$$B = ([-1,-6,2],[3,4,-1],[-1,-3,1])?$$

(c) What does your answer to part (b) tell you about the vectors in B? Explain.

17. In Example 6, verify that $p_{\mathbf{A}_{BB}}(x) = (x-1)^2(x+1)$, $\{[3,1,0],[-2,0,1]\}$ is a basis for the eigenspace E_1, $\{[1,-3,2]\}$ is a basis for the eigenspace E_{-1}, the transition matrices $\mathbf{P}$ and $\mathbf{P}^{-1}$ are as indicated, and, finally, $\mathbf{A}_{DD} = \mathbf{P}^{-1}\mathbf{A}_{BB}\mathbf{P}$ is a diagonal matrix with entries $1, 1$, and -1, respectively, on the main diagonal.

18. Let $L: \mathbb{R}^3 \to \mathbb{R}^3$ be the linear operator whose matrix with respect to the standard basis B for $\mathbb{R}^3$ is

$$\mathbf{A}_{BB} = \frac{1}{9}\begin{bmatrix} 8 & 2 & 2 \\ 2 & 5 & -4 \\ 2 & -4 & 5 \end{bmatrix}.$$

★(a) Calculate and factor $p_{\mathbf{A}_{BB}}(x)$. (Be sure to incorporate $\frac{1}{9}$ correctly into your calculations.)

★(b) Solve for a basis for each eigenspace for L. Combine these to form a basis C for $\mathbb{R}^3$.

★(c) Find the transition matrix $\mathbf{P}$ from C to B.

(d) Calculate $\mathbf{A}_{CC}$ using $\mathbf{A}_{BB}, \mathbf{P}$, and $\mathbf{P}^{-1}$.

(e) Use $\mathbf{A}_{CC}$ to give a geometric description of the operator L, as was done in Example 6.

19. Let L be a linear operator on a vector space $\mathcal{V}$ with ordered basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Suppose that k is a nonzero real number, and let C be the ordered basis $(k\mathbf{v}_1, \ldots, k\mathbf{v}_n)$ for $\mathcal{V}$. Show that $\mathbf{A}_{BB} = \mathbf{A}_{CC}$.

20. Let $\mathcal{V}$ be an n-dimensional vector space, and let $\mathbf{X}$ and $\mathbf{Y}$ be similar $n \times n$ matrices. Prove that there is a linear operator $L: \mathcal{V} \to \mathcal{V}$ and bases B and C such that $\mathbf{X}$ is the matrix for L with respect to B and $\mathbf{Y}$ is the matrix for L with respect to C. (Hint: Suppose that $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}\mathbf{P}$. Choose any basis B for $\mathcal{V}$. Then create the linear operator $L: \mathcal{V} \to \mathcal{V}$ whose matrix with respect to B is $\mathbf{X}$. Let $\mathbf{v}_i$ be the vector so that $[\mathbf{v}_i]_B = i$th column of $\mathbf{P}$. Define C to be $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Prove that C is a basis for $\mathcal{V}$. Then show that $\mathbf{P}^{-1}$ is the transition matrix from B to C and that $\mathbf{Y}$ is the matrix for L with respect to C.)

21. Let $B = ([a,b],[c,d])$ be a basis for $\mathbb{R}^2$. Then $ad - bc \neq 0$ (why?). Let $L: \mathbb{R}^2 \to \mathbb{R}^2$ be a linear operator such that $L([a,b]) = [c,d]$ and $L([c,d]) = [a,b]$. Show that the matrix for L with respect to the standard basis for $\mathbb{R}^2$ is

$$\frac{1}{ad - bc}\begin{bmatrix} cd - ab & a^2 - c^2 \\ d^2 - b^2 & ab - cd \end{bmatrix}.$$

22. Let $L: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation where $L(\mathbf{v})$ is the reflection of $\mathbf{v}$ through the line $y = mx$. (Assume that the initial point of $\mathbf{v}$ is the origin.) Show that the matrix for L with respect to the standard basis for $\mathbb{R}^2$ is

$$\frac{1}{1 + m^2}\begin{bmatrix} 1 - m^2 & 2m \\ 2m & m^2 - 1 \end{bmatrix}.$$

(Hint: Use Exercise 19 in Section 1.2.)

23. Find the set of all matrices with respect to the standard basis for $\mathbb{R}^2$ for all linear operators that

(a) Take all vectors of the form $[0,y]$ to vectors of the form $[0,y']$

(b) Take all vectors of the form $[x,0]$ to vectors of the form $[x',0]$

(c) Satisfy both parts (a) and (b) simultaneously

24. Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces, and let $\mathcal{Y}$ be a subspace of $\mathcal{V}$. Suppose that $L: \mathcal{Y} \to \mathcal{W}$ is a linear transformation. Prove that there is a linear transformation $L': \mathcal{V} \to \mathcal{W}$ such that $L'(\mathbf{y}) = L(\mathbf{y})$ for every $\mathbf{y} \in \mathcal{Y}$. ($L'$ is called an **extension** of L to $\mathcal{V}$.)

►**25.** Prove the uniqueness assertion in Theorem 5.4. (Hint: Let $\mathbf{v}$ be any vector in $\mathcal{V}$. Show that there is only one possible answer for $L(\mathbf{v})$ by expressing $L(\mathbf{v})$ as a linear combination of the $\mathbf{w}_i$'s.)

★**26.** True or False:

(a) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, and $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ is an ordered basis for $\mathcal{V}$, then for any $\mathbf{v} \in \mathcal{V}$, $L(\mathbf{v})$ can be computed if $L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_n)$ are known.

(b) There is a unique linear transformation $L: \mathbb{R}^3 \to \mathcal{P}_3$ such that $L([1,0,0]) = x^3 - x^2, L([0,1,0]) = x^3 - x^2$, and $L([0,0,1]) = x^3 - x^2$.

(c) If $\mathcal{V}, \mathcal{W}$ are nontrivial finite dimensional vector spaces and $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then there is a unique matrix $\mathbf{A}$ corresponding to L.

(d) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and B is a (finite nonempty) ordered basis for $\mathcal{V}$, and C is a (finite nonempty) ordered basis for $\mathcal{W}$, then $[\mathbf{v}]_B = \mathbf{A}_{BC}[L(\mathbf{v})]_C$.

(e) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ is an ordered basis for $\mathcal{V}$, and C is a (finite nonempty) ordered basis for $\mathcal{W}$, then the ith column of $\mathbf{A}_{BC}$ is $[L(\mathbf{v}_i)]_C$.

(f) The matrix for the projection of $\mathbb{R}^3$ onto the xz-plane (with respect to the standard basis) is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

(g) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, and B and D are (finite nonempty) ordered bases for $\mathcal{V}$, and C and E are (finite nonempty) ordered bases for $\mathcal{W}$, then $\mathbf{A}_{DE}\mathbf{P} = \mathbf{Q}\mathbf{A}_{BC}$, where $\mathbf{P}$ is the transition matrix from B to D, and $\mathbf{Q}$ is the transition matrix from C to E.

(h) If $L: \mathcal{V} \to \mathcal{V}$ is a linear operator on a nontrivial finite dimensional vector space, and B and D are ordered bases for $\mathcal{V}$, then $\mathbf{A}_{BB}$ is similar to $\mathbf{A}_{DD}$.

(i) Similar square matrices have identical characteristic polynomials.

(j) If $L_1, L_2: \mathbb{R}^2 \to \mathbb{R}^2$ are linear transformations with matrices $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, respectively, with respect to the standard basis, then the matrix for $L_2 \circ L_1$ with respect to the standard basis equals $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

5.3 THE DIMENSION THEOREM

In this section, we introduce two special subspaces associated with a linear transformation $L: \mathcal{V} \to \mathcal{W}$: the kernel of L (a subspace of $\mathcal{V}$) and the range of L (a subspace of $\mathcal{W}$). We illustrate techniques for calculating bases for both the kernel and range and show their dimensions are related to the rank of any matrix for the linear transformation. We then use this to show that any matrix and its transpose have the same rank.

Kernel and Range

Definition Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. The **kernel** of L, denoted by $\ker(L)$, is the subset of all vectors in $\mathcal{V}$ that map to $\mathbf{0}_{\mathcal{W}}$. That is, $\ker(L) = \{\mathbf{v} \in \mathcal{V} \mid L(\mathbf{v}) = \mathbf{0}_{\mathcal{W}}\}$. The **range** of L, or, $\text{range}(L)$, is the subset of all vectors in $\mathcal{W}$ that are the image of some vector in $\mathcal{V}$. That is, $\text{range}(L) = \{L(\mathbf{v}) \mid \mathbf{v} \in \mathcal{V}\}$.

Remember that the kernel[1] is a subset of the *domain* and that the range is a subset of the *codomain*. Since the kernel of $L: \mathcal{V} \to \mathcal{W}$ is the pre-image of the subspace $\{\mathbf{0}_{\mathcal{W}}\}$ of $\mathcal{W}$, it must be a subspace of $\mathcal{V}$ by Theorem 5.3. That theorem also assures us that the range of L is a subspace of $\mathcal{W}$. Hence, we have

Theorem 5.8 If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then the kernel of L is a subspace of $\mathcal{V}$ and the range of L is a subspace of $\mathcal{W}$.

Example 1

Projection: For $n \geq 3$, consider the linear operator $L: \mathbb{R}^n \to \mathbb{R}^n$ given by $L([a_1, a_2, \ldots, a_n]) = [a_1, a_2, 0, \ldots, 0]$. Now, $\ker(L)$ consists of those elements of the domain that map to $[0, 0, \ldots, 0]$, the zero vector of the codomain. Hence, for vectors in the kernel, $a_1 = a_2 = 0$, but $a_3, \ldots, a_n$ can have any values. Thus,

$$\ker(L) = \left\{ [0, 0, a_3, \ldots, a_n] \middle| a_3, \ldots, a_n \in \mathbb{R} \right\}.$$

Notice that $\ker(L)$ is a subspace of the domain and that $\dim(\ker(L)) = n - 2$, because the standard basis vectors $\mathbf{e}_3, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$ span $\ker(L)$.

Also, $\text{range}(L)$ consists of those elements of the codomain $\mathbb{P}^2$ that are images of domain elements. Hence, $\text{range}(L) = \{[a_1, a_2, 0, \ldots, 0] | a_1, a_2 \in \mathbb{R}\}$. Notice that $\text{range}(L)$ is a subspace of the codomain and that $\dim(\text{range}(L)) = 2$, since the standard basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ span $\text{range}(L)$.

Example 2

Differentiation: Consider the linear transformation $L: \mathcal{P}_3 \to \mathcal{P}_2$ given by $L(ax^3 + bx^2 + cx + d) = 3ax^2 + 2bx + c$. Now, $\ker(L)$ consists of the polynomials in $\mathcal{P}_3$ that map to the zero polynomial in $\mathcal{P}_2$. However, if $3ax^2 + 2bx + c = 0$, we must have $a = b = c = 0$. Hence, $\ker(L) = \left\{ 0x^3 + 0x^2 + 0x + d \middle| d \in \mathbb{R} \right\}$; that is, $\ker(L)$ is just the subset of $\mathcal{P}_3$ of all constant polynomials. Notice that $\ker(L)$ is a subspace of $\mathcal{P}_3$ and that $\dim(\ker(L)) = 1$ because the single polynomial "1" spans $\ker(L)$.

[1] Some textbooks refer to the kernel of L as the **nullspace** of L.

Also, range(L) consists of all polynomials in the codomain $\mathcal{P}_2$ of the form $3ax^2 + 2bx + c$. Since every polynomial $Ax^2 + Bx + C$ of degree 2 or less can be expressed in this form (take $a = A/3,\ b = B/2,\ c = C$), range($L$) is all of $\mathcal{P}_2$. Therefore, range(L) is a subspace of $\mathcal{P}_2$, and $\dim(\text{range}(L)) = 3$.

Example 3

Rotation: Recall that the linear transformation $L: \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

for some (fixed) angle θ, represents the counterclockwise rotation of any vector $[x,y]$ with initial point at the origin through the angle θ.

Now, $\ker(L)$ consists of all vectors in the domain $\mathbb{R}^2$ that map to $[0,0]$ in the codomain $\mathbb{R}^2$. However, only $[0,0]$ itself is rotated by L to the zero vector. Hence, $\ker(L) = \{[0,0]\}$. Notice that $\ker(L)$ is a subspace of $\mathbb{R}^2$, and $\dim(\ker(L)) = 0$.

Also, range(L) is all of the codomain $\mathbb{R}^2$ because every nonzero vector $\mathbf{v}$ in $\mathbb{R}^2$ is the image of the vector of the same length at the angle θ *clockwise* from $\mathbf{v}$. Thus, range(L) $= \mathbb{R}^2$, and so, range(L) is a subspace of $\mathbb{R}^2$ with $\dim(\text{range}(L)) = 2$.

Finding the Kernel from the Matrix of a Linear Transformation

Consider the linear transformation $L: \mathbb{R}^n \to \mathbb{R}^m$ given by $L(\mathbf{X}) = \mathbf{AX}$, where $\mathbf{A}$ is a (fixed) $m \times n$ matrix and $\mathbf{X} \in \mathbb{R}^n$. Now, $\ker(L)$ is the subspace of all vectors $\mathbf{X}$ in the domain $\mathbb{R}^n$ that are solutions of the homogeneous system $\mathbf{AX} = \mathbf{O}$. If $\mathbf{B}$ is the reduced row echelon form matrix for $\mathbf{A}$, we find a basis for $\ker(L)$ by solving for particular solutions to the system $\mathbf{BX} = \mathbf{O}$ by systematically setting each independent variable equal to 1 in turn, while setting the others equal to 0. (You should be familiar with this process from the Diagonalization Method for finding fundamental eigenvectors in Section 3.4.) Thus, $\dim(\ker(L))$ equals the number of independent variables in the system $\mathbf{BX} = \mathbf{O}$.

We present an example of this technique.

Example 4

Let $L: \mathbb{R}^5 \longrightarrow \mathbb{R}^4$ be given by $L(\mathbf{X}) = \mathbf{AX}$, where

$$\mathbf{A} = \begin{bmatrix} 8 & 4 & 16 & 32 & 0 \\ 4 & 2 & 10 & 22 & -4 \\ -2 & -1 & -5 & -11 & 7 \\ 6 & 3 & 15 & 33 & -7 \end{bmatrix}.$$

To solve for $\ker(L)$, we first row reduce $\mathbf{A}$ to

$$\mathbf{B} = \begin{bmatrix} 1 & \frac{1}{2} & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The homogeneous system $\mathbf{BX} = \mathbf{O}$ has independent variables x_2 and x_4, and

$$\begin{cases} x_1 = -\frac{1}{2}x_2 + 2x_4 \\ x_3 = \quad - 3x_4 \\ x_5 = \quad 0 \end{cases}.$$

We construct two particular solutions, first by setting $x_2 = 1$ and $x_4 = 0$ to obtain $\mathbf{v}_1 = [-\frac{1}{2}, 1, 0, 0, 0]$, and then setting $x_2 = 0$ and $x_4 = 1$, yielding $\mathbf{v}_2 = [2, 0, -3, 1, 0]$. The set $\{\mathbf{v}_1, \mathbf{v}_2\}$ forms a basis for $\ker(L)$, and thus, $\dim(\ker(L)) = 2$, the number of independent variables. The entire subspace $\ker(L)$ consists of all linear combinations of the basis vectors; that is,

$$\ker(L) = \{a\mathbf{v}_1 + b\mathbf{v}_2 \,|\, a, b \in \mathbb{R}\} = \left\{ \left[-\frac{1}{2}a + 2b, a, -3b, b, 0 \right] \,\middle|\, a, b \in \mathbb{R} \right\}.$$

Finally, note that we could have eliminated fractions in this basis, just as we did with fundamental eigenvectors in Section 3.4, by replacing $\mathbf{v}_1$ with $2\mathbf{v}_1 = [-1, 2, 0, 0, 0]$. ■

Example 4 illustrates the following general technique:

Method for Finding a Basis for the Kernel of a Linear Transformation (Kernel Method)
Let $L: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation given by $\mathbf{L}(\mathbf{X}) = \mathbf{AX}$ for some $m \times n$ matrix $\mathbf{A}$. To find a basis for $\ker(L)$, perform the following steps:

Step 1: Find $\mathbf{B}$, the reduced row echelon form of $\mathbf{A}$.

Step 2: Solve for one particular solution for each independent variable in the homogeneous system $\mathbf{BX} = \mathbf{O}$. The ith such solution, $\mathbf{v}_i$, is found by setting the ith independent variable equal to 1 and setting all other independent variables equal to 0.

Step 3: The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a basis for $\ker(L)$. (We can replace any $\mathbf{v}_i$ with $c\mathbf{v}_i$, where $c \neq 0$, to eliminate fractions.)

The method for finding a basis for $\ker(L)$ is practically identical to Step 3 of the Diagonalization Method of Section 3.4, in which we create a basis of fundamental eigenvectors for the eigenspace E_λ for a matrix $\mathbf{A}$. This is to be expected, since E_λ is really the kernel of the linear transformation L whose matrix is $(\lambda\mathbf{I}_n - \mathbf{A})$.

Finding the Range from the Matrix of a Linear Transformation

Next, we determine a method for finding a basis for the range of $L: \mathbb{R}^n \to \mathbb{R}^m$ given by $L(\mathbf{X}) = \mathbf{AX}$. In Section 1.5, we saw that $\mathbf{AX}$ can be expressed as a linear combination of the columns of $\mathbf{A}$. In particular, if $\mathbf{X} = [x_1, \ldots x_n]$, then $\mathbf{AX} = x_1$ (1st column of $\mathbf{A}$) $+ \cdots + x_n$ (nth column of $\mathbf{A}$). Thus, range(L) is spanned by the set of columns of $\mathbf{A}$; that is, range$(L) = $ span({columns of $\mathbf{A}$}). Note that $L(\mathbf{e}_i)$ equals the ith column of $\mathbf{A}$. Thus, we can also say that $\{L(\mathbf{e}_1), \ldots, L(\mathbf{e}_n)\}$ spans range(L).

The fact that the columns of $\mathbf{A}$ span range(L) combined with the Independence Test Method yields the following general technique for finding a basis for the range:

Method for Finding a Basis for the Range of a Linear Transformation (Range Method)

Let $L: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation given by $\mathbf{L}(\mathbf{X}) = \mathbf{AX}$, for some $m \times n$ matrix $\mathbf{A}$. To find a basis for $\mathbf{range}(L)$, perform the following steps:

Step 1: Find $\mathbf{B}$, the reduced row echelon form of $\mathbf{A}$.

Step 2: Form the set of those columns of $\mathbf{A}$ whose corresponding columns in $\mathbf{B}$ have nonzero pivots. This set is a basis for $\mathbf{range}(L)$.

Example 5

Consider the linear transformation $L: \mathbb{R}^5 \to \mathbb{R}^4$ given in Example 4. After row reducing the matrix $\mathbf{A}$ for L, we obtained a matrix $\mathbf{B}$ in reduced row echelon form having nonzero pivots in columns $1, 3$, and 5. Hence, columns 1, 3, and 5 of $\mathbf{A}$ form a basis for $\mathbf{range}(L)$. In particular, we get the basis $\{[8,4,-2,6],\ [16,10,-5,15],\ [0,-4,7,-7]\}$, and so $\dim(\mathbf{range}(L)) = 3$. ■

From Examples 4 and 5, we see that $\dim(\ker(L)) + \dim(\text{range}(L)) = 2 + 3 = 5 = \dim(\mathbb{R}^5) = \dim(\text{domain}(L))$, for the given linear transformation L. We can understand why this works by examining our methods for calculating bases for the kernel and range. For $\ker(L)$, we get one basis vector for each independent variable, which corresponds to a nonpivot column of $\mathbf{A}$ after row reducing. For range(L), we get one basis vector for each pivot column of $\mathbf{A}$. Together, these account for the total number of columns of $\mathbf{A}$, which is the dimension of the domain.

The fact that the number of nonzero pivots of $\mathbf{A}$ equals the number of nonzero rows in the reduced row echelon form matrix for $\mathbf{A}$ shows that $\dim(\text{range}(L)) = \text{rank}\,(\mathbf{A})$. This result is stated in the following theorem, which also holds when bases other than the standard bases are used (see Exercise 17).

Theorem 5.9 If $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with matrix $\mathbf{A}$ with respect to any bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, then

(1) $\dim(\text{range}(L)) = \text{rank}(\mathbf{A})$

(2) $\dim(\ker(L)) = n - \text{rank}(\mathbf{A})$

(3) $\dim(\ker(L)) + \dim(\text{range}(L)) = \dim(\text{domain}(L)) = n$.

The Dimension Theorem

The result in part (3) of Theorem 5.9 generalizes to linear transformations between any vector spaces $\mathcal{V}$ and $\mathcal{W}$, as long as the dimension of the domain is finite. We state this important theorem here, but postpone its proof until after a discussion of isomorphism in Section 5.5. An alternate proof of the Dimension Theorem that does not involve the matrix of the linear transformation is outlined in Exercise 18 of this section.

Theorem 5.10 (Dimension Theorem) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\mathcal{V}$ is finite dimensional, then $\text{range}(L)$ is finite dimensional, and

$$\dim(\ker(L)) + \dim(\text{range}(L)) = \dim(\mathcal{V}).$$

We have already seen that for the linear transformation in Examples 4 and 5, the dimensions of the kernel and the range add up to the dimension of the domain, as the Dimension Theorem asserts. Notice the Dimension Theorem holds for the linear transformations in Examples 1 through 3 as well.

Example 6

Consider $L: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T$. Now, $\ker(L) = \{\mathbf{A} \in \mathcal{M}_{nn} \,|\, \mathbf{A} + \mathbf{A}^T = \mathbf{O}_n\}$. However, $\mathbf{A} + \mathbf{A}^T = \mathbf{O}_n$ implies that $\mathbf{A} = -\mathbf{A}^T$. Hence, $\ker(L)$ is precisely the set of all skew-symmetric $n \times n$ matrices.

The range of L is the set of all matrices $\mathbf{B}$ of the form $\mathbf{A} + \mathbf{A}^T$ for some $n \times n$ matrix $\mathbf{A}$. However, if $\mathbf{B} = \mathbf{A} + \mathbf{A}^T$, then $\mathbf{B}^T = \left(\mathbf{A} + \mathbf{A}^T\right)^T = \mathbf{A}^T + \mathbf{A} = \mathbf{B}$, so $\mathbf{B}$ is symmetric. Thus, $\text{range}(L) \subseteq \{\text{symmetric } n \times n \text{ matrices}\}$.

Next, if $\mathbf{B}$ is a symmetric $n \times n$ matrix, then $L(\frac{1}{2}\mathbf{B}) = \frac{1}{2}L(\mathbf{B}) = \frac{1}{2}(\mathbf{B} + \mathbf{B}^T) = \frac{1}{2}(\mathbf{B} + \mathbf{B}) = \mathbf{B}$, and so $\mathbf{B} \in \text{range}(L)$, thus proving $\{\text{symmetric } n \times n \text{ matrices}\} \subseteq \text{range}(L)$. Hence, $\text{range}(L)$ is the set of all symmetric $n \times n$ matrices.

In Exercise 12 of Section 4.6, we found that $\dim(\{\text{skew-symmetric } n \times n \text{ matrices}\}) = (n^2 - n)/2$ and that $\dim(\{\text{symmetric } n \times n \text{ matrices}\}) = \left(n^2 + n\right)/2$. Notice that the Dimension Theorem holds here, since $\dim(\ker(L)) + \dim(\text{range}(L)) = (n^2 - n)/2 + \left(n^2 + n\right)/2 = n^2 = \dim(\mathcal{M}_{nn})$. ■

Rank of the Transpose

We can use the Range Method to prove the following result:[2]

Corollary 5.11 If $\mathbf{A}$ is any matrix, then $\mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{A}^T)$.

Proof. Let $\mathbf{A}$ be an $m \times n$ matrix. Consider the linear transformation $L: \mathbb{R}^n \to \mathbb{R}^m$ with associated matrix $\mathbf{A}$ (using the standard bases). By the Range Method, $\mathrm{range}(L)$ is the span of the column vectors of $\mathbf{A}$. Hence, $\mathrm{range}(L)$ is the span of the row vectors of $\mathbf{A}^T$; that is, $\mathrm{range}(L)$ is the row space of $\mathbf{A}^T$. Thus, $\dim(\mathrm{range}(L)) = \mathrm{rank}(\mathbf{A}^T)$, by the Simplified Span Method. But by Theorem 5.9, $\dim(\mathrm{range}(L)) = \mathrm{rank}(\mathbf{A})$. Hence, $\mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{A}^T)$. □

Example 7

Let $\mathbf{A}$ be the matrix from Examples 4 and 5. We calculated its reduced row echelon form $\mathbf{B}$ in Example 4 and found it has three nonzero rows. Hence, $\mathrm{rank}(\mathbf{A}) = 3$. Now,

$$\mathbf{A}^T = \begin{bmatrix} 8 & 4 & -2 & 6 \\ 4 & 2 & -1 & 3 \\ 16 & 10 & -5 & 15 \\ 32 & 22 & -11 & 33 \\ 0 & -4 & 7 & -7 \end{bmatrix} \quad \text{row reduces to} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{7}{5} \\ 0 & 0 & 1 & -\frac{1}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

showing that $\mathrm{rank}(\mathbf{A}^T) = 3$ as well. ■

In some textbooks, rank(**A**) is called the **row rank** of **A** and rank($\mathbf{A}^T$) is called the **column rank** of **A**. Thus, Corollary 5.11 asserts that the row rank of **A** equals the column rank of **A**.

Recall that rank(**A**) = dim(row space of **A**). Analogous to the concept of row space, we define the **column space** of a matrix **A** as the span of the columns of **A**. In Corollary 5.11, we observed that if $L: \mathbb{R}^n \to \mathbb{R}^m$ with $L(\mathbf{X}) = \mathbf{AX}$ (using the standard bases), then range(L) = span({columns of **A**}) = column space of **A**, and so dim(range(L)) = dim(column space of **A**) = rank($\mathbf{A}^T$). With this new terminology, Corollary 5.11 asserts that dim(row space of **A**) = dim(column space of **A**). Be careful! This statement does not imply that these *spaces* are equal, only that their *dimensions* are equal. In fact, unless **A** is square, they contain vectors of different sizes. Notice that for the matrix **A** in Example 7, the row space of **A** is a subspace of $\mathbb{R}^5$, but the column space of **A** is a subspace of $\mathbb{R}^4$.

[2] In Exercise 18 of Section 4.6, you were asked to prove Corollary 5.11 by essentially the same method given here, only using different notation.

New Vocabulary

column rank (of a matrix)
column space (of a matrix)
Dimension Theorem
kernel (of a linear transformation)
Kernel Method
range (of a linear transformation)
Range Method
row rank (of a matrix)

Highlights

- The kernel of a linear transformation consists of all vectors of the domain that map to the zero vector of the codomain. The kernel is always a subspace of the domain.
- The range of a linear transformation consists of all vectors of the codomain that are images of vectors in the domain. The range is always a subspace of the codomain.
- If $\mathbf{A}$ is the matrix (with respect to any bases) for a linear transformation $L\colon \mathbb{R}^n \to \mathbb{R}^m$, then $\dim(\ker(L)) = n - \operatorname{rank}(\mathbf{A})$ and $\dim(\operatorname{range}(L)) = \operatorname{rank}(\mathbf{A})$.
- Kernel Method: A basis for the kernel of a linear transformation $L(\mathbf{X}) = \mathbf{AX}$ is obtained from the solution set of $\mathbf{BX} = \mathbf{O}$ by letting each independent variable in turn equal 1 and all other independent variables equal 0, where $\mathbf{B}$ is the reduced row echelon form of $\mathbf{A}$.
- Range Method: A basis for the range of a linear transformation $L(\mathbf{X}) = \mathbf{AX}$ is obtained by selecting the columns of $\mathbf{A}$ corresponding to pivot columns in the reduced row echelon form matrix $\mathbf{B}$ for $\mathbf{A}$.
- Dimension Theorem: If $L\colon \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\mathcal{V}$ is finite dimensional, then $\dim(\ker(L)) + \dim(\operatorname{range}(L)) = \dim(\mathcal{V})$.
- The rank of any matrix (= row rank) is equal to the rank of its transpose (= column rank).

EXERCISES FOR SECTION 5.3

1. Let $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 5 & 1 & -1 \\ -3 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

$\star$**(a)** Is $[1, -2, 3]$ in $\ker(L)$? Why or why not?

(b) Is $[2, -1, 4]$ in $\ker(L)$? Why or why not?

$\star$**(c)** Is $[2, -1, 4]$ in $\operatorname{range}(L)$? Why or why not?

(d) Is $[-16, 12, -8]$ in $\operatorname{range}(L)$? Why or why not?

2. Let $L\colon \mathcal{P}_3 \to \mathcal{P}_3$ be given by $L(ax^3+bx^2+cx+d) = 2cx^3+(a+b)x+(d+c)$.

★**(a)** Is x^3-5x^2+3x-6 in $\ker(L)$? Why or why not?

(b) Is $4x^3-4x^2$ in $\ker(L)$? Why or why not?

★**(c)** Is $8x^3-x-1$ in $\operatorname{range}(L)$? Why or why not?

(d) Is $4x^3-3x^2+7$ in $\operatorname{range}(L)$? Why or why not?

3. For each of the following linear transformations $L\colon \mathcal{V} \to \mathcal{W}$, find a basis for $\ker(L)$ and a basis for $\operatorname{range}(L)$. Verify that $\dim(\ker(L)) + \dim(\operatorname{range}(L)) = \dim(\mathcal{V})$.

★**(a)** $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & 5 \\ -2 & 3 & -13 \\ 3 & -3 & 15 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(b) $L\colon \mathbb{R}^3 \to \mathbb{R}^4$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 4 & -2 & 8 \\ 7 & 1 & 5 \\ -2 & -1 & 0 \\ 3 & -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c) $L\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3 & 2 & 11 \\ 2 & 1 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

★**(d)** $L\colon \mathbb{R}^4 \to \mathbb{R}^5$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} -14 & -8 & -10 & 2 \\ -4 & -1 & 1 & -2 \\ -6 & 2 & 12 & -10 \\ 3 & -7 & -24 & 17 \\ 4 & 2 & 2 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

4. For each of the following linear transformations $L\colon \mathcal{V} \to \mathcal{W}$, find a basis for $\ker(L)$ and a basis for $\operatorname{range}(L)$, and verify that $\dim(\ker(L)) + \dim(\operatorname{range}(L)) = \dim(\mathcal{V})$:

★**(a)** $L\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by $L([x_1,x_2,x_3]) = [0,x_2]$

(b) $L\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by $L([x_1,x_2]) = [x_1, x_1+x_2, x_2]$

(c) $L: \mathcal{M}_{22} \to \mathcal{M}_{32}$ given by $L\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{bmatrix} 0 & -a_{12} \\ -a_{21} & 0 \\ 0 & 0 \end{bmatrix}$

★(d) $L: \mathcal{P}_4 \to \mathcal{P}_2$ given by $L(ax^4 + bx^3 + cx^2 + dx + e) = cx^2 + dx + e$

(e) $L: \mathcal{P}_2 \to \mathcal{P}_3$ given by $L(ax^2 + bx + c) = cx^3 + bx^2 + ax$

★(f) $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by $L([x_1, x_2, x_3]) = [x_1, 0, x_1 - x_2 + x_3]$

★(g) $L: \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L(\mathbf{A}) = \mathbf{A}^T$

(h) $L: \mathcal{M}_{33} \to \mathcal{M}_{33}$ given by $L(\mathbf{A}) = \mathbf{A} - \mathbf{A}^T$

★(i) $L: \mathcal{P}_2 \to \mathbb{R}^2$ given by $L(\mathbf{p}) = [\mathbf{p}(1), \mathbf{p}'(1)]$

(j) $L: \mathcal{P}_4 \to \mathbb{R}^3$ given by $L(\mathbf{p}) = [\mathbf{p}(-1), \mathbf{p}(0), \mathbf{p}(1)]$

5. (a) Suppose that $L: \mathcal{V} \to \mathcal{W}$ is the linear transformation given by $L(\mathbf{v}) = \mathbf{0}_{\mathcal{W}}$, for all $\mathbf{v} \in \mathcal{V}$. What is $\ker(L)$? What is $\text{range}(L)$?

(b) Suppose that $L: \mathcal{V} \to \mathcal{V}$ is the linear transformation given by $L(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v} \in \mathcal{V}$. What is $\ker(L)$? What is $\text{range}(L)$?

★6. Consider the mapping $L: \mathcal{M}_{33} \to \mathbb{R}$ given by $L(\mathbf{A}) = \text{trace}(\mathbf{A})$ (see Exercise 14 in Section 1.4). Show that L is a linear transformation. What is $\ker(L)$? What is $\text{range}(L)$? Calculate $\dim(\ker(L))$ and $\dim(\text{range}(L))$.

7. Let $\mathcal{V}$ be a vector space with fixed basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Define $L: \mathcal{V} \to \mathcal{V}$ by $L(\mathbf{v}_1) = \mathbf{v}_2$, $L(\mathbf{v}_2) = \mathbf{v}_3, \ldots,$ $L(\mathbf{v}_{n-1}) = \mathbf{v}_n$, $L(\mathbf{v}_n) = \mathbf{v}_1$. Find $\text{range}(L)$. What is $\ker(L)$?

★8. Consider $L: \mathcal{P}_2 \to \mathcal{P}_4$ given by $L(\mathbf{p}) = x^2\mathbf{p}$. What is $\ker(L)$? What is $\text{range}(L)$? Verify that $\dim(\ker(L)) + \dim(\text{range}(L)) = \dim(\mathcal{P}_2)$.

9. Consider $L: \mathcal{P}_4 \to \mathcal{P}_2$ given by $L(\mathbf{p}) = \mathbf{p}''$. What is $\ker(L)$? What is $\text{range}(L)$? Verify that $\dim(\ker(L)) + \dim(\text{range}(L)) = \dim(\mathcal{P}_4)$.

★10. Consider $L: \mathcal{P}_n \to \mathcal{P}_n$ given by $L(\mathbf{p}) = \mathbf{p}^{(k)}$ (the kth derivative of $\mathbf{p}$), where $k \le n$. What is $\dim(\ker(L))$? What is $\dim(\text{range}(L))$? What happens when $k > n$?

11. Let a be a fixed real number. Consider $L: \mathcal{P}_n \to \mathbb{R}$ given by $L(\mathbf{p}(x)) = \mathbf{p}(a)$ (that is, the evaluation of $\mathbf{p}$ at $x = a$). (Recall from Exercise 18 in Section 5.1 that L is a linear transformation.) Show that $\{x - a, x^2 - a^2, \ldots, x^n - a^n\}$ is a basis for $\ker(L)$. (Hint: What is $\text{range}(L)$?)

★12. Suppose that $L: \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator given by $L(\mathbf{X}) = \mathbf{AX}$, where $|\mathbf{A}| \neq 0$. What is $\ker(L)$? What is $\text{range}(L)$?

13. Let $\mathcal{V}$ be a finite dimensional vector space, and let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator. Show that $\ker(L) = \{\mathbf{0}_{\mathcal{V}}\}$ if and only if $\text{range}(L) = \mathcal{V}$.

14. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Prove directly that $\ker(L)$ is a subspace of $\mathcal{V}$ and that $\text{range}(L)$ is a subspace of $\mathcal{W}$ using Theorem 4.2, that is, without invoking Theorem 5.8.

15. Let $L_1: \mathcal{V} \to \mathcal{W}$ and $L_2: \mathcal{W} \to \mathcal{X}$ be linear transformations.

(a) Show that $\ker(L_1) \subseteq \ker(L_2 \circ L_1)$.

(b) Show that $\text{range}(L_2 \circ L_1) \subseteq \text{range}(L_2)$.

(c) If $\mathcal{V}$ is finite dimensional, prove that $\dim(\text{range}(L_2 \circ L_1)) \leq \dim(\text{range}(L_1))$.

$\star$**16.** Give an example of a linear operator $L: \mathbb{R}^2 \to \mathbb{R}^2$ such that $\ker(L) = \text{range}(L)$.

17. Let $L: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with $m \times n$ matrix $\mathbf{A}$ for L with respect to the standard bases and $m \times n$ matrix $\mathbf{B}$ for L with respect to bases B and C.

(a) Prove that $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{B})$. (Hint: Use Exercise 16 in the Review Exercises of Chapter 2.)

(b) Use part (a) to finish the proof of Theorem 5.9. (Hint: Notice that Theorem 5.9 allows *any* bases to be used for $\mathbb{R}^n$ and $\mathbb{R}^m$. You can assume, from the remarks before Theorem 5.9, that the theorem is true when the standard bases are used for $\mathbb{R}^n$ and $\mathbb{R}^m$.)

18. This exercise outlines an alternate proof of the Dimension Theorem. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation with $\mathcal{V}$ finite dimensional. Figure 5.10 illustrates the relationships among the vectors referenced throughout this exercise.

(a) Let $\{\mathbf{k}_1, \ldots, \mathbf{k}_s\}$ be a basis for $\ker(L)$. Show that there exist vectors $\mathbf{q}_1, \ldots, \mathbf{q}_t$ such that $\{\mathbf{k}_1, \ldots, \mathbf{k}_s, \mathbf{q}_1, \ldots, \mathbf{q}_t\}$ is a basis for $\mathcal{V}$. Express $\dim(\mathcal{V})$ in terms of s and t.

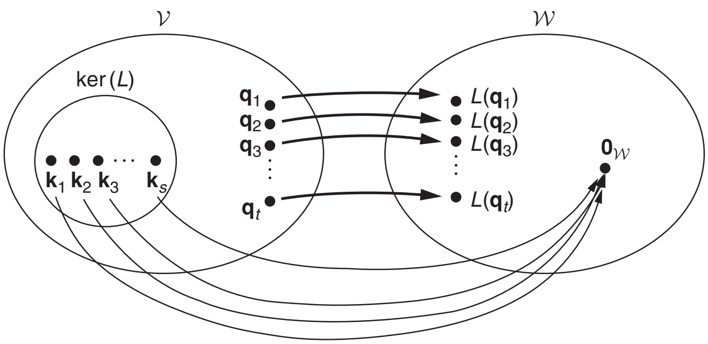

FIGURE 5.10

Images of basis elements in Exercise 18

(b) Use part (a) to show that for every $\mathbf{v} \in \mathcal{V}$, there exist scalars $b_1, \ldots, b_t$ such that $L(\mathbf{v}) = b_1 L(\mathbf{q}_1) + \cdots + b_t L(\mathbf{q}_t)$.

(c) Use part (b) to show that $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ spans range(L). Conclude that $\dim(\text{range}(L)) \leq t$, and, hence, is finite.

(d) Suppose that $c_1 L(\mathbf{q}_1) + \cdots + c_t L(\mathbf{q}_t) = \mathbf{0}_{\mathcal{W}}$. Prove that $c_1\mathbf{q}_1 + \cdots + c_t\mathbf{q}_t \in \ker(L)$.

(e) Use part (d) to show that there are scalars $d_1, \ldots, d_s$ such that $c_1\mathbf{q}_1 + \cdots + c_t\mathbf{q}_t = d_1\mathbf{k}_1 + \cdots + d_s\mathbf{k}_s$.

(f) Use part (e) and the fact that $\{\mathbf{k}_1, \ldots, \mathbf{k}_s, \mathbf{q}_1, \ldots, \mathbf{q}_t\}$ is a basis for $\mathcal{V}$ to prove that $c_1 = c_2 = \cdots = c_t = d_1 = \cdots = d_s = 0$.

(g) Use parts (d) and (f) to conclude that $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ is linearly independent.

(h) Use parts (c) and (g) to prove that $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ is a basis for range(L).

(i) Conclude that $\dim(\ker(L)) + \dim(\text{range}(L)) = \dim(\mathcal{V})$.

19. Prove the following corollary of the Dimension Theorem: Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation with $\mathcal{V}$ finite dimensional. Then $\dim(\ker(L)) \leq \dim(\mathcal{V})$ and $\dim(\text{range}(L)) \leq \dim(\mathcal{V})$.

⋆20. True or False:

(a) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then $\ker(L) = \{L(\mathbf{v}) \mid \mathbf{v} \in \mathcal{V}\}$.

(b) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then range(L) is a subspace of $\mathcal{V}$.

(c) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\dim(\mathcal{V}) = n$, then $\dim(\ker(L)) = n - \dim(\text{range}(L))$.

(d) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\dim(\mathcal{V}) = 5$ and $\dim(\mathcal{W}) = 3$, then the Dimension Theorem implies that $\dim(\ker(L)) = 2$.

(e) If $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $L(\mathbf{X}) = \mathbf{AX}$, then $\dim(\ker(L))$ equals the number of nonpivot columns in the reduced row echelon form matrix for $\mathbf{A}$.

(f) If $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $L(\mathbf{X}) = \mathbf{AX}$, then $\dim(\text{range}(L)) = n - \text{rank}(\mathbf{A})$.

(g) If $\mathbf{A}$ is a 5×5 matrix, and $\text{rank}(\mathbf{A}) = 2$, then $\text{rank}\left(\mathbf{A}^T\right) = 3$.

(h) If $\mathbf{A}$ is any matrix, then the row space of $\mathbf{A}$ equals the column space of $\mathbf{A}$.

5.4 ONE-TO-ONE AND ONTO LINEAR TRANSFORMATIONS

The kernel and the range of a linear transformation are related to the function properties one-to-one and onto. Consequently, in this section we study linear transformations that are one-to-one or onto.

One-to-One and Onto Linear Transformations

One-to-one functions and onto functions are defined and discussed in Appendix B. In particular, Appendix B contains the usual methods for proving that a given function is, or is not, one-to-one or onto. Now, we are interested primarily in linear transformations, so we restate the definitions of *one-to-one* and *onto* specifically as they apply to this type of function.

Definition Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation.

(1) L is **one-to-one** if and only if distinct vectors in $\mathcal{V}$ have different images in $\mathcal{W}$. That is, L is **one-to-one** if and only if, for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$, $L(\mathbf{v}_1) = L(\mathbf{v}_2)$ implies $\mathbf{v}_1 = \mathbf{v}_2$.

(2) L is **onto** if and only if every vector in the codomain $\mathcal{W}$ is the image of some vector in the domain $\mathcal{V}$. That is, L is **onto** if and only if, for every $\mathbf{w} \in \mathcal{W}$, there is some $\mathbf{v} \in \mathcal{V}$ such that $L(\mathbf{v}) = \mathbf{w}$.

Notice that the two descriptions of a one-to-one linear transformation given in this definition are really contrapositives of each other.

Example 1

Rotation: Recall the rotation linear operator $L: \mathbb{R}^2 \to \mathbb{R}^2$ from Example 9 in Section 5.1 given by $L(\mathbf{v}) = \mathbf{A}\mathbf{v}$, where $\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. We will show that L is *both one-to-one and onto.*

To show that L is one-to-one, we take any two arbitrary vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in the domain $\mathbb{R}^2$, assume that $L(\mathbf{v}_1) = L(\mathbf{v}_2)$, and prove that $\mathbf{v}_1 = \mathbf{v}_2$. Now, if $L(\mathbf{v}_1) = L(\mathbf{v}_2)$, then $\mathbf{A}\mathbf{v}_1 = \mathbf{A}\mathbf{v}_2$. Because $\mathbf{A}$ is nonsingular, we can multiply both sides on the left by $\mathbf{A}^{-1}$ to obtain $\mathbf{v}_1 = \mathbf{v}_2$. Hence, L is one-to-one.

To show that L is onto, we must take any arbitrary vector $\mathbf{w}$ in the codomain $\mathbb{R}^2$ and show that there is some vector $\mathbf{v}$ in the domain $\mathbb{R}^2$ that maps to $\mathbf{w}$. Recall that multiplication by $\mathbf{A}^{-1}$ undoes the action of multiplication by $\mathbf{A}$, and so it must represent a *clockwise* rotation through the angle θ. Hence, we can find a pre-image for $\mathbf{w}$ by rotating it *clockwise* through the angle θ; that is, consider $\mathbf{v} = \mathbf{A}^{-1}\mathbf{w} \in \mathbb{R}^2$. When we apply L to $\mathbf{v}$, we rotate it *counterclockwise* through the same angle θ: $L(\mathbf{v}) = \mathbf{A}(\mathbf{A}^{-1}\mathbf{w}) = \mathbf{w}$, thus obtaining the original vector $\mathbf{w}$. Since $\mathbf{v}$ is in the domain and $\mathbf{v}$ maps to $\mathbf{w}$ under L, L is onto. ∎

Example 2

Differentiation: Consider the linear transformation $L: \mathcal{P}_3 \to \mathcal{P}_2$ given by $L(\mathbf{p}) = \mathbf{p}'$. We will show that L is *onto but not one-to-one.*

To show that L is not one-to-one, we must find two different vectors $\mathbf{p}_1$ and $\mathbf{p}_2$ in the domain $\mathcal{P}_3$ that have the same image. Consider $\mathbf{p}_1 = x + 1$ and $\mathbf{p}_2 = x + 2$. Since $L(\mathbf{p}_1) = L(\mathbf{p}_2) = 1$, L is not one-to-one.

To show that L is onto, we must take an arbitrary vector $\mathbf{q}$ in $\mathcal{P}_2$ and find some vector $\mathbf{p}$ in $\mathcal{P}_3$ such that $L(\mathbf{p}) = \mathbf{q}$. Consider the vector $\mathbf{p} = \int \mathbf{q}(x)\,dx$ with zero constant term. Because $L(\mathbf{p}) = \mathbf{q}$, we see that L is onto.

If in Example 2 we had used $\mathcal{P}_3$ for the codomain instead of $\mathcal{P}_2$, the linear transformation would not have been onto because x^3 would have no pre-image (why?). This provides an example of a linear transformation that is neither one-to-one nor onto. Also, Exercise 6 illustrates a linear transformation that is one-to-one but not onto. These examples, together with Examples 1 and 2, show that the concepts of one-to-one and onto are independent of each other; that is, there are linear transformations that have either property with or without the other.

Theorem B.1 in Appendix B shows that the composition of one-to-one linear transformations is one-to-one, and similarly, the composition of onto linear transformations is onto.

Kernel and Range

The next theorem gives an alternate way of characterizing one-to-one linear transformations and onto linear transformations.

Theorem 5.12 Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, and let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then:

(1) L is one-to-one if and only if $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$ (or, equivalently, if and only if $\dim(\ker(L)) = 0$), and

(2) If $\mathcal{W}$ is finite dimensional, then L is onto if and only if $\dim(\text{range}(L)) = \dim(\mathcal{W})$.

Thus, a linear transformation whose kernel contains a nonzero vector cannot be one-to-one.

Proof. First suppose that L is one-to-one, and let $\mathbf{v} \in \ker(L)$. We must show that $\mathbf{v} = \mathbf{0}_\mathcal{V}$. Now, $L(\mathbf{v}) = \mathbf{0}_\mathcal{W}$. However, by Theorem 5.1, $L(\mathbf{0}_\mathcal{V}) = \mathbf{0}_\mathcal{W}$. Because $L(\mathbf{v}) = L(\mathbf{0}_\mathcal{V})$ and L is one-to-one, we must have $\mathbf{v} = \mathbf{0}_\mathcal{V}$.

Conversely, suppose that $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$. We must show that L is one-to-one. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$, with $L(\mathbf{v}_1) = L(\mathbf{v}_2)$. We must show that $\mathbf{v}_1 = \mathbf{v}_2$. Now, $L(\mathbf{v}_1) - L(\mathbf{v}_2) = \mathbf{0}_\mathcal{W}$, implying that $L(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}_\mathcal{W}$. Hence, $\mathbf{v}_1 - \mathbf{v}_2 \in \ker(L)$, by definition of the kernel. Since $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}_\mathcal{V}$ and so $\mathbf{v}_1 = \mathbf{v}_2$.

Finally, note that, by definition, L is onto if and only if $\mathbf{range}(L) = \mathcal{W}$, and therefore part (2) of the theorem follows immediately from Theorem 4.16. □

Example 3

Consider the linear transformation $L: \mathcal{M}_{22} \to \mathcal{M}_{23}$ given by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a-b & 0 & c-d \\ c+d & a+b & 0 \end{bmatrix}$. If $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \ker(L)$, then $a-b = c-d = c+d = a+b = 0$. Solving these equations yields $a = b = c = d = 0$, and so $\ker(L)$ contains only the zero matrix $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$; that is, $\dim(\ker(L)) = 0$. Thus, by part (1) of Theorem 5.12, L is one-to-one. However, by the Dimension Theorem, $\dim(\mathbf{range}(L)) = \dim(\mathcal{M}_{22}) - \dim(\ker(L)) = \dim(\mathcal{M}_{22}) = 4$. Hence, by part (2) of Theorem 5.12, L is not onto. In particular, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \notin \mathbf{range}(L)$.

On the other hand, consider $M: \mathcal{M}_{23} \to \mathcal{M}_{22}$ given by $M\left(\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\right) = \begin{bmatrix} a+b & a+c \\ d+e & d+f \end{bmatrix}$. It is easy to see that M is onto, since $M\left(\begin{bmatrix} 0 & b & c \\ 0 & e & f \end{bmatrix}\right) = \begin{bmatrix} b & c \\ e & f \end{bmatrix}$, and thus every 2×2 matrix is in $\mathbf{range}(M)$. Thus, by part (2) of Theorem 5.12, $\dim(\mathbf{range}(M)) = \dim(\mathcal{M}_{22}) = 4$. Then, by the Dimension Theorem, $\ker(M) = \dim(\mathcal{M}_{23}) - \dim(\mathbf{range}(M)) = 6 - 4 = 2$. Hence, by part (1) of Theorem 5.12, M is not one-to-one. In particular, $\begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \in \ker(L)$. ■

Spanning and Linear Independence

The next theorem shows that the one-to-one property is related to linear independence, while the onto property is related to spanning.

Theorem 5.13 Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces, and let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then:

(1) If L is one-to-one, and T is a linearly independent subset of $\mathcal{V}$, then $L(T)$ is linearly independent in $\mathcal{W}$.

(2) If L is onto, and S spans $\mathcal{V}$, then $L(S)$ spans $\mathcal{W}$.

Proof. Suppose that L is one-to-one, and T is a linearly independent subset of $\mathcal{V}$. To prove that $L(T)$ is linearly independent in $\mathcal{W}$, it is enough to show that any finite subset of $L(T)$ is linearly independent. Suppose $\{L(\mathbf{x}_1), \ldots, L(\mathbf{x}_n)\}$ is a finite subset of $L(T)$, for vectors $\mathbf{x}_1,\ldots,\mathbf{x}_n \in T$, and suppose $b_1L(\mathbf{x}_1)+\cdots+b_nL(\mathbf{x}_n)=\mathbf{0}_{\mathcal{W}}$. Then, $L(b_1\mathbf{x}_1+\cdots+b_n\mathbf{x}_n)=\mathbf{0}_{\mathcal{W}}$, implying that $b_1\mathbf{x}_1+\cdots+b_n\mathbf{x}_n \in \ker(L)$. But since L is one-to-one, Theorem 5.12 tells us that $\ker(L)=\{\mathbf{0}_{\mathcal{V}}\}$. Hence, $b_1\mathbf{x}_1+\cdots+b_n\mathbf{x}_n=\mathbf{0}_{\mathcal{V}}$. Then, because the vectors in T are linearly independent, $b_1=b_2=\cdots=b_n=0$. Therefore, $\{L(\mathbf{x}_1),\ldots,L(\mathbf{x}_n)\}$ is linearly independent. Hence, $L(T)$ is linearly independent.

Now suppose that L is onto, and S spans $\mathcal{V}$. To prove that $L(S)$ spans $\mathcal{W}$, we must show that any vector $\mathbf{w}\in\mathcal{W}$ can be expressed as a linear combination of vectors in $L(S)$. Since L is onto, there is a $\mathbf{v}\in\mathcal{V}$ such that $L(\mathbf{v})=\mathbf{w}$. Since S spans $\mathcal{V}$, there are scalars $a_1,\ldots,a_n$ and vectors $\mathbf{v}_1,\ldots,\mathbf{v}_n\in S$ such that $\mathbf{v}=a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n$. Thus, $\mathbf{w}=L(\mathbf{v})=L(a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n)=a_1L(\mathbf{v}_1)+\cdots+a_nL(\mathbf{v}_n)$. Hence, $L(S)$ spans $\mathcal{W}$. □

An almost identical proof gives the following useful generalization of part (2) of Theorem 5.13: For any linear transformation $L:\mathcal{V}\to\mathcal{W}$, and any subset S of $\mathcal{V}$, $L(S)$ spans the subspace $L(\text{span}(S))$ of $\mathcal{W}$. In particular, if S spans $\mathcal{V}$, then $L(S)$ spans range(L). (See Exercise 8.)

Example 4

Consider the linear transformation $L: P_2 \to P_3$ given by $L(ax^2+bx+c)=bx^3+cx^2+ax$. It is easy to see that $\ker(L)=\{\mathbf{0}\}$ since $L(ax^2+bx+c)=0x^3+0x^2+0x+0$ only if $a=b=c=0$, and so L is one-to-one by Theorem 5.12. Consider the linearly independent set $T=\{x^2+x,\ x+1\}$ in P_2. Notice that $L(T)=\{x^3+x,\ x^3+x^2\}$, and that $L(T)$ is linearly independent, as predicted by part (1) of Theorem 5.13.

Next, let $\mathcal{W}=\{[x,0,z]\}$ be the xz-plane in $\mathbb{R}^3$. Clearly, $\dim(\mathcal{W})=2$. Consider $L:\mathbb{R}^3\to\mathcal{W}$, where L is the projection of $\mathbb{R}^3$ onto the xz-plane; that is, $L([x,y,z])=[x,0,z]$. It is easy to check that $S=\left\{[2,-1,3],[1,-2,0],[4,3,-1]\right\}$ spans $\mathbb{R}^3$ using the Simplified Span Method. Part (2) of Theorem 5.13 then asserts that $L(S)=\left\{[2,0,3],[1,0,0],[4,0,-1]\right\}$ spans $\mathcal{W}$. In fact, $\{[2,0,3],[1,0,0]\}$ alone spans $\mathcal{W}$, since $\dim(\text{span}(\{[2,0,3],[1,0,0]\}))=2=\dim(\mathcal{W})$. ■

In Section 5.5, we will consider isomorphisms, which are linear transformations that are simultaneously one-to-one and onto. We will see that such functions faithfully carry vector space properties from the domain to the codomain.

New Vocabulary

one-to-one linear transformation — onto linear transformation

Highlights

- A linear transformation is one-to-one if no two distinct vectors of the domain map to the same image in the codomain.
- A linear transformation $L:\ \mathcal{V}\to\mathcal{W}$ is one-to-one if and only if $\ker(L)=\{\mathbf{0}_{\mathcal{V}}\}$ (or, equivalently, if and only if $\dim(\ker(L))=0$).
- If a linear transformation is one-to-one, then the image of every linearly independent subset of the domain is linearly independent.

- A linear transformation is onto if every vector in the codomain is the image of some vector from the domain.
- A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is onto if and only if range$(L) = \mathcal{W}$ (or, equivalently, if and only if dim(range(L)) = dim$(\mathcal{W})$ when $\mathcal{W}$ is finite dimensional).
- If a linear transformation is onto, then the image of every spanning set for the domain spans the codomain.

EXERCISES FOR SECTION 5.4

1. Which of the following linear transformations are one-to-one? Which are onto? Justify your answers without using row reduction.

★**(a)** $L\colon \mathbb{R}^3 \to \mathbb{R}^4$ given by $L([x,y,z]) = [y,z,-y,0]$

(b) $L\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by $L([x,y,z]) = [x+y, y+z]$

★**(c)** $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by $L([x,y,z]) = [2x, x+y+z, -y]$

(d) $L\colon \mathcal{P}_3 \to \mathcal{P}_2$ given by $L(ax^3+bx^2+cx+d) = ax^2+bx+c$

★**(e)** $L\colon \mathcal{P}_2 \to \mathcal{P}_2$ given by $L(ax^2+bx+c) = (a+b)x^2+(b+c)x+(a+c)$

(f) $L\colon \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} d & b+c \\ b-c & a \end{bmatrix}$

★**(g)** $L\colon \mathcal{M}_{23} \to \mathcal{M}_{22}$ given by $L\left(\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\right) = \begin{bmatrix} a & -c \\ 2e & d+f \end{bmatrix}$

★**(h)** $L\colon \mathcal{P}_2 \to \mathcal{M}_{22}$ given by $L(ax^2+bx+c) = \begin{bmatrix} a+c & 0 \\ b-c & -3a \end{bmatrix}$

2. Which of the following linear transformations are one-to-one? Which are onto? Justify your answers by using row reduction to determine the dimensions of the kernel and range.

★**(a)** $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -4 & -3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

★**(b)** $L\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by $L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -3 & 4 \\ -6 & 9 \\ 7 & -8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

★**(c)** $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by $L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -7 & 4 & -2 \\ 16 & -7 & 2 \\ 4 & -3 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(d) $L: \mathbb{R}^4 \to \mathbb{R}^3$ given by $L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} -5 & 3 & 1 & 18 \\ -2 & 1 & 1 & 6 \\ -7 & 3 & 4 & 19 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$

3. In each of the following cases, the matrix for a linear transformation with respect to some ordered bases for the domain and codomain is given. Which of these linear transformations are one-to-one? Which are onto? Justify your answers by using row reduction to determine the dimensions of the kernel and range.

★(a) $L: \mathcal{P}_2 \to \mathcal{P}_2$ having matrix $\begin{bmatrix} 1 & -3 & 0 \\ -4 & 13 & -1 \\ 8 & -25 & 2 \end{bmatrix}$

(b) $L: \mathcal{M}_{22} \to \mathcal{M}_{22}$ having matrix $\begin{bmatrix} 6 & -9 & 2 & 8 \\ 10 & -6 & 12 & 4 \\ -3 & 3 & -4 & -4 \\ 8 & -9 & 9 & 11 \end{bmatrix}$

★(c) $L: \mathcal{M}_{22} \to \mathcal{P}_3$ having matrix $\begin{bmatrix} 2 & 3 & -1 & 1 \\ 5 & 2 & -4 & 7 \\ 1 & 7 & 1 & -4 \\ -2 & 19 & 7 & -19 \end{bmatrix}$

4. Suppose that $m > n$.

(a) Show there is no onto linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

(b) Show there is no one-to-one linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$.

5. Let $\mathbf{A}$ be a fixed $n \times n$ matrix, and consider $L: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L(\mathbf{B}) = \mathbf{AB} - \mathbf{BA}$.

(a) Show that L is not one-to-one. (Hint: Consider $L(\mathbf{I}_n)$.)

(b) Use part (a) to show that L is not onto.

6. Define $L: \mathcal{U}_3 \to \mathcal{M}_{33}$ by $L(\mathbf{A}) = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$. Prove that L is one-to-one but is *not* onto.

7. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation between vector spaces. Suppose that for every linearly independent set T in $\mathcal{V}$, $L(T)$ is linearly independent in $\mathcal{W}$. Prove that L is one-to-one. (Hint: Prove $\ker(L) = \{\mathbf{0}_{\mathcal{V}}\}$ using a proof by contradiction.)

8. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation between vector spaces, and let S be a subset of $\mathcal{V}$.

(a) Prove that $L(S)$ spans the subspace $L(\mathrm{span}(S))$.

(b) Show that if S spans $\mathcal{V}$, then $L(S)$ spans range(L).

(c) Show that if $L(S)$ spans $\mathcal{W}$, then L is onto.

★9. True or False:

(a) A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is one-to-one if for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$, $\mathbf{v}_1 = \mathbf{v}_2$ implies $L(\mathbf{v}_1) = L(\mathbf{v}_2)$.

(b) A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is onto if for all $\mathbf{v} \in \mathcal{V}$, there is some $\mathbf{w} \in \mathcal{W}$ such that $L(\mathbf{v}) = \mathbf{w}$.

(c) A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is one-to-one if $\ker(L)$ contains no vectors other than $\mathbf{0}_{\mathcal{V}}$.

(d) If L is a linear transformation and S spans the domain of L, then $L(S)$ spans the range of L.

(e) Suppose $\mathcal{V}$ is a finite dimensional vector space. A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is not one-to-one if $\dim(\ker(L)) \neq 0$.

(f) Suppose $\mathcal{W}$ is a finite dimensional vector space. A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is not onto if $\dim(\text{range}(L)) < \dim(\mathcal{W})$.

(f) If L is a linear transformation and T is a linearly independent subset of the domain of L, then $L(T)$ is linearly independent.

(g) If L is a linear transformation $L: \mathcal{V} \to \mathcal{W}$, and S is a subset of $\mathcal{V}$ such that $L(S)$ spans $\mathcal{W}$, then S spans $\mathcal{V}$.

5.5 ISOMORPHISM

In this section, we examine methods for determining whether two vector spaces are equivalent, or *isomorphic*. Isomorphism is important because if certain algebraic results are true in one of two isomorphic vector spaces, corresponding results hold true in the other as well. It is the concept of isomorphism that has allowed us to apply our techniques and formal methods to vector spaces other than $\mathbb{R}^n$.

Isomorphisms: Invertible Linear Transformations

We restate here the definition from Appendix B for the inverse of a function as it applies to linear transformations.

> **Definition** Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then L is an **invertible linear transformation** if and only if there is a function $M: \mathcal{W} \to \mathcal{V}$ such that $(M \circ L)(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v} \in \mathcal{V}$, and $(L \circ M)(\mathbf{w}) = \mathbf{w}$, for all $\mathbf{w} \in \mathcal{W}$. Such a function M is called an **inverse** of L.

If the inverse M of $L: \mathcal{V} \to \mathcal{W}$ exists, then it is unique by Theorem B.3 and is usually denoted by $L^{-1}: \mathcal{W} \to \mathcal{V}$.

Definition A linear transformation $L: \mathcal{V} \to \mathcal{W}$ that is both one-to-one and onto is called an **isomorphism** from $\mathcal{V}$ to $\mathcal{W}$.

The next result shows that the previous two definitions actually refer to the same class of linear transformations.

Theorem 5.14 Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then L is an isomorphism if and only if L is an invertible linear transformation. Moreover, if L is invertible, then L^{-1} is also a linear transformation.

Notice that Theorem 5.14 also asserts that whenever L is an isomorphism, L^{-1} is an isomorphism as well because L^{-1} is an invertible linear transformation (with L as its inverse).

Proof. The "if and only if" part of Theorem 5.14 follows directly from Theorem B.2. Thus, we only need to prove the last assertion in Theorem 5.14. That is, suppose $L: \mathcal{V} \to \mathcal{W}$ is invertible (and thus, an isomorphism) with inverse L^{-1}. We need to prove L^{-1} is a linear transformation. To do this, we must show both of the following properties hold:

(1) $L^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = L^{-1}(\mathbf{w}_1) + L^{-1}(\mathbf{w}_2)$, for all $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}$

(2) $L^{-1}(c\mathbf{w}) = cL^{-1}(\mathbf{w})$, for all $c \in \mathbb{R}$, and for all $\mathbf{w} \in \mathcal{W}$.

Property (1): Because L is an isomorphism, L is one-to-one. Hence, if we can show that $L(L^{-1}(\mathbf{w}_1 + \mathbf{w}_2)) = L(L^{-1}(\mathbf{w}_1) + L^{-1}(\mathbf{w}_2))$, we will be done. But,

$$\begin{aligned} L(L^{-1}(\mathbf{w}_1) + L^{-1}(\mathbf{w}_2)) &= L(L^{-1}(\mathbf{w}_1)) + L(L^{-1}(\mathbf{w}_2)) \\ &= \mathbf{w}_1 + \mathbf{w}_2 \\ &= L(L^{-1}(\mathbf{w}_1 + \mathbf{w}_2)). \end{aligned}$$

Property (2): Again, because L is an isomorphism, L is one-to-one. Hence, if we can show that $L(L^{-1}(c\mathbf{w})) = L(cL^{-1}(\mathbf{w}))$, we will be done. But,

$$\begin{aligned} L(cL^{-1}(\mathbf{w})) &= cL(L^{-1}(\mathbf{w})) \\ &= c\mathbf{w} \\ &= L(L^{-1}(c\mathbf{w})). \end{aligned}$$

Because both properties (1) and (2) hold, L^{-1} is a linear transformation. □

Example 1

Recall the rotation linear operator $L: \mathbb{R}^2 \to \mathbb{R}^2$ with

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

given in Example 9 in Section 5.1. In Example 1 in Section 5.4, we proved that L is both one-to-one and onto. Hence, L is an isomorphism and has an inverse, L^{-1}. Because L represents a *counterclockwise* rotation of vectors through the angle θ, then L^{-1} must represent a *clockwise* rotation through the angle θ, as we saw in Example 1 of Section 5.4. Equivalently, L^{-1} can be thought of as a *counterclockwise* rotation through the angle $-\theta$. Thus,

$$L^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Of course, L^{-1} is also an isomorphism.

The next theorem gives a simple method for determining whether a linear transformation between finite dimensional vector spaces is an isomorphism.

Theorem 5.15 Let $\mathcal{V}$ and $\mathcal{W}$ both be nontrivial finite dimensional vector spaces with ordered bases B and C, respectively, and let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then L is an isomorphism if and only if the matrix representation $\mathbf{A}_{BC}$ for L with respect to B and C is nonsingular.

To prove one half of Theorem 5.15, let $\mathbf{A}_{BC}$ be the matrix for L with respect to B and C, and let $\mathbf{D}_{CB}$ be the matrix for L^{-1} with respect to C and B. Theorem 5.7 then shows that $\mathbf{D}_{CB}\mathbf{A}_{BC} = \mathbf{I}_n$, with $n = \dim(\mathcal{V})$, and $\mathbf{A}_{BC}\mathbf{D}_{CB} = \mathbf{I}_k$, with $k = \dim(\mathcal{W})$. By Exercise 21 in Section 2.4, $n = k$, and $(\mathbf{A}_{BC})^{-1} = \mathbf{D}_{CB}$, so $\mathbf{A}_{BC}$ is nonsingular. The proof of the converse is straightforward, and you are asked to give the details in Exercise 8. Notice, in particular, that the matrix for any isomorphism must be a square matrix.

Example 2

Consider $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by $L(\mathbf{v}) = \mathbf{A}\mathbf{v}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now, $\mathbf{A}$ is nonsingular ($|\mathbf{A}| = 1 \neq 0$). Hence, by Theorem 5.15, L is an isomorphism. Geometrically, L represents a shear in the z-direction (see Table 5.1).

Theorem B.4 in Appendix B shows that the composition of isomorphisms results in an isomorphism. In particular, the inverse of the composition $L_2 \circ L_1$ is $L_1^{-1} \circ L_2^{-1}$. That is, the transformations must be undone in *reverse* order to arrive at the correct inverse. (Compare this with part (3) of Theorem 2.11 for matrix multiplication.)

When an isomorphism exists between two vector spaces, properties from the domain are carried over to the codomain by the isomorphism. In particular, the following theorem, which follows immediately from Theorem 5.13, shows that spanning sets map to spanning sets, and linearly independent sets map to linearly independent sets.

Theorem 5.16 Suppose $L: \mathcal{V} \to \mathcal{W}$ is an isomorphism. Let S span $\mathcal{V}$ and let T be a linearly independent subset of $\mathcal{V}$. Then $L(S)$ spans $\mathcal{W}$ and $L(T)$ is linearly independent.

Isomorphic Vector Spaces

Definition Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces. Then $\mathcal{V}$ is **isomorphic** to $\mathcal{W}$, denoted $\mathcal{V} \cong \mathcal{W}$, if and only if there exists an isomorphism $L: \mathcal{V} \to \mathcal{W}$.

If $\mathcal{V} \cong \mathcal{W}$, there is some isomorphism $L: \mathcal{V} \to \mathcal{W}$. Then by Theorem 5.14, $L^{-1}: \mathcal{W} \to \mathcal{V}$ is also an isomorphism, so $\mathcal{W} \cong \mathcal{V}$. Hence, we usually speak of such $\mathcal{V}$ and $\mathcal{W}$ as being *isomorphic to each other*.

Also notice that if $\mathcal{V} \cong \mathcal{W}$ and $\mathcal{W} \cong \mathcal{X}$, then there are isomorphisms $L_1: \mathcal{V} \to \mathcal{W}$ and $L_2: \mathcal{W} \to \mathcal{X}$. But then $L_2 \circ L_1: \mathcal{V} \to \mathcal{X}$ is an isomorphism, and so $\mathcal{V} \cong \mathcal{X}$. In other words, two vector spaces such as $\mathcal{V}$ and $\mathcal{X}$ that are both isomorphic to the same vector space $\mathcal{W}$ are isomorphic to each other.

Example 3

Consider $L_1: \mathbb{R}^4 \to \mathcal{P}_3$ given by $L_1([a,b,c,d]) = ax^3 + bx^2 + cx + d$ and $L_2: \mathcal{M}_{22} \to \mathcal{P}_3$ given by $L_2\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ax^3 + bx^2 + cx + d$. L_1 and L_2 are certainly both isomorphisms. Hence, $\mathbb{R}^4 \cong \mathcal{P}_3$ and $\mathcal{M}_{22} \cong \mathcal{P}_3$. Thus, the composition $L_2^{-1} \circ L_1: \mathbb{R}^4 \to \mathcal{M}_{22}$ is also an isomorphism, and so $\mathbb{R}^4 \cong \mathcal{M}_{22}$. Notice that all of these vector spaces have dimension 4. ■

Next, we show that finite dimensional vector spaces $\mathcal{V}$ and $\mathcal{W}$ must have the same dimension for an isomorphism to exist between them.

Theorem 5.17 Suppose $\mathcal{V} \cong \mathcal{W}$ and $\mathcal{V}$ is finite dimensional. Then $\mathcal{W}$ is finite dimensional and $\dim(\mathcal{V}) = \dim(\mathcal{W})$.

Proof. Since $\mathcal{V} \cong \mathcal{W}$, there is an isomorphism $L: \mathcal{V} \to \mathcal{W}$. Let $\dim(\mathcal{V}) = n$, and let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for $\mathcal{V}$. By Theorem 5.16, $L(B) = \{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ both spans $\mathcal{W}$ and is linearly independent, and so must be a basis for $\mathcal{W}$. Also, because L is a one-to-one function, $|L(B)| = |B| = n$. Therefore, $\dim(\mathcal{V}) = \dim(\mathcal{W})$. □

Theorem 5.17 implies that there is no possible isomorphism from, say, $\mathbb{R}^3$ to $\mathcal{P}_4$ or from $\mathcal{M}_{22}$ to $\mathbb{R}^3$, because the dimensions of the spaces do not agree. Notice that Theorem 5.17 gives another confirmation of the fact that any matrix for an isomorphism must be square.

Isomorphism of n-Dimensional Vector Spaces

Example 3 hints that any two finite dimensional vector spaces of the same dimension are isomorphic. This result, which is one of the most important in all linear algebra, is a corollary of the next theorem.

> **Theorem 5.18** If $\mathcal{V}$ is any n-dimensional vector space, then $\mathcal{V} \cong \mathbb{R}^n$.

Proof. Suppose that $\mathcal{V}$ is a vector space with $\dim(\mathcal{V}) = n$. If we can find an isomorphism $L: \mathcal{V} \to \mathbb{R}^n$, then $\mathcal{V} \cong \mathbb{R}^n$, and we will be done. Let $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an ordered basis for $\mathcal{V}$. Consider the mapping $L(\mathbf{v}) = [\mathbf{v}]_B$, for all $\mathbf{v} \in \mathcal{V}$. Now, L is a linear transformation by Example 4 in Section 5.1. Also,

$$\mathbf{v} \in \ker(L) \;\Leftrightarrow\; [\mathbf{v}]_B = [0, \ldots, 0] \;\Leftrightarrow\; \mathbf{v} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n \;\Leftrightarrow\; \mathbf{v} = \mathbf{0}.$$

Hence, $\ker(L) = \{\mathbf{0}_{\mathcal{V}}\}$, and L is one-to-one.

If $\mathbf{a} = [a_1, \ldots, a_n] \in \mathbb{R}^n$, then $L(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = [a_1, \ldots, a_n]$, showing that $\mathbf{a} \in \text{range}(L)$. Hence, L is onto, and so L is an isomorphism. □

In particular, Theorem 5.18 tells us that $\mathcal{P}_n \cong \mathbb{R}^{n+1}$ and that $\mathcal{M}_{mn} \cong \mathbb{R}^{mn}$. Also, the proof of Theorem 5.18 illustrates that coordinatization of vectors in an n-dimensional vector space $\mathcal{V}$ automatically gives an isomorphism of $\mathcal{V}$ with $\mathbb{R}^n$.

By the remarks before Example 3, Theorem 5.18 implies the following converse of Theorem 5.17:

> **Corollary 5.19** Any two n-dimensional vector spaces $\mathcal{V}$ and $\mathcal{W}$ are isomorphic. That is, if $\dim(\mathcal{V}) = \dim(\mathcal{W})$, then $\mathcal{V} \cong \mathcal{W}$.

For example, suppose that $\mathcal{V}$ and $\mathcal{W}$ are both vector spaces with $\dim(\mathcal{V}) = \dim(\mathcal{W}) = 47$. Then by Corollary 5.19, $\mathcal{V} \cong \mathcal{W}$, and by Theorem 5.18, $\mathcal{V} \cong \mathcal{W} \cong \mathbb{R}^{47}$.

Isomorphism and the Methods

We now have the means to justify the use of the Simplified Span Method and the Independence Test Method on vector spaces other than $\mathbb{R}^n$. Suppose $\mathcal{V} \cong \mathbb{R}^n$. By using the coordinatization isomorphism or its inverse as the linear transformation L in Theorem 5.16, we see that spanning sets in $\mathcal{V}$ are mapped to spanning sets in $\mathbb{R}^n$, and vice versa. Similarly, linearly independent sets in $\mathcal{V}$ are mapped to linearly independent sets in $\mathbb{R}^n$, and vice versa. This is illustrated in the following example.

Example 4

Consider the subset $S = \{x^3 - 2x^2 + x - 2,\ x^3 + x^2 + x + 1,\ x^3 - 5x^2 + x - 5,\ x^3 - x^2 - x + 1\}$ of $\mathcal{P}_3$. We use the coordinatization isomorphism $L\colon \mathcal{P}_3 \to \mathbb{R}^4$ with respect to the standard basis of $\mathcal{P}_3$ to obtain $L(S) = \{[1, -2, 1, -2], [1, 1, 1, 1], [1, -5, 1, -5], [1, -1, -1, 1]\}$, a subset of $\mathbb{R}^4$ corresponding to S. Row reducing

$$\begin{bmatrix} 1 & -2 & 1 & -2 \\ 1 & 1 & 1 & 1 \\ 1 & -5 & 1 & -5 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad \text{to obtain} \quad \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

shows, by the Simplified Span Method, that $\mathbf{span}(\{[1,-2,1,-2],[1,1,1,1],[1,-5,1,-5],[1,-1,-1,1]\}) = \mathbf{span}(\{[1,0,0,1],\ [0,1,0,1],\ [0,0,1,-1]\})$. Since L^{-1} is an isomorphism, Theorem 5.16 shows that $L^{-1}(\{[1,0,0,1],[0,1,0,1],[0,0,1,-1]\}) = \{x^3 + 1,\ x^2 + 1,\ x - 1\}$ spans the same subspace of $\mathcal{P}_3$ that S does. That is, $\mathbf{span}(\{x^3 + 1,\ x^2 + 1,\ x - 1\}) = \mathbf{span}(S)$.

Similarly, row reducing

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -2 & 1 & -5 & -1 \\ 1 & 1 & 1 & -1 \\ -2 & 1 & -5 & 1 \end{bmatrix} \quad \text{to obtain} \quad \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

shows, by the Independence Test Method, that $\{[1,-2,1,-2],[1,1,1,1],[1,-1,-1,1]\}$ is a linearly independent subset of $\mathbb{R}^4$, and that $[1,-5,1,-5] = 2[1,-2,1,-2] - [1,1,1,1] + 0[1,-1,-1,1]$. Since L^{-1} is an isomorphism, Theorem 5.16 shows us that $L^{-1}\left(\{[1,-2,1,-2],[1,1,1,1],[1,-1,-1,1]\}\right) = \left\{x^3 - 2x^2 + x - 2, x^3 + x^2 + x + 1, x^3 - x^2 - x + 1\right\}$ is a linearly independent subset of $\mathcal{P}_3$. The fact that L^{-1} is a linear transformation also assures us that $x^3 - 5x^2 + x - 5 = 2\left(x^3 - 2x^2 + x - 2\right) - \left(x^3 + x^2 + x + 1\right) + 0\left(x^3 - x^2 - x + 1\right)$. ■

In addition to preserving dimension, spanning, and linear independence, isomorphisms keep intact most other properties of vector spaces and the linear transformations between them. In particular, the next theorem shows that when we coordinatize the domain and codomain of a linear transformation, the kernel and the range are preserved.

Theorem 5.20 Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation between nontrivial finite dimensional vector spaces, and let $L_1: \mathcal{V} \to \mathbb{R}^n$ and $L_2: \mathcal{W} \to \mathbb{R}^m$ be coordinatization isomorphisms with respect to some ordered bases B and C for $\mathcal{V}$ and $\mathcal{W}$, respectively. Let $M = L_2 \circ L \circ L_1^{-1}: \mathbb{R}^n \to \mathbb{R}^m$, so that $M([\mathbf{v}]_B) = [L(\mathbf{v})]_C$. Then,

(1) $L_1^{-1}(\ker(M)) = \ker(L) \subseteq \mathcal{V}$,

(2) $L_2^{-1}(\text{range}(M)) = \text{range}(L) \subseteq \mathcal{W}$,

(3) $\dim(\ker(M)) = \dim(\ker(L))$, and

(4) $\dim(\text{range}(M)) = \dim(\text{range}(L))$.

Figure 5.11 illustrates the situation in Theorem 5.20. The linear transformation M in Theorem 5.20 is merely an "$\mathbb{R}^n \to \mathbb{R}^m$" version of L, using coordinatized vectors instead of the actual vectors in $\mathcal{V}$ and $\mathcal{W}$. Because L_1^{-1} and L_2^{-1} are isomorphisms, parts (1) and (2) of the theorem show that the subspace $\ker(L)$ of $\mathcal{V}$ is isomorphic to the subspace $\ker(M)$ of $\mathbb{R}^n$, and that the subspace range(L) of $\mathcal{W}$ is isomorphic to the subspace range(M) of $\mathbb{R}^m$. Parts (3) and (4) of the theorem follow directly from parts (1) and (2) because isomorphic finite dimensional vector spaces must have the same dimension. You are asked to prove a more general version of Theorem 5.20 as well as other related statements in Exercises 17 and 18.

The importance of Theorem 5.20 is that it justifies our use of the Kernel Method and the Range Method of Section 5.3 when vector spaces other than $\mathbb{R}^n$ are involved. Suppose that we want to find $\ker(L)$ and range(L) for a given linear transformation L: $\mathcal{V} \to \mathcal{W}$. We begin by coordinatizing the domain $\mathcal{V}$ and codomain $\mathcal{W}$ using coordinatization isomorphisms L_1 and L_2 as in Theorem 5.20. (For simplicity, we can assume B and C are the standard bases for $\mathcal{V}$ and $\mathcal{W}$, respectively.) The mapping M created in Theorem 5.20 is thus an equivalent "$\mathbb{R}^n \to \mathbb{R}^m$" version of L. By applying the Kernel and Range Methods to M, we can find bases for $\ker(M)$ and range(M).

$$\begin{tikzcd} \mathcal{V} \arrow[r, "L"] \arrow[d, "L_1"'] & \mathcal{W} \arrow[d, "L_2"] \\ \mathbb{R}^n \arrow[r, "M"'] & \mathbb{R}^m \end{tikzcd}$$

FIGURE 5.11

The linear transformations L and M and the isomorphisms L_1 and L_2 in Theorem 5.20

However, parts (1) and (2) of the theorem assure us that $\ker(L)$ is isomorphic to $\ker(M)$, and, similarly, that $\text{range}(L)$ is isomorphic to $\text{range}(M)$. Therefore, by reversing the coordinatizations, we can find bases for $\ker(L)$ and $\text{range}(L)$. In fact, this is exactly the approach that was used without justification in Section 5.3 to determine bases for the kernel and range for linear transformations involving vector spaces other than $\mathbb{R}^n$.

Proving the Dimension Theorem Using Isomorphism

Recall the Dimension Theorem:

(Dimension Theorem) If $L\colon \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\mathcal{V}$ is finite dimensional, then $\mathbf{range}(L)$ is finite dimensional, and

$$\dim(\ker(L)) + \dim(\mathbf{range}(L)) = \dim(\mathcal{V}).$$

In Section 5.3, we stated the Dimension Theorem in its full generality, but only *proved* it for linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. We now supply the general proof, assuming that the special case for linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ has already been proved.

Proof. The theorem is obviously true if $\mathcal{V}$ is the trivial vector space. Suppose B is a finite, nonempty ordered basis for $\mathcal{V}$. Then, by the comments directly after Theorem 5.13 regarding spanning sets and range, $\mathbf{range}(L)$ is spanned by the finite set $L(B)$, and so $\mathbf{range}(L)$ is finite dimensional. Since L does not interact at all with the vectors in $\mathcal{W}$ outside $\mathbf{range}(L)$, we can consider adjusting L so that its codomain is just the subspace $\mathbf{range}(L)$ of $\mathcal{W}$. That is, without loss of generality, we can let $\mathcal{W} = \mathbf{range}(L)$. Hence, we can assume that $\mathcal{W}$ is finite dimensional.

Let $L_1\colon \mathcal{V} \to \mathbb{R}^n$ and $L_2\colon \mathcal{W} \to \mathbb{R}^m$ be coordinatization transformations with respect to some ordered bases for $\mathcal{V}$ and $\mathcal{W}$, respectively. Applying the special case of the Dimension Theorem to the linear transformation $L_2 \circ L \circ L_1^{-1}\colon \mathbb{R}^n \to \mathbb{R}^m$, we get

$$\begin{aligned}
\dim(\mathcal{V}) = n = \dim(\mathbb{R}^n) &= \dim\left(\text{domain}\left(L_2 \circ L \circ L_1^{-1}\right)\right) \\
&= \dim\left(\ker\left(L_2 \circ L \circ L_1^{-1}\right)\right) + \dim\left(\mathbf{range}\left(L_2 \circ L \circ L_1^{-1}\right)\right) \\
&= \dim(\ker(L)) + \dim(\mathbf{range}(L)), \text{ by parts (3) and (4) of Theorem 5.20.} \qquad \square
\end{aligned}$$

Suppose that $\mathcal{V}$ and $\mathcal{W}$ are finite dimensional vector spaces and $L\colon \mathcal{V} \to \mathcal{W}$ is a linear transformation. If $\dim(\mathcal{V}) = \dim(\mathcal{W})$, the next result, which requires the full generality of the Dimension Theorem, asserts that we need only check that L is *either* one-to-one or onto to know that L has the other property as well.

Corollary 5.21 Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces with $\dim(\mathcal{V}) = \dim(\mathcal{W})$. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then L is one-to-one if and only if L is onto.

Proof. Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces with $\dim(\mathcal{V}) = \dim(\mathcal{W})$, and let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then

L is one-to-one	$\Leftrightarrow$	$\dim(\ker(L)) = 0$	by Theorem 5.12
	$\Leftrightarrow$	$\dim(\mathcal{V}) = \dim(\text{range}(L))$	by the Dimension Theorem
	$\Leftrightarrow$	$\dim(\mathcal{W}) = \dim(\text{range}(L))$	because $\dim(\mathcal{V}) = \dim(\mathcal{W})$
	$\Leftrightarrow$	L is onto.	by Theorem 4.16

□

Example 5

Consider $L: \mathcal{P}_2 \to \mathbb{R}^3$ given by $L(\mathbf{p}) = [\mathbf{p}(0), \mathbf{p}(1), \mathbf{p}(2)]$. Now, $\dim(\mathcal{P}_2) = \dim(\mathbb{R}^3) = 3$. Hence, by Corollary 5.21, if L is either one-to-one or onto, it has the other property as well.

We will show that L is one-to-one using Theorem 5.12. If $\mathbf{p} \in \ker(L)$, then $L(\mathbf{p}) = \mathbf{0}$, and so $\mathbf{p}(0) = \mathbf{p}(1) = \mathbf{p}(2) = 0$. Hence, $\mathbf{p}$ is a polynomial of degree ≤ 2 touching the x-axis at $x = 0, x = 1$, and $x = 2$. Since the graph of $\mathbf{p}$ must be either a parabola or a line, it cannot touch the x-axis at three distinct points unless its graph is the line $y = 0$. That is, $\mathbf{p} = \mathbf{0}$ in $\mathcal{P}_2$. Therefore, $\ker(L) = \{\mathbf{0}\}$, and L is one-to-one.

Now, by Corollary 5.21, L is onto. Thus, given any 3-vector $[a, b, c]$, there is some $\mathbf{p} \in \mathcal{P}_2$ such that $\mathbf{p}(0) = a$, $\mathbf{p}(1) = b$, and $\mathbf{p}(2) = c$. (This example is generalized further in Exercise 21.)

■

So far, we have proved many important results concerning the concepts of one-to-one, onto, and isomorphism. For convenience, these and other useful properties from the exercises are summarized in Table 5.2.

New Vocabulary

inverse of a linear transformation
invertible linear transformation
isomorphic vector spaces
isomorphism

Highlights

- A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is invertible if and only if there is a function $M: \mathcal{W} \to \mathcal{V}$ such that $L \circ M$ and $M \circ L$ are the identity linear operators on $\mathcal{W}$ and $\mathcal{V}$, respectively.
- If a linear transformation has an inverse, its inverse is also a linear transformation.
- An isomorphism is a linear transformation that is both one-to-one and onto.

Table 5.2 Conditions on linear transformations that are one-to-one, onto, or isomorphisms

Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation, and let B be a basis for $\mathcal{V}$.

L is *one-to-one*	
$\Leftrightarrow$ $\ker(L) = \{\mathbf{0}_{\mathcal{V}}\}$	Theorem 5.12
$\Leftrightarrow$ $\dim(\ker(L)) = 0$	Theorem 5.12
$\Leftrightarrow$ the image of every linearly independent set in $\mathcal{V}$ is linearly independent in $\mathcal{W}$	Theorem 5.13 and Exercise 7 in Section 5.4
L is *onto*	
$\Leftrightarrow$ $\text{range}(L) = \mathcal{W}$	Definition
$\Leftrightarrow$ $\dim(\text{range}(L)) = \dim(\mathcal{W})$	Theorem 4.16*
$\Leftrightarrow$ the image of *every* spanning set for $\mathcal{V}$ is a spanning set for $\mathcal{W}$	Theorem 5.13
$\Leftrightarrow$ the image of *some* spanning set for $\mathcal{V}$ is a spanning set for $\mathcal{W}$	Exercise 8 in Section 5.4
L is an *isomorphism*	
$\Leftrightarrow$ L is both one-to-one and onto	Definition
$\Leftrightarrow$ L is invertible (that is, $L^{-1}: \mathcal{W} \to \mathcal{V}$ exists)	Theorem 5.14
$\Leftrightarrow$ the matrix for L (with respect to *every* pair of ordered bases for $\mathcal{V}$ and $\mathcal{W}$) is nonsingular	Theorem 5.15*
$\Leftrightarrow$ the matrix for L (with respect to *some* pair of ordered bases for $\mathcal{V}$ and $\mathcal{W}$) is nonsingular	Theorem 5.15*
$\Leftrightarrow$ the images of vectors in B are distinct and $L(B)$ is a basis for $\mathcal{W}$	Exercise 14
$\Leftrightarrow$ L is one-to-one and $\dim(\mathcal{V}) = \dim(\mathcal{W})$	Corollary 5.21*
$\Leftrightarrow$ L is onto and $\dim(\mathcal{V}) = \dim(\mathcal{W})$	Corollary 5.21*
Furthermore, if $L: \mathcal{V} \to \mathcal{W}$ is an isomorphism, then	
(1) $\dim(\mathcal{V}) = \dim(\mathcal{W})$	Theorem 5.17*
(2) L^{-1} is an isomorphism from $\mathcal{W}$ to $\mathcal{V}$	Theorem 5.14
(3) for any subspace $\mathcal{Y}$ of $\mathcal{V}$, $\dim(\mathcal{Y}) = \dim(L(\mathcal{Y}))$	Exercise 16*

**True only in the finite dimensional case*

- A linear transformation is an isomorphism if and only if it is an invertible linear transformation.
- A linear transformation (involving nontrivial finite dimensional vector spaces) is an isomorphism if and only if the matrix for the linear transformation (with respect to any ordered bases) is nonsingular.
- Under an isomorphism, the image of every linearly independent subset of the domain is linearly independent.
- Under an isomorphism, the image of every spanning set for the domain spans the codomain.
- Under an isomorphism, the dimension of every subspace of the domain is equal to the dimension of its image.
- If two vector spaces $\mathcal{V}$ and $\mathcal{W}$ have the same (finite) dimension, a linear transformation $L\colon \mathcal{V} \to \mathcal{W}$ is one-to-one if and only if it is onto.
- Finite-dimensional vector spaces are isomorphic if and only if they have the same dimension.
- All n-dimensional vector spaces are isomorphic to $\mathbb{R}^n$ (and to each other).
- The Simplified Span Method and the Independence Test Method can be justified for sets of vectors in any finite dimensional vector space $\mathcal{V}$ by applying a coordinatization isomorphism from $\mathcal{V}$ to $\mathbb{R}^n$. Similarly, the Kernel Method and the Range Method can be justified for any linear transformation $L\colon \mathcal{V} \to \mathcal{W}$ where $\mathcal{V}$ is finite dimensional by applying coordinatization isomorphisms between $\mathcal{V}$ and $\mathbb{R}^n$ and between $\mathcal{W}$ and $\mathbb{R}^m$.

EXERCISES FOR SECTION 5.5

1. Each part of this exercise gives matrices for linear operators L_1 and L_2 on $\mathbb{R}^3$ with respect to the standard basis. For each part, do the following:

(i) Show that L_1 and L_2 are isomorphisms.

(ii) Find L_1^{-1} and L_2^{-1}.

(iii) Calculate $L_2 \circ L_1$ directly.

(iv) Calculate $(L_2 \circ L_1)^{-1}$ by inverting the appropriate matrix.

(v) Calculate $L_1^{-1} \circ L_2^{-1}$ directly from your answer to (ii) and verify that the answer agrees with the result you obtained in (iv).

$\star$**(a)** $L_1\colon \begin{bmatrix} 0 & -2 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $L_2\colon \begin{bmatrix} 1 & 0 & 0 \\ -2 & 0 & 1 \\ 0 & -3 & 0 \end{bmatrix}$

(b) L_1: $\begin{bmatrix} -4 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 0 \end{bmatrix}$, L_2: $\begin{bmatrix} 0 & 3 & -1 \\ 1 & 0 & 0 \\ 0 & -2 & 1 \end{bmatrix}$

★(c) L_1: $\begin{bmatrix} -9 & 2 & 1 \\ -6 & 1 & 1 \\ 5 & 0 & -2 \end{bmatrix}$, L_2: $\begin{bmatrix} -4 & 2 & 1 \\ -3 & 1 & 0 \\ -5 & 2 & 1 \end{bmatrix}$

2. Show that $L: \mathcal{M}_{mn} \to \mathcal{M}_{nm}$ given by $L(\mathbf{A}) = \mathbf{A}^T$ is an isomorphism.

3. Let $\mathbf{A}$ be a fixed nonsingular $n \times n$ matrix.

(a) Show that $L_1: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L_1(\mathbf{B}) = \mathbf{AB}$ is an isomorphism. (Hint: Be sure to show first that L_1 is a linear operator.)

(b) Show that $L_2: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L_2(\mathbf{B}) = \mathbf{ABA}^{-1}$ is an isomorphism.

4. Show that $L: \mathcal{P}_n \to \mathcal{P}_n$ given by $L(\mathbf{p}) = \mathbf{p} + \mathbf{p}'$ is an isomorphism. (Hint: First show that L is a linear operator.)

5. Let $R: \mathbb{R}^2 \to \mathbb{R}^2$ be the operator that reflects a vector through the line $y = x$; that is, $R([a,b]) = [b,a]$.

★(a) Find the matrix for R with respect to the standard basis for $\mathbb{R}^2$.

(b) Show that R is an isomorphism.

(c) Prove that $R^{-1} = R$ using the matrix from part (a).

(d) Give a geometric explanation for the result in part (c).

6. Prove that the change of basis process is essentially an isomorphism; that is, if B and C are two different finite bases for a vector space $\mathcal{V}$, with $\dim(\mathcal{V}) = n$, then the mapping L: $\mathbb{R}^n \to \mathbb{R}^n$ given by $L([\mathbf{v}]_B) = [\mathbf{v}]_C$ is an isomorphism. (Hint: First show that L is a linear operator.)

7. Let $\mathcal{V}$, $\mathcal{W}$, and $\mathcal{X}$ be vector spaces. Let $L_1: \mathcal{V} \to \mathcal{W}$ and $L_2: \mathcal{V} \to \mathcal{W}$ be linear transformations. Let $M: \mathcal{W} \to \mathcal{X}$ be an isomorphism. If $M \circ L_1 = M \circ L_2$, show that $L_1 = L_2$.

►8. Prove Theorem 5.15.

9. (a) Explain why $\mathcal{M}_{mn} \cong \mathcal{M}_{nm}$.

(b) Explain why $\mathcal{P}_{4n+3} \cong \mathcal{M}_{4,n+1}$.

(c) Explain why the subspace of upper triangular matrices in $\mathcal{M}_{nn}$ is isomorphic to $\mathbb{R}^{n(n+1)/2}$. Is the subspace still isomorphic to $\mathbb{R}^{n(n+1)/2}$ if *upper* is replaced by *lower*?

10. Let $\mathcal{V}$ be a vector space. Show that a linear operator $L: \mathcal{V} \to \mathcal{V}$ is an isomorphism if and only if $L \circ L$ is an isomorphism.

11. Let $\mathcal{V}$ be a nontrivial vector space. Suppose that $L: \mathcal{V} \to \mathcal{V}$ is a linear operator.

(a) If $L \circ L$ is the zero transformation, show that L is not an isomorphism.

(b) If $L \circ L = L$ and L is not the identity transformation, show that L is not an isomorphism.

12. Let $L: \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator with matrix $\mathbf{A}$ (using the standard basis for $\mathbb{R}^n$). Prove that L is an isomorphism if and only if the columns of $\mathbf{A}$ are linearly independent.

⋆13. **(a)** Suppose that $L: \mathbb{R}^6 \to \mathcal{P}_5$ is a linear transformation and that L is not onto. Is L one-to-one? Why or why not?

(b) Suppose that $L: \mathcal{M}_{22} \to \mathcal{P}_3$ is a linear transformation and that L is not one-to-one. Is L onto? Why or why not?

14. Let $L: \mathcal{V} \to \mathcal{W}$ be a linear transformation between vector spaces, and let B be a basis for $\mathcal{V}$.

(a) Show that if L is an isomorphism, then $L(B)$ is a basis for $\mathcal{W}$.

(b) Prove that if $L(B)$ is a basis for $\mathcal{W}$, and the images of vectors in B are distinct, then L is an isomorphism. (Hint: Use Exercise 8(c) in Section 5.4 to show L is onto. Then show $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$ using a proof by contradiction.)

(c) Define $T: \mathbb{R}^3 \to \mathbb{R}^2$ by $T(\mathbf{X}) = \begin{bmatrix} 3 & 5 & 3 \\ 1 & 2 & 1 \end{bmatrix} \mathbf{X}$, and let B be the standard basis in $\mathbb{R}^3$. Show that $T(B)$ is a basis for $\mathbb{R}^2$, but T is not an isomorphism.

(d) Explain why part (c) does not provide a counterexample to part (b).

15. Let $L: \mathcal{V} \to \mathcal{W}$ be an isomorphism between finite dimensional vector spaces, and let B be a basis for $\mathcal{V}$. Show that for all $\mathbf{v} \in \mathcal{V}$, $[\mathbf{v}]_B = [L(\mathbf{v})]_{L(B)}$. (Hint: Use the fact from Exercise 14(a) that $L(B)$ is a basis for $\mathcal{W}$.)

►16. Let $L: \mathcal{V} \to \mathcal{W}$ be an isomorphism, with $\mathcal{V}$ finite dimensional. If $\mathcal{Y}$ is any subspace of $\mathcal{V}$, prove that $\dim(L(\mathcal{Y})) = \dim(\mathcal{Y})$.

17. Suppose $T: \mathcal{V} \to \mathcal{W}$ is a linear transformation, and $T_1: \mathcal{X} \to \mathcal{V}$ and $T_2: \mathcal{W} \to \mathcal{Y}$ are isomorphisms.

(a) Prove that $\ker(T_2 \circ T) = \ker(T)$.

(b) Prove that $\text{range}(T \circ T_1) = \text{range}(T)$.

►(c) Prove that $T_1(\ker(T \circ T_1)) = \ker(T)$.

(d) Show that $\dim(\ker(T)) = \dim(\ker(T \circ T_1))$. (Hint: Use part (c) and Exercise 16.)

►(e) Prove that $\text{range}(T_2 \circ T) = T_2(\text{range}(T))$.

(f) Show that $\dim(\text{range}(T)) = \dim(\text{range}(T_2 \circ T))$. (Hint: Use part (e) and Exercise 16.)

18. Suppose $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, and that $L_1: \mathcal{V} \to \mathbb{R}^n$ and $L_2: \mathcal{W} \to \mathbb{R}^m$ are isomorphisms. Let $M = L_2 \circ L \circ L_1^{-1}$.

▶**(a)** Use part (c) of Exercise 17 with $T = L_2 \circ L$ and $T_1 = L_1^{-1}$ to prove that $L_1^{-1}(\ker(M)) = \ker(L_2 \circ L)$.

▶**(b)** Use part (a) of this exercise together with part (a) of Exercise 17 to prove that $L_1^{-1}(\ker(M)) = \ker(L)$.

▶**(c)** Use part (b) of this exercise together with Exercise 16 to prove that $\dim(\ker(M)) = \dim(\ker(L))$.

(d) Use part (e) of Exercise 17 to prove that $L_2^{-1}(\text{range}(M)) = \text{range}\left(L \circ L_1^{-1}\right)$. (Hint: Let $T = L \circ L_1^{-1}$ and $T_2 = L_2$. Then apply L_2^{-1} to both sides.)

(e) Use part (d) of this exercise together with part (b) of Exercise 17 to prove that $L_2^{-1}(\text{range}(M)) = \text{range}(L)$.

(f) Use part (e) of this exercise together with Exercise 16 to prove that $\dim(\text{range}(M)) = \dim(\text{range}(L))$.

19. We show in this exercise that any isomorphism from $\mathbb{R}^2$ to $\mathbb{R}^2$ is the composition of certain types of reflections, contractions/dilations, and shears. (See Exercise 11 in Section 5.1 for the definition of a shear.) Note that if $a \neq 0$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{ad-bc}{a} \end{bmatrix} \begin{bmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{bmatrix},$$

and if $c \neq 0$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{bc-ad}{c} \end{bmatrix} \begin{bmatrix} 1 & \frac{d}{c} \\ 0 & 1 \end{bmatrix}.$$

(a) Use the given equations to show that every nonsingular 2×2 matrix can be expressed as a product of matrices, each of which is in one of the following forms:

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

(b) Show that when $k \geq 0$, multiplying either of the first two matrices in part (a) times the vector $[x, y]$ represents a contraction/dilation along the x-coordinate or the y-coordinate.

(c) Show that when $k < 0$, multiplying either of the first two matrices in part (a) times the vector $[x, y]$ represents a contraction/dilation along the

x-coordinate or the y-coordinate, followed by a reflection through one of the axes. $\left(\text{Hint: } \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -k & 0 \\ 0 & 1 \end{bmatrix}.\right)$

(d) Explain why multiplying either of the third or fourth matrices in part (a) times $[x,y]$ represents a shear.

(e) Explain why multiplying the last matrix in part (a) times $[x,y]$ represents a reflection through the line $y = x$.

(f) Using parts (a) through (e), show that any isomorphism from $\mathbb{R}^2$ to $\mathbb{R}^2$ is the composition of a finite number of the following linear operators: reflection through an axis, reflection through $y = x$, contraction/dilation of the x- or y-coordinate, shear in the x- or y-direction.

20. Express the linear transformation $L: \mathbb{R}^2 \to \mathbb{R}^2$ that rotates the plane 45° in a counterclockwise direction as a composition of the transformations described in part (f) of Exercise 19.

21. **(a)** Let x_1, x_2, x_3 be distinct real numbers. Use an argument similar to that in Example 5 to show that for any given $a, b, c \in \mathbb{R}$, there is a polynomial $\mathbf{p} \in \mathcal{P}_2$ such that $\mathbf{p}(x_1) = a$, $\mathbf{p}(x_2) = b$, and $\mathbf{p}(x_3) = c$.

(b) For each choice of $x_1, x_2, x_3, a, b, c \in \mathbb{R}$, show that the polynomial $\mathbf{p}$ from part (a) is unique.

(c) Recall from algebra that a nonzero polynomial of degree n can have at most n roots. Use this fact to prove that if $x_1, \ldots, x_{n+1} \in \mathbb{R}$, with $x_1, \ldots, x_{n+1}$ distinct, then for any given $a_1, \ldots, a_{n+1} \in \mathbb{R}$, there is a unique polynomial $\mathbf{p} \in \mathcal{P}_n$ such that $\mathbf{p}(x_1) = a_1$, $\mathbf{p}(x_2) = a_2, \ldots, \mathbf{p}(x_n) = a_n$, and $\mathbf{p}(x_{n+1}) = a_{n+1}$.

22. Define $L: \mathcal{P} \to \mathcal{P}$ by $L(\mathbf{p}(x)) = x\mathbf{p}(x)$.

(a) Show that L is one-to-one but not onto.

(b) Explain why L does not contradict Corollary 5.21.

★23. True or False:

(a) If the inverse L^{-1} of a linear transformation L exists, then L^{-1} is also a linear transformation.

(b) A linear transformation is an isomorphism if and only if it is invertible.

(c) If $L: \mathcal{V} \to \mathcal{V}$ is a linear operator, and the matrix for L with respect to the finite basis B for $\mathcal{V}$ is $\mathbf{A}_{BB}$, then L is an isomorphism if and only if $|\mathbf{A}_{BB}| = 0$.

(d) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then L is one-to-one if and only if L is onto.

(e) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $M: \mathcal{X} \to \mathcal{V}$ is an isomorphism, then $\ker(L \circ M) = \ker(L)$.

(f) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $M: \mathcal{X} \to \mathcal{V}$ is an isomorphism, then $\text{range}(L \circ M) = \text{range}(L)$.

(g) If $L: \mathcal{V} \to \mathcal{W}$ is an isomorphism and $\mathbf{w}_1, \ldots, \mathbf{w}_n \in \mathcal{W}$, then for every set of scalars $a_1, \ldots, a_n$, $L^{-1}(a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n) = a_1L^{-1}(\mathbf{w}_1) + \cdots + a_nL^{-1}(\mathbf{w}_n)$.

(h) $\mathbb{R}^{28} \cong \mathcal{P}_{27} \cong \mathcal{M}_{74}$.

(i) If $L: \mathbb{R}^6 \to \mathcal{M}_{32}$ is not one-to-one, then it is not onto.

5.6 DIAGONALIZATION OF LINEAR OPERATORS

In Section 3.4, we examined a method for diagonalizing certain square matrices. In this section, we generalize this process to diagonalize certain linear operators.

Eigenvalues, Eigenvectors, and Eigenspaces for Linear Operators

We define eigenvalues and eigenvectors for linear operators in a manner analogous to their definitions for matrices.

Definition Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator. A real number λ is said to be an **eigenvalue** of L if and only if there is a nonzero vector $\mathbf{v} \in \mathcal{V}$ such that $L(\mathbf{v}) = \lambda\mathbf{v}$. Also, any nonzero vector $\mathbf{v}$ such that $L(\mathbf{v}) = \lambda\mathbf{v}$ is said to be an **eigenvector** for L corresponding to the eigenvalue λ.

If L is a linear operator on $\mathbb{R}^n$ given by multiplication by a square matrix $\mathbf{A}$ (that is, $L(\mathbf{v}) = \mathbf{A}\mathbf{v}$), then the eigenvalues and eigenvectors for L are merely the eigenvalues and eigenvectors of the matrix $\mathbf{A}$, since $L(\mathbf{v}) = \lambda\mathbf{v}$ if and only if $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Hence, all of the results regarding eigenvalues and eigenvectors for matrices in Section 3.4 apply to this type of operator. Let us now consider an example involving a different type of linear operator.

Example 1

Consider $L: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T$. Then every nonzero $n \times n$ symmetric matrix $\mathbf{S}$ is an eigenvector for L corresponding to the eigenvalue $\lambda_1 = 2$ because $L(\mathbf{S}) = \mathbf{S} + \mathbf{S}^T = \mathbf{S} + \mathbf{S}$ (since $\mathbf{S}$ is symmetric) $= 2\mathbf{S}$. Similarly, every nonzero skew-symmetric $n \times n$ matrix $\mathbf{V}$ is an eigenvector for L corresponding to the eigenvalue $\lambda_2 = 0$ because $L(\mathbf{V}) = \mathbf{V} + \mathbf{V}^T = \mathbf{V} + (-\mathbf{V}) = \mathbf{O}_{nn} = 0\mathbf{V}$. ■

We now define an eigenspace for a linear operator.

Definition Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator on $\mathcal{V}$. Let λ be an eigenvalue for L. Then E_λ, the **eigenspace of** λ, is defined to be the set of all eigenvectors for L corresponding to λ, together with the zero vector $\mathbf{0}_\mathcal{V}$ of $\mathcal{V}$. That is, $E_\lambda = \{\mathbf{v} \in \mathcal{V} \mid L(\mathbf{v}) = \lambda\mathbf{v}\}$.

Just as the eigenspace of an $n \times n$ matrix is a subspace of $\mathbb{R}^n$ (see Theorem 4.4), the eigenspace of a linear operator $L: \mathcal{V} \to \mathcal{V}$ is a subspace of the vector space $\mathcal{V}$. This can be proved directly by showing that the eigenspace is nonempty and closed under vector addition and scalar multiplication, and then applying Theorem 4.2.

Example 2

Recall the operator $L: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ from Example 1 given by $L(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T$. We have already seen that the eigenspace E_2 for L contains all symmetric $n \times n$ matrices. In fact, these are the only elements of E_2 because

$$L(\mathbf{A}) = 2\mathbf{A} \Longrightarrow \mathbf{A} + \mathbf{A}^T = 2\mathbf{A} \Longrightarrow \mathbf{A} + \mathbf{A}^T = \mathbf{A} + \mathbf{A} \Longrightarrow \mathbf{A}^T = \mathbf{A}.$$

Hence, $E_2 = \{$symmetric $n \times n$ matrices$\}$, which we know to be a subspace of $\mathcal{M}_{nn}$ having dimension $n(n+1)/2$.

Similarly, the eigenspace $E_0 = \{$skew-symmetric $n \times n$ matrices$\}$.

The Characteristic Polynomial of a Linear Operator

Frequently, we analyze a linear operator L on a finite dimensional vector space $\mathcal{V}$ by looking at its matrix with respect to some basis for $\mathcal{V}$. In particular, to solve for the eigenvalues of L, we first find an ordered basis B for $\mathcal{V}$, and then solve for the matrix representation $\mathbf{A}$ of L with respect to B. For this matrix $\mathbf{A}$, we have $[L(\mathbf{v})]_B = \mathbf{A}[\mathbf{v}]_B$. Thus, finding the eigenvalues of $\mathbf{A}$ gives the eigenvalues of L.

Example 3

Let $L: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear operator given by $L([a,b]) = [b,a]$; that is, a reflection about the line $y = x$. We will calculate the eigenvalues for L two ways — first, using the standard basis for $\mathbb{R}^2$, and then, using a nonstandard basis.

Since $L(\mathbf{i}) = \mathbf{j}$ and $L(\mathbf{j}) = \mathbf{i}$, the matrix for L with respect to the standard basis is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \text{ Then } p_\mathbf{A}(x) = \begin{vmatrix} x & -1 \\ -1 & x \end{vmatrix} = x^2 - 1 = (x-1)(x+1).$$

Hence, the eigenvalues for $\mathbf{A}$ (and L) are $\lambda_1 = 1$ and $\lambda_2 = -1$. Solving the homogeneous system $(1\mathbf{I}_2 - \mathbf{A})\mathbf{v} = \mathbf{0}$ yields $\mathbf{v}_1 = [1,1]$ as an eigenvector corresponding to $\lambda_1 = 1$. Similarly, we obtain $\mathbf{v}_2 = [1,-1]$, for $\lambda_2 = -1$.

Notice that this result makes sense geometrically. The vector $\mathbf{v}_1$ runs parallel to the line of reflection and thus L leaves $\mathbf{v}_1$ unchanged; $L(\mathbf{v}_1) = \lambda_1\mathbf{v}_1 = \mathbf{v}_1$. On the other hand, $\mathbf{v}_2$ is perpendicular to the axis of reflection, and so L reverses its direction; $L(\mathbf{v}_2) = \lambda_2\mathbf{v}_2 = -\mathbf{v}_2$.

Now, instead of using the standard basis in $\mathbb{R}^2$, let us find the matrix representation of L with respect to $B = (\mathbf{v}_1, \mathbf{v}_2)$. Since $[L(\mathbf{v}_1)]_B = [1,0]$ and $[L(\mathbf{v}_2)]_B = [0,-1]$ (why?), the matrix for L with respect to B is

$$\mathbf{D} = \begin{array}{cc} [L(\mathbf{v}_1)]_B & [L(\mathbf{v}_2)]_B \end{array} \atop \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

a diagonal matrix with the eigenvalues for L on the main diagonal. Notice that

$$p_{\mathbf{D}}(x) = \begin{vmatrix} x-1 & 0 \\ 0 & x+1 \end{vmatrix} = (x-1)(x+1) = p_{\mathbf{A}}(x),$$

giving us (of course) the same eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$ for L. ■

Example 3 illustrates how two different matrix representations for the same linear operator (using different ordered bases) produce the same characteristic polynomial. Theorem 5.6 and Exercise 6 in Section 3.4 together show that this is true in general. Therefore, we can define the characteristic polynomial of a linear operator as follows, without concern about which particular ordered basis is used:

Definition Let L be a linear operator on a nontrivial finite dimensional vector space $\mathcal{V}$. Suppose $\mathbf{A}$ is the matrix representation of L with respect to some ordered basis for $\mathcal{V}$. Then the **characteristic polynomial of** L, $p_L(x)$, is defined to be $p_{\mathbf{A}}(x)$.

Example 4

Consider $L: \mathcal{P}_2 \to \mathcal{P}_2$ determined by $L(\mathbf{p}(x)) = x^2\mathbf{p}''(x) + (3x-2)\mathbf{p}'(x) + 5\mathbf{p}(x)$. You can check that $L(x^2) = 13x^2 - 4x$, $L(x) = 8x - 2$, and $L(1) = 5$. Thus, the matrix representation of L with respect to the standard basis $S = (x^2, x, 1)$ is

$$\mathbf{A} = \begin{bmatrix} 13 & 0 & 0 \\ -4 & 8 & 0 \\ 0 & -2 & 5 \end{bmatrix}.$$

Hence,

$$p_L(x) = p_{\mathbf{A}}(x) = \begin{vmatrix} x-13 & 0 & 0 \\ 4 & x-8 & 0 \\ 0 & 2 & x-5 \end{vmatrix} = (x-13)(x-8)(x-5),$$

since this is the determinant of a lower triangular matrix. The eigenvalues of L are the roots of $p_L(x)$, namely, $\lambda_1 = 13, \lambda_2 = 8$, and $\lambda_3 = 5$. ■

Criterion for Diagonalization

Given a linear operator L on a finite dimensional vector space $\mathcal{V}$, our goal is to find a basis B for $\mathcal{V}$ such that the matrix for L with respect to B is diagonal, as in Example 3. But, just as every square matrix cannot be diagonalized, neither can every linear operator.

Definition A linear operator L on a finite dimensional vector space $\mathcal{V}$ is **diagonalizable** if and only if the matrix representation of L with respect to some ordered basis for $\mathcal{V}$ is a diagonal matrix.

The next result indicates precisely which linear operators are diagonalizable.

Theorem 5.22 Let L be a linear operator on a nontrivial n-dimensional vector space $\mathcal{V}$. Then L is diagonalizable if and only if there is a set of n linearly independent eigenvectors for L.

Proof. Suppose that L is diagonalizable. Then there is an ordered basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ for $\mathcal{V}$ such that the matrix representation for L with respect to B is a diagonal matrix $\mathbf{D}$. Now, B is a linearly independent set. If we can show that each vector $\mathbf{v}_i$ in B, for $1 \leq i \leq n$, is an eigenvector corresponding to some eigenvalue for L, then B will be a set of n linearly independent eigenvectors for L. Now, for each $\mathbf{v}_i$, we have $[L(\mathbf{v}_i)]_B = \mathbf{D}[\mathbf{v}_i]_B = \mathbf{D}\mathbf{e}_i = d_{ii}\mathbf{e}_i = d_{ii}[\mathbf{v}_i]_B = [d_{ii}\mathbf{v}_i]_B$, where d_{ii} is the (i,i) entry of $\mathbf{D}$. Since coordinatization of vectors with respect to B is an isomorphism, we have $L(\mathbf{v}_i) = d_{ii}\mathbf{v}_i$, and so each $\mathbf{v}_i$ is an eigenvector for L corresponding to the eigenvalue d_{ii}.

Conversely, suppose that $B = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a set of n linearly independent eigenvectors for L, corresponding to the (not necessarily distinct) eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively. Since B contains $n = \dim(\mathcal{V})$ linearly independent vectors, B is a basis for $\mathcal{V}$, by part (2) of Theorem 4.13. We show that the matrix $\mathbf{A}$ for L with respect to B is, in fact, diagonal. Now, for $1 \leq i \leq n$,

$$i\text{th column of } \mathbf{A} = [L(\mathbf{w}_i)]_B = [\lambda_i \mathbf{w}_i]_B = \lambda_i[\mathbf{w}_i]_B = \lambda_i \mathbf{e}_i.$$

Thus, $\mathbf{A}$ is a diagonal matrix, and so L is diagonalizable. □

Example 5

In Example 3, $L: \mathbb{R}^2 \to \mathbb{R}^2$ was defined by $L([a,b]) = [b,a]$. In that example, we found a set of two linearly independent eigenvectors for L, namely, $\mathbf{v}_1 = [1,1]$ and $\mathbf{v}_2 = [1,-1]$. Since $\dim(\mathbb{R}^2) = 2$, Theorem 5.22 indicates that L is diagonalizable. In fact, in Example 3, we computed the matrix for L with respect to the ordered basis $(\mathbf{v}_1, \mathbf{v}_2)$ for $\mathbb{R}^2$ to be the diagonal matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. ■

Example 6

Consider the linear operator $L: \mathbb{R}^2 \to \mathbb{R}^2$ that rotates the plane counterclockwise through an angle of $\frac{\pi}{4}$. Now, every nonzero vector $\mathbf{v}$ is moved to $L(\mathbf{v})$, which is not parallel to $\mathbf{v}$, since $L(\mathbf{v})$ forms a $45°$ angle with $\mathbf{v}$. Hence, L has *no* eigenvectors, and so a set of two linearly independent eigenvectors cannot be found for L. Therefore, by Theorem 5.22, L is not diagonalizable. ■

Linear Independence of Eigenvectors

Theorem 5.22 asserts that finding enough *linearly independent* eigenvectors is crucial to the diagonalization process. The next theorem gives a condition under which a set of eigenvectors is guaranteed to be linearly independent.

Theorem 5.23 Let L be a linear operator on a vector space $\mathcal{V}$, and let $\lambda_1, \ldots, \lambda_t$ be distinct eigenvalues for L. If $\mathbf{v}_1, \ldots, \mathbf{v}_t$ are eigenvectors for L corresponding to $\lambda_1, \ldots, \lambda_t$, respectively, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_t\}$ is linearly independent. That is, eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof. We proceed by induction on t.

Base Step: Suppose that $t = 1$. Any eigenvector $\mathbf{v}_1$ for λ_1 is nonzero, so $\{\mathbf{v}_1\}$ is linearly independent.

Inductive Step: Let $\lambda_1, \ldots, \lambda_{k+1}$ be distinct eigenvalues for L, and let $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ be corresponding eigenvectors. Our inductive hypothesis is that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent. We must prove that $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is linearly independent. Suppose that $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1} = \mathbf{0}_{\mathcal{V}}$. Showing that $a_1 = a_2 = \cdots = a_k = a_{k+1} = 0$ will finish the proof. Now,

$$\begin{aligned} & L(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1}) = L(\mathbf{0}_{\mathcal{V}}) \\ \Longrightarrow\ & a_1L(\mathbf{v}_1) + \cdots + a_kL(\mathbf{v}_k) + a_{k+1}L(\mathbf{v}_{k+1}) = L(\mathbf{0}_{\mathcal{V}}) \\ \Longrightarrow\ & a_1\lambda_1\mathbf{v}_1 + \cdots + a_k\lambda_k\mathbf{v}_k + a_{k+1}\lambda_{k+1}\mathbf{v}_{k+1} = \mathbf{0}_{\mathcal{V}}. \end{aligned}$$

Multiplying both sides of the original equation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1} = \mathbf{0}_{\mathcal{V}}$ by λ_{k+1} yields

$$a_1\lambda_{k+1}\mathbf{v}_1 + \cdots + a_k\lambda_{k+1}\mathbf{v}_k + a_{k+1}\lambda_{k+1}\mathbf{v}_{k+1} = \mathbf{0}_{\mathcal{V}}.$$

Subtracting the last two equations containing λ_{k+1} gives

$$a_1(\lambda_1 - \lambda_{k+1})\mathbf{v}_1 + \cdots + a_k(\lambda_k - \lambda_{k+1})\mathbf{v}_k = \mathbf{0}_{\mathcal{V}}.$$

Hence, our inductive hypothesis implies that

$$a_1(\lambda_1 - \lambda_{k+1}) = \cdots = a_k(\lambda_k - \lambda_{k+1}) = 0.$$

Since the eigenvalues $\lambda_1, \ldots, \lambda_{k+1}$ are distinct, none of the factors $\lambda_i - \lambda_{k+1}$ in these equations can equal zero, for $1 \le i \le k$. Thus, $a_1 = a_2 = \cdots = a_k = 0$. Finally, plugging these values into the earlier equation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1} = \mathbf{0}_\mathcal{V}$ gives $a_{k+1}\mathbf{v}_{k+1} = \mathbf{0}_\mathcal{V}$. Since $\mathbf{v}_{k+1} \neq \mathbf{0}_\mathcal{V}$, we must have $a_{k+1} = 0$ as well. □

Example 7

Consider the linear operator $L: \mathbb{R}^3 \to \mathbb{R}^3$ given by $L(\mathbf{x}) = \mathbf{Ax}$, where

$$\mathbf{A} = \begin{bmatrix} 31 & -14 & -92 \\ -50 & 28 & 158 \\ 18 & -9 & -55 \end{bmatrix}.$$

It can be shown that the characteristic polynomial for $\mathbf{A}$ is $p_\mathbf{A}(x) = x^3 - 4x^2 + x + 6 = (x+1)(x-2)(x-3)$. Hence, the eigenvalues for $\mathbf{A}$ are $\lambda_1 = -1$, $\lambda_2 = 2$, and $\lambda_3 = 3$. A quick check verifies that $[2,-2,1], [10,1,3]$, and $[1,2,0]$ are eigenvectors, respectively, for the distinct eigenvalues λ_1, λ_2, and λ_3. Therefore, by Theorem 5.23, the set $B = \{[2,-2,1], [10,1,3], [1,2,0]\}$ is linearly independent (verify!). In fact, since $\dim(\mathbb{R}^3) = 3$, this set B is a basis for $\mathbb{R}^3$.

Also note that L is diagonalizable by Theorem 5.22, since there are three linearly independent eigenvectors for L and $\dim(\mathbb{R}^3) = 3$. In fact, the matrix for L with respect to B is

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

This can be verified by computing $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$, where

$$\mathbf{P} = \begin{bmatrix} 2 & 10 & 1 \\ -2 & 1 & 2 \\ 1 & 3 & 0 \end{bmatrix}$$

is the transition matrix from B-coordinates to standard coordinates, that is, the matrix whose columns are the vectors in B (see Exercise 8(b) in Section 4.7). ■

As illustrated in Example 7, Theorems 5.22 and 5.23 combine to prove the following:

Corollary 5.24 If L is a linear operator on an n-dimensional vector space and L has n distinct eigenvalues, then L is diagonalizable.

The converse to this corollary is false, since it is possible to get n linearly independent eigenvectors from fewer than n eigenvalues (see Exercise 6).

The proof of the following generalization of Theorem 5.23 is left as Exercises 15 and 16.

Theorem 5.25 Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator on a finite dimensional vector space $\mathcal{V}$, and let $B_1, B_2, \ldots, B_k$ be bases for eigenspaces $E_{\lambda_1}, \ldots, E_{\lambda_k}$ for L, where $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues for L. Then $B_i \cap B_j = \emptyset$ for $1 \le i < j \le k$, and $B_1 \cup B_2 \cup \cdots \cup B_k$ is a linearly independent subset of $\mathcal{V}$.

This theorem asserts that for a given operator on a finite dimensional vector space, the bases for distinct eigenspaces are disjoint, and the union of two or more bases from distinct eigenspaces always constitutes a linearly independent set.

Example 8

Consider the linear operator $L: \mathbb{R}^4 \to \mathbb{R}^4$ given by $L(\mathbf{x}) = \mathbf{A}\mathbf{x}$, for the matrix $\mathbf{A}$ in Example 6 of Section 3.4; namely,

$$\mathbf{A} = \begin{bmatrix} -4 & 7 & 1 & 4 \\ 6 & -16 & -3 & -9 \\ 12 & -27 & -4 & -15 \\ -18 & 43 & 7 & 24 \end{bmatrix}.$$

In that example, we showed there were precisely three eigenvalues for $\mathbf{A}$ (and hence, for L): $\lambda_1 = -1$, $\lambda_2 = 2$, and $\lambda_3 = 0$. In the row reduction of $[(-1)\mathbf{I}_4 - \mathbf{A} \,|\, \mathbf{0}]$ in that example, we found two independent variables, and so $\dim(\mathbf{E}_{\lambda_1}) = 2$. We also discovered fundamental eigenvectors $\mathbf{X}_1 = [-2, -1, 1, 0]$ and $\mathbf{X}_2 = [-1, -1, 0, 1]$ for λ_1. Therefore, $\{\mathbf{X}_1, \mathbf{X}_2\}$ is a basis for $\mathbf{E}_{\lambda_1}$. Similarly, we can verify that $\dim(\mathbf{E}_{\lambda_2}) = \dim(\mathbf{E}_{\lambda_3}) = 1$. We found a fundamental eigenvector $\mathbf{X}_3 = [1, -2, -4, 6]$ for λ_2, and a fundamental eigenvector $\mathbf{X}_4 = [1, -3, -3, 7]$ for λ_3. Thus, $\{\mathbf{X}_3\}$ is a basis for $\mathbf{E}_{\lambda_2}$, and $\{\mathbf{X}_4\}$ is a basis for $\mathbf{E}_{\lambda_3}$. Now, by Theorem 5.25, the union $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4\}$ of these bases is a linearly independent subset of $\mathbb{R}^4$. Of course, since $\dim(\mathbb{R}^4) = 4$, $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4\}$ is also a basis for $\mathbb{R}^4$. Hence, by Theorem 5.22, L is diagonalizable. ■

Method for Diagonalizing a Linear Operator

Theorem 5.25 suggests a method for diagonalizing a given linear operator $L: \mathcal{V} \to \mathcal{V}$, when possible. This method, outlined below, illustrates how to find a basis B so that the matrix for L with respect to B is diagonal. In the case where $\mathcal{V} = \mathbb{R}^n$ and the standard basis is used, we simply apply the Diagonalization Method of Section 3.4 to the matrix for L to find a basis for $\mathcal{V}$. In other cases, we first need to choose a basis C for $\mathcal{V}$. Next we find the matrix for L with respect to C, and then use the Diagonalization Method on this matrix to obtain a basis Z of eigenvectors in $\mathbb{R}^n$. Finally, the desired basis B for $\mathcal{V}$ consists of the vectors in $\mathcal{V}$ whose coordinatization with respect to C are the vectors in Z.

Method for Diagonalizing a Linear Operator (if possible) (Generalized Diagonalization Method)
Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator on an n-dimensional vector space $\mathcal{V}$.

Step 1: Find a basis C for $\mathcal{V}$ (if $\mathcal{V} = \mathbb{R}^n$, we can use the standard basis), and calculate the matrix representation $\mathbf{A}$ of L with respect to C.

Step 2: Apply the Diagonalization Method of Section 3.4 to $\mathbf{A}$ in order to obtain all of the eigenvalues $\lambda_1, \ldots, \lambda_k$ of $\mathbf{A}$ and a basis in $\mathbb{R}^n$ for each eigenspace E_{λ_i} of $\mathbf{A}$ (by solving an appropriate homogeneous system if necessary). If the union of the bases of the E_{λ_i} contains fewer than n elements, then L is not diagonalizable, and we stop. Otherwise, let $Z = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ be an ordered basis for $\mathbb{R}^n$ consisting of the union of the bases for the E_{λ_i}.

Step 3: Reverse the C-coordinatization isomorphism on the vectors in Z to obtain an ordered basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ for $\mathcal{V}$; that is, $[\mathbf{v}_i]_C = \mathbf{w}_i$.

The matrix representation for L with respect to B is the diagonal matrix $\mathbf{D}$ whose (i,i) entry d_{ii} is the eigenvalue for L corresponding to $\mathbf{v}_i$. In most practical situations, the transition matrix $\mathbf{P}$ from B- to C-coordinates is useful; $\mathbf{P}$ is the $n \times n$ matrix whose columns are $[\mathbf{v}_1]_C, \ldots, [\mathbf{v}_n]_C$ — that is, $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$. Note that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

If we have a linear operator on $\mathbb{R}^n$ and use the standard basis for C, then the C-coordinatization isomorphism in this method is merely the identity mapping. In this case, Steps 1 and 3 are a lot easier to perform, as we see in the next example.

Example 9

We use the preceding method to diagonalize the operator $L: \mathbb{R}^4 \to \mathbb{R}^4$ given by $L(\mathbf{v}) = \mathbf{A}\mathbf{v}$, where

$$\mathbf{A} = \begin{bmatrix} 5 & 0 & -8 & 8 \\ 8 & 1 & -16 & 16 \\ -4 & 0 & 9 & -8 \\ -8 & 0 & 16 & -15 \end{bmatrix}.$$

Step 1: Since $\mathcal{V} = \mathbb{R}^4$, we let C be the standard basis for $\mathbb{R}^4$. Then no additional work needs to be done here, since the matrix representation for L with respect to C is simply $\mathbf{A}$ itself.

Step 2: We apply the Diagonalization Method of Section 3.4 to $\mathbf{A}$. A lengthy computation produces the characteristic polynomial

$$p_{\mathbf{A}}(x) = x^4 - 6x^2 + 8x - 3 = (x-1)^3(x+3).$$

Thus, the eigenvalues for $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = -3$.

To obtain a basis for the eigenspace E_{λ_1}, we row reduce

$$[1\mathbf{I}_4 - \mathbf{A}\,|\,\mathbf{0}] = \left[\begin{array}{cccc|c} -4 & 0 & 8 & -8 & 0 \\ -8 & 0 & 16 & -16 & 0 \\ 4 & 0 & -8 & 8 & 0 \\ 8 & 0 & -16 & 16 & 0 \end{array}\right] \quad \text{to obtain} \quad \left[\begin{array}{cccc|c} 1 & 0 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

There are three independent variables, so $\dim(E_{\lambda_1}) = 3$. As in Section 3.4, we set each independent variable in turn to 1, while setting the others equal to 0. This yields three linearly independent fundamental eigenvectors: $\mathbf{w}_1 = [0,1,0,0], \mathbf{w}_2 = [2,0,1,0]$, and $\mathbf{w}_3 = [-2,0,0,1]$. Thus, $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is a basis for E_{λ_1}. A similar procedure yields $\dim(E_{\lambda_2}) = 1$, and a fundamental eigenvector $\mathbf{w}_4 = [1,2,-1,-2]$ for E_{λ_2}. Also, $\{\mathbf{w}_4\}$ is a basis for E_{λ_2}. Since $\dim(\mathcal{V}) = 4$ and since we obtained four fundamental eigenvectors overall from the Diagonalization Method, L is diagonalizable. We form the union $Z = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\}$ of the bases for E_{λ_1} and E_{λ_2}.

Step 3: Since C is the standard basis for $\mathbb{R}^4$ and the C-coordinatization isomorphism is the identity mapping, no additional work needs to be done here. We simply let $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$, where $\mathbf{v}_1 = \mathbf{w}_1, \mathbf{v}_2 = \mathbf{w}_2, \mathbf{v}_3 = \mathbf{w}_3$, and $\mathbf{v}_4 = \mathbf{w}_4$. That is, $B = ([0,1,0,0],[2,0,1,0],[-2,0,0,1],[1,2,-1,-2])$. B is an ordered basis for $\mathcal{V} = \mathbb{R}^4$.

Notice that the matrix representation of L with respect to B is the 4×4 diagonal matrix $\mathbf{D}$ with each d_{ii} equal to the eigenvalue for $\mathbf{v}_i$, for $1 \le i \le 4$. In particular,

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

Also, the transition matrix $\mathbf{P}$ from B-coordinates to standard coordinates is formed by using $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and $\mathbf{v}_4$ as columns. Hence,

$$\mathbf{P} = \begin{bmatrix} 0 & 2 & -2 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{bmatrix}, \quad \text{and its inverse is} \quad \mathbf{P}^{-1} = \begin{bmatrix} 2 & 1 & -4 & 4 \\ -1 & 0 & 3 & -2 \\ -2 & 0 & 4 & -3 \\ -1 & 0 & 2 & -2 \end{bmatrix}.$$

You should verify that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$.

■

In the next example, the linear operator is not originally defined as a matrix multiplication, and so Steps 1 and 3 of the process require additional work.

Example 10

Let $L\colon \mathcal{P}_3 \to \mathcal{P}_3$ be given by $L(\mathbf{p}(x)) = x\mathbf{p}'(x) + \mathbf{p}(x+1)$. We want to find an ordered basis B for $\mathcal{P}_3$ such that the matrix representation of L with respect to B is diagonal.

Step 1: Let $C = (x^3, x^2, x, 1)$, the standard basis for $\mathcal{P}_3$. We need the matrix for L with respect to C. Calculating directly, we get

$$\begin{aligned} L(x^3) &= x(3x^2) + (x+1)^3 &&= 4x^3 + 3x^2 + 3x + 1, \\ L(x^2) &= x(2x) + (x+1)^2 &&= 3x^2 + 2x + 1, \\ L(x) &= x(1) + (x+1) &&= 2x + 1, \\ \text{and } L(1) &= x(0) + 1 &&= 1. \end{aligned}$$

Thus, the matrix for L with respect to C is

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

Step 2: We now apply the Diagonalization Method of Section 3.4 to $\mathbf{A}$. The characteristic polynomial of $\mathbf{A}$ is $p_{\mathbf{A}}(x) = (x-4)(x-3)(x-2)(x-1)$, since $\mathbf{A}$ is lower triangular. Thus, the eigenvalues for $\mathbf{A}$ are $\lambda_1 = 4, \lambda_2 = 3, \lambda_3 = 2$, and $\lambda_4 = 1$. Solving for a basis for each eigenspace of $\mathbf{A}$ gives: basis for $E_{\lambda_1} = \{[6, 18, 27, 17]\}$, basis for $E_{\lambda_2} = \{[0, 2, 4, 3]\}$, basis for $E_{\lambda_3} = \{[0, 0, 1, 1]\}$, and basis for $E_{\lambda_4} = \{[0, 0, 0, 1]\}$. Since $\dim(\mathcal{P}_3) = 4$ and since we obtained four distinct eigenvectors, L is diagonalizable. The union

$$Z = \{[6, 18, 27, 17], [0, 2, 4, 3], [0, 0, 1, 1], [0, 0, 0, 1]\}$$

of these eigenspaces is a linearly independent set by Theorem 5.25, and hence, Z is a basis for $\mathbf{R}^4$.

Step 3: Reversing the C-coordinatization isomorphism on the vectors in Z yields the ordered basis $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ for $\mathcal{P}_3$, where $\mathbf{v}_1 = 6x^3 + 18x^2 + 27x + 17$, $\mathbf{v}_2 = 2x^2 + 4x + 3$, $\mathbf{v}_3 = x + 1$, and $\mathbf{v}_4 = 1$. The diagonal matrix

$$\mathbf{D} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is the matrix representation of L in B-coordinates and has the eigenvalues of L appearing on the main diagonal. Finally, the transition matrix $\mathbf{P}$ from B-coordinates to C-coordinates is

$$\mathbf{P} = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 18 & 2 & 0 & 0 \\ 27 & 4 & 1 & 0 \\ 17 & 3 & 1 & 1 \end{bmatrix}.$$

It can quickly be verified that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

■

Geometric and Algebraic Multiplicity

As we have seen, the number of eigenvectors in a basis for each eigenspace is crucial in determining whether a given linear operator is diagonalizable, and so we often need to consider the dimension of each eigenspace.

Definition Let L be a linear operator on a finite dimensional vector space, and let λ be an eigenvalue for L. Then the dimension of the eigenspace E_λ is called the **geometric multiplicity of** λ.

Example 11

In Example 9, we studied a linear operator on $\mathbb{R}^4$ having eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -3$. In that example, we found $\dim(E_{\lambda_1}) = 3$ and $\dim(E_{\lambda_2}) = 1$. Hence, the geometric multiplicity of λ_1 is 3 and the geometric multiplicity of λ_2 is 1.

We define the algebraic multiplicity of a linear operator in a manner analogous to the matrix-related definition in Section 3.4.

Definition Let L be a linear operator on a finite dimensional vector space, and let λ be an eigenvalue for L. Suppose that $(x - \lambda)^k$ is the highest power of $(x - \lambda)$ that divides $p_L(x)$. Then k is called the **algebraic multiplicity of** λ.

In Section 3.4, we suggested, but did not prove, the following relationship between the algebraic and geometric multiplicities of an eigenvalue.

Theorem 5.26 Let L be a linear operator on a finite dimensional vector space $\mathcal{V}$, and let λ be an eigenvalue for L. Then

$$1 \leq (\text{geometric multiplicity of } \lambda) \leq (\text{algebraic multiplicity of } \lambda).$$

The proof of Theorem 5.26 uses the following lemma:

Lemma 5.27 Let $\mathbf{A}$ be an $n \times n$ matrix symbolically represented by $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}$, where $\mathbf{B}$ is an $m \times m$ submatrix, $\mathbf{C}$ is an $m \times (n - m)$ submatrix, $\mathbf{O}$ is an $(n - m) \times m$ zero submatrix, and $\mathbf{D}$ is an $(n - m) \times (n - m)$ submatrix. Then, $|\mathbf{A}| = |\mathbf{B}| \cdot |\mathbf{D}|$.

Lemma 5.27 follows from Exercise 14 in Section 3.2. (We suggest you complete that exercise if you have not already done so.)

Proof. Proof of Theorem 5.26: Let $\mathcal{V}, L$, and λ be as given in the statement of the theorem, and let k represent the geometric multiplicity of λ. By definition, the eigenspace E_λ must contain at least one nonzero vector, and thus $k = \dim(E_\lambda) \geq 1$. Thus, the first inequality in the theorem is proved.

Next, choose a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for E_λ and expand it to an ordered basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$ for $\mathcal{V}$. Let $\mathbf{A}$ be the matrix representation for L with respect to B. Notice that for $1 \leq i \leq k$, the ith column of $\mathbf{A} = [L(\mathbf{v}_i)]_B = [\lambda \mathbf{v}_i]_B = \lambda[\mathbf{v}_i]_B = \lambda \mathbf{e}_i$. Thus, $\mathbf{A}$ has the form

$$\mathbf{A} = \begin{bmatrix} \lambda \mathbf{I}_k & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{bmatrix},$$

where $\mathbf{C}$ is a $k \times (n-k)$ submatrix, $\mathbf{O}$ is an $(n-k) \times k$ zero submatrix, and $\mathbf{D}$ is an $(n-k) \times (n-k)$ submatrix.

The form of $\mathbf{A}$ makes it straightforward to calculate the characteristic polynomial of L:

$$\begin{aligned} p_L(x) = p_{\mathbf{A}}(x) = |x\mathbf{I}_n - \mathbf{A}| &= \left| x\mathbf{I}_n - \begin{bmatrix} \lambda \mathbf{I}_k & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{bmatrix} \right| \\ &= \begin{vmatrix} (x-\lambda)\mathbf{I}_k & -\mathbf{C} \\ \mathbf{O} & x\mathbf{I}_{n-k} - \mathbf{D} \end{vmatrix} \\ &= |(x-\lambda)\mathbf{I}_k| \cdot |x\mathbf{I}_{n-k} - \mathbf{D}| \quad \text{by Lemma 5.27} \\ &= (x-\lambda)^k \cdot p_{\mathbf{D}}(x). \end{aligned}$$

Let l be the number of factors of $x - \lambda$ in $p_{\mathbf{D}}(x)$. (Note that $l \geq 0$, with $l = 0$ if $p_{\mathbf{D}}(\lambda) \neq 0$.) Then, altogether, $(x-\lambda)^{k+l}$ is the largest power of $x - \lambda$ that divides $p_L(x)$. Hence,

$$\text{geometric multiplicity of } \lambda = k \leq k + l = \text{algebraic multiplicity of } \lambda. \qquad \square$$

Example 12

Consider the linear operator $L: \mathbb{R}^4 \to \mathbb{R}^4$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -2 & 1 & 0 & -1 \\ 4 & 4 & 3 & 2 \\ 16 & 0 & -8 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

In Exercise 3(a), you are asked to verify that $p_L(x) = (x-3)^3(x+5)$. Thus, the eigenvalues for L are $\lambda_1 = 3$ and $\lambda_2 = -5$. Notice that the algebraic multiplicity of λ_1 is 3 and the algebraic multiplicity of λ_2 is 1.

Next we find the eigenspaces of λ_1 and λ_2 by solving appropriate homogeneous systems. Let $\mathbf{A}$ be the matrix for L. For $\lambda_1 = 3$, we solve $(3\mathbf{I}_4 - \mathbf{A})\mathbf{v} = \mathbf{0}$ by row reducing

$$\left[\begin{array}{rrrr|r} -2 & -2 & 0 & -1 & 0 \\ 2 & 2 & 0 & 1 & 0 \\ -4 & -4 & 0 & -2 & 0 \\ -16 & 0 & 8 & 8 & 0 \end{array}\right] \quad \text{to obtain} \quad \left[\begin{array}{rrrr|r} 1 & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, a basis for E_3 is $\{[1,-1,2,0],[1,-2,0,2]\}$, and so the geometric multiplicity of λ_1 is 2, which is less than its algebraic multiplicity.

In Exercise 3(b), you are asked to solve an appropriate system to show that the eigenspace for $\lambda_2 = -5$ has dimension 1, with $\{[-1,1,-2,8]\}$ being a basis for E_{-5}. Thus, the geometric multiplicity of λ_2 is 1. Hence, the geometric and algebraic multiplicities of λ_2 are actually equal. ■

The eigenvalue λ_2 in Example 12 also illustrates the principle that if the algebraic multiplicity of an eigenvalue is 1, then its geometric multiplicity must also be 1. This follows immediately from Theorem 5.26.

Multiplicities and Diagonalization

Theorem 5.26 gives us a way to use algebraic and geometric multiplicities to determine whether a linear operator is diagonalizable. Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator, with $\dim(\mathcal{V}) = n$. Then $p_L(x)$ has degree n. Therefore, the sum of the algebraic multiplicities for all eigenvalues can be at most n. Now, for L to be diagonalizable, L must have n linearly independent eigenvectors by Theorem 5.22. This can only happen if the sum of the geometric multiplicities of all eigenvalues for L equals n. Theorem 5.26 then forces the geometric multiplicity of every eigenvalue to equal its algebraic multiplicity (why?). We have therefore proven the following alternative characterization of diagonalizability:

> **Theorem 5.28** Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator with $\dim(\mathcal{V}) = n$. Then L is diagonalizable if and only if both of the following conditions hold: (1) the sum of the algebraic multiplicities over all eigenvalues of L equals n, and (2) the geometric multiplicity of each eigenvalue equals its algebraic multiplicity.

Theorem 5.28 gives another justification that the operator L on $\mathbb{R}^4$ in Example 9 is diagonalizable. The eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -3$ have algebraic multiplicities 3 and 1, respectively, and $3 + 1 = 4 = \dim(\mathbb{R}^4)$. Also, the eigenvalues respectively have geometric multiplicities 3 and 1, which equal their algebraic multiplicities. These conditions ensure L is diagonalizable, as we demonstrated in that example.

Example 13

Theorem 5.28 shows the operator on $\mathbb{R}^4$ in Example 12 is not diagonalizable because the geometric multiplicity of $\lambda_1 = 3$ is 2, while its algebraic multiplicity is 3.

Example 14

Let $L: \mathbb{R}^3 \to \mathbb{R}^3$ be a rotation about the z-axis through an angle of $\frac{\pi}{3}$. Then the matrix for L with respect to the standard basis is

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

as described in Table 5.1. Using $\mathbf{A}$, we calculate $p_L(x) = x^3 - 2x^2 + 2x - 1 = (x-1)(x^2 - x + 1)$, where the quadratic factor has no real roots. Therefore, $\lambda = 1$ is the only eigenvalue, and its algebraic multiplicity is 1. Hence, by Theorem 5.28, L is not diagonalizable because the sum of the algebraic multiplicities of its eigenvalues equals 1, which is less than $\dim(\mathbb{R}^3) = 3$.

The Cayley-Hamilton Theorem

We conclude this section with an interesting relationship between a matrix and its characteristic polynomial. If $p(x) = a_n x^n + a_{n-1}x^{n-1} \cdots + a_1 x + a_0$ is any polynomial and $\mathbf{A}$ is an $n \times n$ matrix, we define $p(\mathbf{A})$ to be the $n \times n$ matrix given by $p(\mathbf{A}) = a_n \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} \cdots + a_1 \mathbf{A} + a_0 \mathbf{I}_n$.

Theorem 5.29 (Cayley-Hamilton Theorem) Let $\mathbf{A}$ be an $n \times n$ matrix, and let $p_{\mathbf{A}}(x)$ be its characteristic polynomial. Then $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}_n$.

The Cayley-Hamilton Theorem is an important result in advanced linear algebra. We have placed its proof in Appendix A for the interested reader.

Example 15

Let $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 4 & -1 \end{bmatrix}$. Then $p_{\mathbf{A}}(x) = x^2 - 2x - 11$ (verify!). The Cayley-Hamilton Theorem states that $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}_2$. To check this, note that

$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{A}^2 - 2\mathbf{A} - 11\mathbf{I}_2 = \begin{bmatrix} 17 & 4 \\ 8 & 9 \end{bmatrix} - \begin{bmatrix} 6 & 4 \\ 8 & -2 \end{bmatrix} - \begin{bmatrix} 11 & 0 \\ 0 & 11 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

♦ **Application:** You have now covered the prerequisites for Section 8.9, "Differential Equations."

New Vocabulary

algebraic multiplicity (of an eigenvalue)
Cayley-Hamilton Theorem
characteristic polynomial (for a linear operator)
diagonalizable linear operator
eigenspace (for an eigenvalue of a linear operator)
eigenvalue of a linear operator
eigenvector of a linear operator
Generalized Diagonalization Method (for a linear operator)
geometric multiplicity (of an eigenvalue)

Highlights

- A linear operator L on a finite dimensional vector space $\mathcal{V}$ is diagonalizable if the matrix for L with respect to some ordered basis for $\mathcal{V}$ is diagonal.
- A linear operator L on an n-dimensional vector space $\mathcal{V}$ is diagonalizable if and only if n linearly independent eigenvectors exist for L.
- Eigenvectors corresponding to distinct eigenvalues are linearly independent.
- A linear operator L on an n-dimensional vector space $\mathcal{V}$ is diagonalizable if n distinct eigenvalues exist for L.
- If L is a linear operator, the union of bases for distinct eigenspaces of L is a linearly independent set.
- The Diagonalization Method of Section 3.4 applies to any matrix $\mathbf{A}$ for a linear operator on a finite dimensional vector space, and if $\mathbf{A}$ is diagonalizable, the method can be used to find the eigenvalues of $\mathbf{A}$, a basis of fundamental eigenvectors for $\mathbf{A}$, and a diagonal matrix similar to $\mathbf{A}$.
- The geometric multiplicity of an eigenvalue is the dimension of its eigenspace.
- The algebraic multiplicity of an eigenvalue λ for a linear operator L is the highest power of $(x - \lambda)$ that divides the characteristic polynomial $p_L(x)$.
- The geometric multiplicity of an eigenvalue is always less than or equal to its algebraic multiplicity.
- A linear operator L on an n-dimensional vector space is diagonalizable if and only if both of the following conditions hold: (1) the sum of all the algebraic multiplicities of all the eigenvalues of L is equal to n, and (2) the geometric multiplicity of each eigenvalue equals its algebraic multiplicity.

- If $\mathbf{A}$ is an $n \times n$ matrix with characteristic polynomial $p_{\mathbf{A}}(x)$, then $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}_n$. That is, every matrix is a "root" of its characteristic polynomial (Cayley-Hamilton Theorem).

EXERCISES FOR SECTION 5.6

1. For each of the following, let L be a linear operator on $\mathbb{R}^n$ represented by the given matrix with respect to the standard basis. Find all eigenvalues for L, and find a basis for the eigenspace corresponding to each eigenvalue. Compare the geometric and algebraic multiplicities of each eigenvalue.

★(a) $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$

(b) $\begin{bmatrix} 3 & 0 \\ 4 & 2 \end{bmatrix}$

★(c) $\begin{bmatrix} 0 & 1 & 1 \\ -1 & 4 & -1 \\ -1 & 5 & -2 \end{bmatrix}$

★(d) $\begin{bmatrix} 2 & 0 & 0 \\ 4 & -3 & -6 \\ -4 & 5 & 8 \end{bmatrix}$

(e) $\begin{bmatrix} 7 & 1 & 2 \\ -11 & -2 & -3 \\ -24 & -3 & -7 \end{bmatrix}$

(f) $\begin{bmatrix} -13 & 10 & 12 & 19 \\ 1 & 5 & 7 & -2 \\ -2 & -1 & -1 & 3 \\ -9 & 8 & 10 & 13 \end{bmatrix}$

2. Each of the following represents a linear operator L on a vector space $\mathcal{V}$. Let C be the standard basis in each case, and let $\mathbf{A}$ be the matrix representation of L with respect to C. Follow Steps 1 and 2 of the Generalized Diagonalization Method to determine whether L is diagonalizable. If L is diagonalizable, finish the method by performing Step 3. In particular, find the following:

(i) An ordered basis B for $\mathcal{V}$ consisting of eigenvectors for L

(ii) The diagonal matrix $\mathbf{D}$ that is the matrix representation of L with respect to B

(iii) The transition matrix $\mathbf{P}$ from B to C

Finally, check your work by verifying that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.

(a) $L: \mathbb{R}^4 \to \mathbb{R}^4$ given by $L([x_1, x_2, x_3, x_4]) = [x_2, x_1, x_4, x_3]$

★(b) $L: \mathcal{P}_2 \to \mathcal{P}_2$ given by $L(\mathbf{p}(x)) = (x-1)\mathbf{p}'(x)$

(c) $L: \mathcal{P}_2 \to \mathcal{P}_2$ given by $L(\mathbf{p}(x)) = x^2\mathbf{p}''(x) + \mathbf{p}'(x) - 3\mathbf{p}(x)$

★(d) $L: \mathcal{P}_2 \to \mathcal{P}_2$ given by $L(\mathbf{p}(x)) = (x-3)^2\mathbf{p}''(x) + x\mathbf{p}'(x) - 5\mathbf{p}(x)$

★(e) $L: \mathbb{R}^2 \to \mathbb{R}^2$ such that L is the counterclockwise rotation about the origin through an angle of $\frac{\pi}{3}$ radians

(f) $L: \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L(\mathbf{K}) = \mathbf{K}^T$

(g) $L: \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L(\mathbf{K}) = \mathbf{K} - \mathbf{K}^T$

$\star$**(h)** $L: \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L(\mathbf{K}) = \begin{bmatrix} -4 & 3 \\ -10 & 7 \end{bmatrix}\mathbf{K}$

3. Consider the linear operator $L: \mathbb{R}^4 \to \mathbb{R}^4$ from Example 12.

(a) Verify that $p_L(x) = (x-3)^3(x+5) = x^4 - 4x^3 - 18x^2 + 108x - 135$. (Hint: Use a cofactor expansion along the third column.)

(b) Show that $\{[-1,1,-2,8]\}$ is a basis for the eigenspace E_{-5} for L by solving an appropriate homogeneous system.

4. Let $L: \mathcal{P}_2 \to \mathcal{P}_2$ be the translation operator given by $L(\mathbf{p}(x)) = \mathbf{p}(x+a)$, for some (fixed) real number a.

$\star$**(a)** Find all eigenvalues for L when $a = 1$, and find a basis for each eigenspace.

(b) Find all eigenvalues for L when a is an arbitrary nonzero number, and find a basis for each eigenspace.

5. Let $\mathbf{A}$ be an $n \times n$ upper triangular matrix with all main diagonal entries equal. Show that $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ is a diagonal matrix.

6. Explain why Examples 8 and 9 provide counterexamples to the converse of Corollary 5.24.

$\star$**7.** **(a)** Give an example of a 3×3 upper triangular matrix having an eigenvalue λ with algebraic multiplicity 3 and geometric multiplicity 2.

(b) Give an example of a 3×3 upper triangular matrix, one of whose eigenvalues has algebraic multiplicity 2 and geometric multiplicity 2.

8. **(a)** Suppose that L is a linear operator on a nontrivial finite dimensional vector space. Prove L is an isomorphism if and only if 0 is not an eigenvalue for L.

(b) Let L be an isomorphism from a vector space to itself. Suppose that λ is an eigenvalue for L having eigenvector $\mathbf{v}$. Prove that $\mathbf{v}$ is an eigenvector for L^{-1} corresponding to the eigenvalue $1/\lambda$.

9. Let L be a linear operator on a nontrivial finite dimensional vector space $\mathcal{V}$, and let B be an ordered basis for $\mathcal{V}$. Also, let $\mathbf{A}$ be the matrix for L with respect to B. Assume that $\mathbf{A}$ is a diagonalizable matrix. Prove that there is an ordered basis C for $\mathcal{V}$ such that the matrix representation of L with respect to C is diagonal and hence that L is a diagonalizable operator.

10. Let $\mathbf{A}$ be an $n \times n$ matrix. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$ of eigenvectors for $\mathbf{A}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Show that $|\mathbf{A}| = \lambda_1\lambda_2\cdots\lambda_n$.

11. Let L be a linear operator on an n-dimensional vector space, with $\{\lambda_1, \ldots, \lambda_k\}$ equal to the set of all distinct eigenvalues for L. Show that $\sum_{i=1}^{k}$(geometric multiplicity of λ_i) $\leq n$.

12. Let L be a nontrivial linear operator on a nontrivial finite dimensional vector space $\mathcal{V}$. Show that if L is diagonalizable, then every root of $p_L(x)$ is real.

13. Let $\mathbf{A}$ and $\mathbf{B}$ be commuting $n \times n$ matrices.

(a) Show that if λ is an eigenvalue for $\mathbf{A}$ and $\mathbf{v} \in E_\lambda$ (the eigenspace for $\mathbf{A}$ associated with λ), then $\mathbf{Bv} \in E_\lambda$.

(b) Prove that if $\mathbf{A}$ has n distinct eigenvalues, then $\mathbf{B}$ is diagonalizable.

14. **(a)** Let $\mathbf{A}$ be a fixed 2×2 matrix with distinct eigenvalues λ_1 and λ_2. Show that the linear operator $L: \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L(\mathbf{K}) = \mathbf{AK}$ is diagonalizable with eigenvalues λ_1 and λ_2, each having multiplicity 2. (Hint: Use eigenvectors for $\mathbf{A}$ to help create eigenvectors for L.)

(b) Generalize part (a) as follows: Let $\mathbf{A}$ be a fixed diagonalizable $n \times n$ matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Show that the linear operator $L: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L(\mathbf{K}) = \mathbf{AK}$ is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_k$. In addition, show that, for each i, the geometric multiplicity of λ_i for L is n times the geometric multiplicity of λ_i for $\mathbf{A}$.

►**15.** Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator on a finite dimensional vector space $\mathcal{V}$. Suppose that λ_1 and λ_2 are distinct eigenvalues for L and that B_1 and B_2 are bases for the eigenspaces E_{λ_1} and E_{λ_2} for L. Prove that $B_1 \cap B_2$ is empty.

►**16.** Let $L: \mathcal{V} \to \mathcal{V}$ be a linear operator on a finite dimensional vector space $\mathcal{V}$. Suppose that $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues for L and that $B_i = \{\mathbf{v}_{i1}, \ldots, \mathbf{v}_{ik_i}\}$ is a basis for the eigenspace E_{λ_i}, for $1 \leq i \leq n$. The goal of this exercise is to show that $B = \cup_{i=1}^{n} B_i$ is linearly independent. Suppose that $\sum_{i=1}^{n} \sum_{j=1}^{k_i} a_{ij}\mathbf{v}_{ij} = \mathbf{0}$.

(a) Let $\mathbf{u}_i = \sum_{j=1}^{k_i} a_{ij}\mathbf{v}_{ij}$. Show that $\mathbf{u}_i \in E_{\lambda_i}$.

(b) Note that $\sum_{i=1}^{n} \mathbf{u}_i = \mathbf{0}$. Use Theorem 5.23 to show that $\mathbf{u}_i = \mathbf{0}$, for $1 \leq i \leq n$.

(c) Conclude that $a_{ij} = 0$, for $1 \leq i \leq n$ and $1 \leq j \leq k_i$.

(d) Explain why parts (a) through (c) prove that B is linearly independent.

17. Verify that the Cayley-Hamilton Theorem holds for the matrix in Example 7.

⋆**18.** True or False:

(a) If $L: \mathcal{V} \to \mathcal{V}$ is a linear operator and λ is an eigenvalue for L, then $E_\lambda = \{\lambda L(\mathbf{v}) \mid \mathbf{v} \in \mathcal{V}\}$.

(b) If L is a linear operator on a finite dimensional vector space $\mathcal{V}$ and $\mathbf{A}$ is a matrix for L with respect to some ordered basis for $\mathcal{V}$, then $p_L(x) = p_{\mathbf{A}}(x)$.

(c) If $\dim(\mathcal{V}) = 5$, a linear operator L on $\mathcal{V}$ is diagonalizable when L has five linearly independent eigenvectors.

(d) Eigenvectors for a given linear operator L are linearly independent if and only if they correspond to distinct eigenvalues of L.

(e) If L is a linear operator on a finite dimensional vector space, then the union of bases for distinct eigenspaces for L is a linearly independent set.

(f) If $L: \mathbb{R}^6 \to \mathbb{R}^6$ is a diagonalizable linear operator, then the union of bases for all the distinct eigenspaces of L is actually a basis for $\mathbb{R}^6$.

(g) If L is a diagonalizable linear operator on a finite dimensional vector space $\mathcal{V}$, the Generalized Diagonalization Method produces a basis B for $\mathcal{V}$ so that the matrix for L with respect to B is diagonal.

(h) If L is a linear operator on a finite dimensional vector space $\mathcal{V}$ and λ is an eigenvalue for L, then the algebraic multiplicity of λ is never greater than the geometric multiplicity of λ.

(i) If $\dim(\mathcal{V}) = 7$ and $L: \mathcal{V} \to \mathcal{V}$ is a linear operator, then L is diagonalizable whenever the sum of the algebraic multiplicities of all the eigenvalues equals 7.

(j) If $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix}$, then $(1\mathbf{I}_2 - \mathbf{A})(4\mathbf{I}_2 - \mathbf{A}) = \mathbf{O}_2$.

REVIEW EXERCISES FOR CHAPTER 5

1. Which of the following are linear transformations? Prove your answer is correct.

★**(a)** $f: \mathbb{R}^3 \to \mathbb{R}^3$ given by $f([x,y,z]) = [4z - y, 3x + 1, 2y + 5x]$

(b) $g: \mathcal{P}_3 \to \mathcal{M}_{32}$ given by $g(ax^3 + bx^2 + cx + d) = \begin{bmatrix} 4b - c & 3d - a \\ 2d + 3a & 4c \\ 5a + c + 2d & 2b - 3d \end{bmatrix}$

(c) $h: \mathbb{R}^2 \to \mathbb{R}^2$ given by $h([x,y]) = \left[2\sqrt{xy}, -3x^2y\right]$

2. Find the image of $[2, -3]$ under the linear transformation that rotates every vector $[x,y]$ in $\mathbb{R}^2$ counterclockwise about the origin through $\theta = 2\pi/3$. Use three decimal places in your answer.

★**3.** Let $\mathbf{B}$ and $\mathbf{C}$ be fixed $n \times n$ matrices, with $\mathbf{B}$ nonsingular. Show that the mapping $f: \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $f(\mathbf{A}) = \mathbf{CAB}^{-1}$ is a linear operator.

★**4.** Suppose $L: \mathbb{R}^3 \to \mathbb{R}^3$ is a linear operator and $L([1,0,0]) = [-3,2,4]$, $L([0,1,0]) = [5,-1,3]$, and $L([0,0,1]) = [-4,0,-2]$. Find $L([6,2,-7])$. Find $L([x,y,z])$, for any $[x,y,z] \in \mathbb{R}^3$.

5. Let $L_1: \mathcal{V} \to \mathcal{W}$ and $L_2: \mathcal{W} \to \mathcal{X}$ be linear transformations. Suppose $\mathcal{V}'$ is a subspace of $\mathcal{V}$ and $\mathcal{X}'$ is a subspace of $\mathcal{X}$.

(a) Prove that $(L_2 \circ L_1)(\mathcal{V}')$ is a subspace of $\mathcal{X}$.

★**(b)** Prove that $(L_2 \circ L_1)^{-1}(\mathcal{X}')$ is a subspace of $\mathcal{V}$.

6. For each of the following linear transformations $L: \mathcal{V} \to \mathcal{W}$, find the matrix $\mathbf{A}_{BC}$ for L with respect to the given bases B for $\mathcal{V}$ and C for $\mathcal{W}$ using the method of Theorem 5.5:

★**(a)** $L: \mathbb{R}^3 \to \mathbb{R}^2$ given by $L([x,y,z]) = [3y+2z, 4x-7y]$ with $B = ([-5,-3,-2],[3,0,1],[5,2,2])$ and $C = ([4,3],[-3,-2])$

(b) $L: \mathcal{M}_{22} \to \mathcal{P}_2$ given by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (2d+c-3a)x^2 + (4b-a)x + (2b+3d-5c)$ with $B = \left(\begin{bmatrix} 3 & 4 \\ -7 & 2 \end{bmatrix}, \begin{bmatrix} -2 & -2 \\ 3 & -2 \end{bmatrix}, \begin{bmatrix} 3 & 2 \\ -2 & 3 \end{bmatrix}, \begin{bmatrix} -6 & -3 \\ 3 & -4 \end{bmatrix}\right)$ and $C = (-6x^2+x+5, 7x^2-6x+2, 2x^2-2x+1)$

7. In each case, find the matrix $\mathbf{A}_{DE}$ for the given linear transformation $L: \mathcal{V} \to \mathcal{W}$ with respect to the given bases D and E by first finding the matrix for L with respect to the standard bases B and C for $\mathcal{V}$ and $\mathcal{W}$, respectively, and then using the method of Theorem 5.6.

(a) $L: \mathbb{R}^4 \to \mathbb{R}^3$ given by $L([a,b,c,d]) = [2a+b-3c, 3b+a-4d, c-2d]$ with $D = ([-4,7,3,0],[2,-1,-1,2],[3,-2,-2,3],[-2,2,1,1])$ and $E = ([-2,-1,2],[-6,2,-1],[3,-2,2])$

★**(b)** $L: \mathcal{P}_2 \to \mathcal{M}_{22}$ given by

$$L(ax^2+bx+c) = \begin{bmatrix} 6a-b-c & 3b+2c \\ 2a-4c & a-5b+c \end{bmatrix}$$

with $D = (-5x^2+2x+5, 3x^2-x-1, -2x^2+x+3)$ and $E = \left(\begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 2 \\ 2 & 7 \end{bmatrix}\right)$

8. Find the matrix with respect to the standard bases for the composition $L_3 \circ L_2 \circ L_1: \mathbb{R}^3 \to \mathbb{R}^3$ if L_1 is a reflection through the yz-plane, L_2 is a rotation about the z-axis of $90°$, and L_3 is a projection onto the xz-plane.

9. Suppose $L: \mathbb{R}^3 \to \mathbb{R}^3$ is the linear operator whose matrix with respect to the standard basis B for $\mathbb{R}^3$ is $\mathbf{A}_{BB} = \frac{1}{41}\begin{bmatrix} 23 & 36 & 12 \\ 36 & -31 & -24 \\ -12 & 24 & 49 \end{bmatrix}$.

★**(a)** Find $p_{\mathbf{A}_{BB}}(x)$. (Be sure to incorporate $\frac{1}{41}$ correctly into your calculations.)

(b) Find all eigenvalues for $\mathbf{A}_{BB}$ and fundamental eigenvectors for each eigenvalue.

(c) Combine the fundamental eigenvectors to form a basis C for $\mathbb{R}^3$.

(d) Find $\mathbf{A}_{CC}$. (Hint: Use $\mathbf{A}_{BB}$ and the transition matrix $\mathbf{P}$ from C to B.)

(e) Use $\mathbf{A}_{CC}$ to give a geometric description of the operator L, as was done in Example 6 of Section 5.2.

10. Consider the linear transformation $L: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 3 & 1 & -3 & 5 \\ 2 & 1 & -1 & 2 \\ 2 & 3 & 5 & -6 \\ 1 & 4 & 10 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

★**(a)** Find a basis for $\ker(L)$ and a basis for $\operatorname{range}(L)$.

(b) Verify that $\dim(\ker(L)) + \dim(\operatorname{range}(L)) = \dim(\mathbb{R}^4)$.

(c) Is $[-18, 26, -4, 2]$ in $\ker(L)$? Is $[-18, 26, -6, 2]$ in $\ker(L)$? Why or why not?

(d) Is $[8, 3, -11, -23]$ in $\operatorname{range}(L)$? Why or why not?

11. For $L: \mathcal{M}_{32} \rightarrow \mathcal{P}_3$ given by $L\left(\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\right) = (a-f)x^3 + (b-2c)x^2 + (d-3f)x$, find a basis for $\ker(L)$ and a basis for $\operatorname{range}(L)$, and verify that $\dim(\ker(L)) + \dim(\operatorname{range}(L)) = \dim(\mathcal{M}_{32})$.

★**12.** Let $L_1: \mathcal{V} \rightarrow \mathcal{W}$ and $L_2: \mathcal{W} \rightarrow \mathcal{X}$ be linear transformations.

(a) Show that $\dim(\ker(L_1)) \leq \dim(\ker(L_2 \circ L_1))$.

(b) Find linear transformations $L_1, L_2: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ for which $\dim(\ker(L_1)) < \dim(\ker(L_2 \circ L_1))$.

13. Let $\mathbf{A}$ be a fixed $m \times n$ matrix, and let $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $M: \mathbb{R}^m \rightarrow \mathbb{R}^n$ be given by $L(\mathbf{X}) = \mathbf{AX}$ and $M(\mathbf{Y}) = \mathbf{A}^T\mathbf{Y}$.

(a) Prove that $\dim(\ker(L)) - \dim(\ker(M)) = n - m$.

(b) Prove that if L is onto, then M is one-to-one.

(c) Is the converse to part (b) true? Prove or disprove.

14. Consider $L: \mathcal{P}_3 \rightarrow \mathcal{M}_{22}$ given by $L\left(ax^3 + bx^2 + cx + d\right) = \begin{bmatrix} a-d & 2b \\ b & c+d \end{bmatrix}$.

(a) Without using row reduction, determine whether L is one-to-one and whether L is onto.

(b) What is $\dim(\ker(L))$? What is $\dim(\operatorname{range}(L))$?

15. In each case, use row reduction to determine whether the given linear transformation L is one-to-one and whether L is onto, and find $\dim(\ker(L))$ and $\dim(\operatorname{range}(L))$.

★**(a)** $L: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ given by $L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2 & -1 & 1 \\ -11 & 3 & -3 \\ 13 & -8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(b) $L\colon \mathbb{R}^4 \to \mathcal{P}_2$ having matrix $\begin{bmatrix} 6 & 3 & 21 & 5 \\ 3 & 2 & 10 & 2 \\ 2 & -1 & 9 & 1 \end{bmatrix}$ with respect to the standard bases for $\mathbb{R}^4$ and $\mathcal{P}_2$

16. **(a)** Prove that any linear transformation from $\mathcal{P}_3$ to $\mathbb{R}^3$ is not one-to-one.

(b) Prove that any linear transformation from $\mathcal{P}_2$ to $\mathcal{M}_{22}$ is not onto.

17. Let $L\colon \mathcal{V} \to \mathcal{W}$ be a linear transformation.

(a) Suppose L is one-to-one and $L(\mathbf{v}_1) = cL(\mathbf{v}_2)$ with $c \neq 0$ for some vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$. Show that $\mathbf{v}_1 = c\mathbf{v}_2$, and explain why this result agrees with part (1) of Theorem 5.13.

(b) Suppose L is onto and $\mathbf{w} \in \mathcal{W}$. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ and suppose that $L(a\mathbf{v}_1 + b\mathbf{v}_2) \neq \mathbf{w}$ for all $a, b \in \mathbb{R}$. Prove that $\{\mathbf{v}_1, \mathbf{v}_2\}$ does not span $\mathcal{V}$. (Hint: Use part (2) of Theorem 5.13.)

18. Consider the linear operators L_1 and L_2 on $\mathbb{R}^4$ having the given matrices with respect to the standard basis:

$$L_1\colon \begin{bmatrix} 3 & 6 & 1 & 1 \\ 5 & 2 & -2 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & -1 & -2 & -1 \end{bmatrix}, \quad L_2\colon \begin{bmatrix} 9 & 8 & 5 & 4 \\ 9 & 13 & 4 & 7 \\ 5 & 9 & 2 & 5 \\ -5 & -2 & -2 & 0 \end{bmatrix}.$$

★**(a)** Show that L_1 and L_2 are isomorphisms.

★**(b)** Calculate the matrices for $L_2 \circ L_1$, L_1^{-1}, and L_2^{-1}.

(c) Verify that the matrix for $(L_2 \circ L_1)^{-1}$ agrees with the matrix for $L_1^{-1} \circ L_2^{-1}$.

19. **(a)** Show that a shear in the z-direction with factor k (see Table 5.1 in Section 5.2) is an isomorphism from $\mathbb{R}^3$ to itself.

(b) Calculate the inverse isomorphism of the shear in part (a). Describe the effect of the inverse geometrically.

20. Consider the subspace $\mathcal{W}$ of $\mathcal{M}_{nn}$ consisting of all $n \times n$ symmetric matrices, and let $\mathbf{B}$ be a fixed $n \times n$ nonsingular matrix.

(a) Prove that if $\mathbf{A} \in \mathcal{W}$, then $\mathbf{B}^T\mathbf{A}\mathbf{B} \in \mathcal{W}$.

(b) Prove that the linear operator on $\mathcal{W}$ given by $L(\mathbf{A}) = \mathbf{B}^T\mathbf{A}\mathbf{B}$ is an isomorphism. (Hint: Show either that L is one-to-one or that L is onto, and then use Corollary 5.21.)

21. Consider the subspace $\mathcal{W}$ of $\mathcal{P}_4$ consisting of all polynomials of the form $ax^4 + bx^3 + cx^2$, for some $a, b, c \in \mathbb{R}$.

★**(a)** Prove that $L\colon \mathcal{W} \to \mathcal{P}_3$ given by $L(\mathbf{p}) = \mathbf{p}' + \mathbf{p}''$ is one-to-one.

(b) Is L an isomorphism from $\mathcal{W}$ to $\mathcal{P}_3$?

(c) Find a vector in $\mathcal{P}_3$ that is not in range(L).

22. For each of the following, let L be the indicated linear operator.

(i) Find all eigenvalues for L, and a basis of fundamental eigenvectors for each eigenspace.

(ii) Compare the geometric and algebraic multiplicities of each eigenvalue, and determine whether L is diagonalizable.

(iii) If L is diagonalizable, find an ordered basis B of eigenvectors for L, a diagonal matrix $\mathbf{D}$ that is the matrix for L with respect to the basis B, and the transition matrix $\mathbf{P}$ from B to the standard basis.

⋆**(a)** $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ having matrix $\begin{bmatrix} -9 & 18 & -16 \\ 32 & -63 & 56 \\ 44 & -84 & 75 \end{bmatrix}$ with respect to the standard basis

(b) $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ having matrix $\begin{bmatrix} -1 & -3 & 3 \\ 3 & -1 & -1 \\ -1 & 1 & -3 \end{bmatrix}$ with respect to the standard basis

⋆**(c)** $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ having matrix $\begin{bmatrix} -97 & 20 & 12 \\ -300 & 63 & 36 \\ -300 & 60 & 39 \end{bmatrix}$ with respect to the standard basis

(d) $L\colon \mathcal{P}_3 \to \mathcal{P}_3$ given by $L(\mathbf{p}(x)) = (x-1)\mathbf{p}'(x) - 2\mathbf{p}(x)$

23. Show that $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by reflection through the plane determined by the linearly independent vectors $[a,b,c]$ and $[d,e,f]$ is diagonalizable, and state a diagonal matrix $\mathbf{D}$ that is similar to the matrix for L with respect to the standard basis for $\mathbb{R}^3$, as well as a basis of eigenvectors for L. (Hint: Use Exercise 8(a) in Section 3.1 to find a vector that is orthogonal to both $[a,b,c]$ and $[d,e,f]$. Then, follow the strategy outlined in the last paragraph of Example 6 in Section 5.2.)

24. Verify that the Cayley-Hamilton Theorem holds for the matrix in Example 12 of Section 5.6. (Hint: See part (a) of Exercise 3 in Section 5.6.)

⋆**25.** True or False:

(a) There is only one linear transformation $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $L(\mathbf{i}) = \mathbf{j}$ and $L(\mathbf{j}) = \mathbf{i}$.

(b) There is only one linear transformation $L\colon \mathbb{R}^3 \to \mathbb{R}^2$ such that $L(\mathbf{i}) = \mathbf{j}$ and $L(\mathbf{j}) = \mathbf{i}$.

(c) The matrix with respect to the standard basis for a clockwise rotation about the origin through an angle of 45° in $\mathbb{R}^2$ is $\left(\frac{\sqrt{2}}{2}\right)\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

(d) If $L\colon \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\mathcal{Y}$ is a subspace of $\mathcal{V}$, then $T\colon \mathcal{Y} \to \mathcal{W}$ given by $T(\mathbf{y}) = L(\mathbf{y})$ for all $\mathbf{y} \in \mathcal{Y}$ is a linear transformation.

(e) Let $\mathbf{B}$ be a fixed $m \times n$ matrix, and let $L: \mathbb{R}^n \to \mathbb{R}^m$ be given by $L(\mathbf{X}) = \mathbf{BX}$. Then $\mathbf{B}$ is the matrix for L with respect to the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$.

(f) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation between nontrivial finite dimensional vector spaces, and if $\mathbf{A}_{BC}$ and $\mathbf{A}_{DE}$ are matrices for L with respect to the bases B and D for $\mathcal{V}$ and C and E for $\mathcal{W}$, then $\mathbf{A}_{BC}$ and $\mathbf{A}_{DE}$ are similar matrices.

(g) There is a linear operator L on $\mathbb{R}^5$ such that $\ker(L) = \text{range}(L)$.

(h) If $\mathbf{A}$ is an $m \times n$ matrix and $L: \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation $L(\mathbf{X}) = \mathbf{AX}$, then $\dim(\text{range}(L)) = \dim(\text{row space of } \mathbf{A})$.

(i) If $\mathbf{A}$ is an $m \times n$ matrix and $L: \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation $L(\mathbf{X}) = \mathbf{AX}$, then $\text{range}(L) = \text{column space of } \mathbf{A}$.

(j) The Dimension Theorem shows that if $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\mathcal{V}$ is finite dimensional, then $\mathcal{W}$ is also finite dimensional.

(k) A linear transformation $L: \mathcal{V} \to \mathcal{W}$ is one-to-one if and only if $\ker(L)$ is empty.

(l) If $\mathcal{V}$ is a finite dimensional vector space, then a linear transformation $L: \mathcal{V} \to \mathcal{W}$ is one-to-one if and only if $\dim(\text{range}(L)) = \dim(\mathcal{V})$.

(m) Every linear transformation is either one-to-one or onto or both.

(n) If $\mathcal{V}$ is a finite dimensional vector space and $L: \mathcal{V} \to \mathcal{W}$ is an onto linear transformation, then $\mathcal{W}$ is finite dimensional.

(o) If $L: \mathcal{V} \to \mathcal{W}$ is a one-to-one linear transformation and T is a linearly independent subset of $\mathcal{V}$, then $L(T)$ is a linearly independent subset of $\mathcal{W}$.

(p) If $L: \mathcal{V} \to \mathcal{W}$ is a one-to-one and onto function between vector spaces, then L is a linear transformation.

(q) If $\mathcal{V}$ and $\mathcal{W}$ are nontrivial finite dimensional vector spaces, and $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation, then L is an isomorphism if and only if the matrix for L with respect to some bases for $\mathcal{V}$ and $\mathcal{W}$ is square.

(r) If $L: \mathbb{R}^3 \to \mathbb{R}^3$ is the isomorphism that reflects vectors through the plane $2x + 3y - z = 0$, then $L^{-1} = L$.

(s) Every nontrivial vector space $\mathcal{V}$ is isomorphic to $\mathbb{R}^n$ for some n.

(t) If $\mathcal{W}_1$ and $\mathcal{W}_2$ are two planes through the origin in $\mathbb{R}^3$, then there exists an isomorphism $L: \mathcal{W}_1 \to \mathcal{W}_2$.

(u) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $M: \mathcal{W} \to \mathcal{X}$ is an isomorphism, then $\ker(M \circ L) = \ker(L)$.

(v) If $L: \mathcal{V} \to \mathcal{W}$ is a linear transformation and $M: \mathcal{W} \to \mathcal{X}$ is an isomorphism, then $\text{range}(M \circ L) = \text{range}(L)$.

(w) If $\mathbf{A}$ is an $n \times n$ matrix and λ is an eigenvalue for $\mathbf{A}$, then E_λ is the kernel of the linear operator on $\mathbb{R}^n$ whose matrix with respect to the standard basis is $(\lambda\mathbf{I}_n - \mathbf{A})$.

(x) If L is a linear operator on an n-dimensional vector space $\mathcal{V}$ such that L has n distinct eigenvalues, then the algebraic multiplicity for each eigenvalue is 1.

(y) If L is a linear operator on a nontrivial finite dimensional vector space $\mathcal{V}$, x^2 is a factor of $p_L(x)$, and $\dim(E_0) = 1$, then L is not diagonalizable.

(z) If L is a linear operator on a nontrivial finite dimensional vector space $\mathcal{V}$ and $B_1, \ldots, B_k$ are bases for k different eigenspaces for L, then $B_1 \cup B_2 \cup \cdots \cup B_k$ is a basis for a subspace of $\mathcal{V}$.