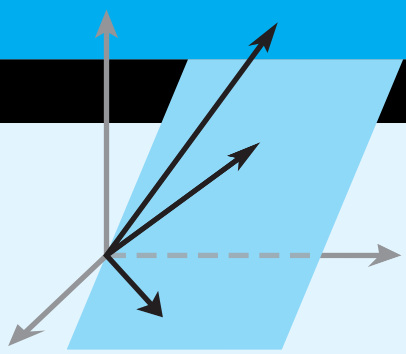


CHAPTER 8

General Linear Transformations

CHAPTER CONTENTS

INTRODUCTION

In earlier sections we studied linear transformations from R^n to R^m. In this chapter we will define and study linear transformations from a general vector space V to a general vector space W. The results we will obtain here have important applications in physics, engineering, and various branches of mathematics.

8.1 General Linear Transformations

Up to now our study of linear transformations has focused on transformations from R^n to R^m. In this section we will turn our attention to linear transformations involving general vector spaces. We will illustrate ways in which such transformations arise, and we will establish a fundamental relationship between general n-dimensional vector spaces and R^n.

Definitions and Terminology

In Section 1.8 we defined a *matrix transformation* $T_A: R^n \to R^m$ to be a mapping of the form

$$T_A(\mathbf{x}) = A\mathbf{x}$$

in which A is an $m \times n$ matrix. We subsequently established in Theorem 1.8.3 that the matrix transformations are precisely the *linear transformations* from R^n to R^m, that is, the transformations with the linearity properties

$$T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{and} \quad T(k\mathbf{u}) = kT(\mathbf{u})$$

We will use these two properties as the starting point for defining more general linear transformations.

DEFINITION 1 If $T: V \to W$ is a mapping from a vector space V to a vector space W, then T is called a ***linear transformation*** from V to W if the following two properties hold for all vectors **u** and **v** in V and for all scalars k:

(i) $T(k\mathbf{u}) = kT(\mathbf{u})$ **[Homogeneity property]**

(ii) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ **[Additivity property]**

In the special case where $V = W$, the linear transformation T is called a ***linear operator*** on the vector space V.

The homogeneity and additivity properties of a linear transformation $T: V \to W$ can be used in combination to show that if $\mathbf{v}_1$ and $\mathbf{v}_2$ are vectors in V and k_1 and k_2 are any scalars, then

$$T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2) = k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2)$$

More generally, if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are vectors in V and $k_1, k_2, \ldots, k_r$ are any scalars, then

$$T(k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r) = k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2) + \cdots + k_rT(\mathbf{v}_r) \tag{1}$$

The following theorem is an analog of parts (a) and (d) of Theorem 1.8.2.

THEOREM 8.1.1 *If $T: V \to W$ is a linear transformation, then:*

(a) $T(\mathbf{0}) = \mathbf{0}$.

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

Use the two parts of Theorem 8.1.1 to prove that

$$T(-\mathbf{v}) = -T(\mathbf{v})$$

for all $\mathbf{v}$ in V.

Proof Let $\mathbf{u}$ be any vector in V. Since $0\mathbf{u} = \mathbf{0}$, it follows from the homogeneity property in Definition 1 that

$$T(\mathbf{0}) = T(0\mathbf{u}) = 0T(\mathbf{u}) = \mathbf{0}$$

which proves (a).

We can prove part (b) by rewriting $T(\mathbf{u} - \mathbf{v})$ as

$$\begin{aligned} T(\mathbf{u} - \mathbf{v}) &= T\big(\mathbf{u} + (-1)\mathbf{v}\big) \\ &= T(\mathbf{u}) + (-1)T(\mathbf{v}) \\ &= T(\mathbf{u}) - T(\mathbf{v}) \end{aligned}$$

We leave it for you to justify each step. ◀

▶ EXAMPLE 1 Matrix Transformations

Because we have based the definition of a general linear transformation on the homogeneity and additivity properties of *matrix transformations*, it follows that every matrix transformation $T_A: R^n \to R^m$ is also a linear transformation in this more general sense with $V = R^n$ and $W = R^m$.

▶ EXAMPLE 2 The Zero Transformation

Let V and W be any two vector spaces. The mapping $T: V \to W$ such that $T(\mathbf{v}) = \mathbf{0}$ for every $\mathbf{v}$ in V is a linear transformation called the ***zero transformation***. To see that T is linear, observe that

$$T(\mathbf{u} + \mathbf{v}) = \mathbf{0}, \quad T(\mathbf{u}) = \mathbf{0}, \quad T(\mathbf{v}) = \mathbf{0}, \quad \text{and} \quad T(k\mathbf{u}) = \mathbf{0}$$

Therefore,

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{and} \quad T(k\mathbf{u}) = kT(\mathbf{u})$$

▶ EXAMPLE 3 The Identity Operator

Let V be any vector space. The mapping $I: V \to V$ defined by $I(\mathbf{v}) = \mathbf{v}$ is called the ***identity operator*** on V. We will leave it for you to verify that I is linear.

① I(kV) = kV = k · T(U)

② I(U+V) = U+V = I(U) + I(V)

▶ EXAMPLE 4 Dilation and Contraction Operators

If V is a vector space and k is any scalar, then the mapping $T: V \to V$ given by $T(\mathbf{x}) = k\mathbf{x}$ is a linear operator on V, for if c is any scalar and if $\mathbf{u}$ and $\mathbf{v}$ are any vectors in V, then

$$T(c\mathbf{u}) = k(c\mathbf{u}) = c(k\mathbf{u}) = cT(\mathbf{u})$$
$$T(\mathbf{u}+\mathbf{v}) = k(\mathbf{u}+\mathbf{v}) = k\mathbf{u} + k\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v})$$

If $0 < k < 1$, then T is called the ***contraction*** of V with factor k, and if $k > 1$, it is called the ***dilation*** of V with factor k.

▶ EXAMPLE 5 A Linear Transformation from P_n to P_{n+1}

Let $\mathbf{p} = p(x) = c_0 + c_1x + \cdots + c_nx^n$ be a polynomial in P_n, and define the transformation $T: P_n \to P_{n+1}$ by

$$T(\mathbf{p}) = T(p(x)) = xp(x) = c_0x + c_1x^2 + \cdots + c_nx^{n+1}$$

This transformation is linear because for any scalar k and any polynomials $\mathbf{p}_1$ and $\mathbf{p}_2$ in P_n we have

$$T(k\mathbf{p}) = T(kp(x)) = x(kp(x)) = k(xp(x)) = kT(\mathbf{p})$$

and

$$T(\mathbf{p}_1 + \mathbf{p}_2) = T(p_1(x) + p_2(x)) = x(p_1(x) + p_2(x))$$
$$= xp_1(x) + xp_2(x) = T(\mathbf{p}_1) + T(\mathbf{p}_2)$$

▶ EXAMPLE 6 A Linear Transformation Using the Dot Product

Let $\mathbf{v}_0$ be any fixed vector in R^n, and let $T: R^n \to R$ be the transformation

$$T(\mathbf{x}) = \langle \mathbf{x} \cdot \mathbf{v}_0 \rangle$$

that maps a vector $\mathbf{x}$ to its dot product with $\mathbf{v}_0$. This transformation is linear, for if k is any scalar, and if $\mathbf{u}$ and $\mathbf{v}$ are any vectors in R^n, then it follows from properties of the dot product in Theorem 3.2.2 that

$$T(k\mathbf{u}) = (k\mathbf{u}) \cdot \mathbf{v}_0 = k(\mathbf{u} \cdot \mathbf{v}_0) = kT(\mathbf{u})$$
$$T(\mathbf{u}+\mathbf{v}) = (\mathbf{u}+\mathbf{v}) \cdot \mathbf{v}_0 = (\mathbf{u} \cdot \mathbf{v}_0) + (\mathbf{v} \cdot \mathbf{v}_0) = T(\mathbf{u}) + T(\mathbf{v})$$

▶ EXAMPLE 7 Transformations on Matrix Spaces

Let M_{nn} be the vector space of $n \times n$ matrices. In each part determine whether the transformation is linear.

(a) $T_1(A) = A^T$ (b) $T_2(A) = \det(A)$

Solution (a) It follows from parts (b) and (d) of Theorem 1.4.8 that

$$T_1(kA) = (kA)^T = kA^T = kT_1(A)$$
$$T_1(A+B) = (A+B)^T = A^T + B^T = T_1(A) + T_1(B)$$

so T_1 is linear.

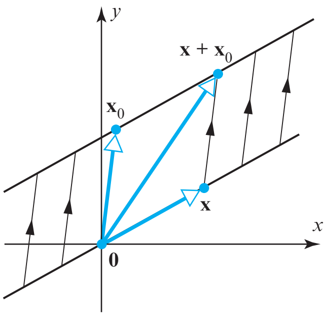
k>1

Solution (b) It follows from Formula (1) of Section 2.3 that

$$T_2(kA) = \det(kA) = k^n \det(A) = k^nT_2(A)$$ ≠ k det(A)

Thus, T_2 is not homogeneous and hence not linear if $n > 1$. Note that additivity also fails because we showed in Example 1 of Section 2.3 that $\det(A+B)$ and $\det(A) + \det(B)$ are not generally equal.

$T_2(A+B) = \det(A+B)$
$\neq \det(A) + \det B$



▲ Figure 8.1.1 $T(\mathbf{x}) = \mathbf{x} + \mathbf{x}_0$ translates each point $\mathbf{x}$ along a line parallel to $\mathbf{x}_0$ through a distance $\|\mathbf{x}_0\|$.

▶ EXAMPLE 8 Translation Is Not Linear

Part (*a*) of Theorem 8.1.1 states that a linear transformation maps $\mathbf{0}$ to $\mathbf{0}$. This property is useful for identifying transformations that are *not* linear. For example, if $\mathbf{x}_0$ is a fixed nonzero vector in R^2, then the transformation

$$T(\mathbf{x}) = \mathbf{x} + \mathbf{x}_0$$

has the geometric effect of translating each point $\mathbf{x}$ in a direction parallel to $\mathbf{x}_0$ through a distance of $\|\mathbf{x}_0\|$ (Figure 8.1.1). This cannot be a linear transformation since $T(\mathbf{0}) = \mathbf{x}_0$, so T does not map $\mathbf{0}$ to $\mathbf{0}$.

▶ EXAMPLE 9 The Evaluation Transformation

Let V be a subspace of $F(-\infty, \infty)$, let

$$x_1, x_2, \ldots, x_n$$

be a sequence of distinct real numbers, and let $T: V \to R^n$ be the transformation

$$T(f) = \big(f(x_1), f(x_2), \ldots, f(x_n)\big) \tag{2}$$

that associates with f the n-tuple of function values at $x_1, x_2, \ldots, x_n$. We call this the ***evaluation transformation*** on V at $x_1, x_2, \ldots, x_n$. Thus, for example, if

$$x_1 = -1, \quad x_2 = 2, \quad x_3 = 4$$

and if $f(x) = x^2 - 1$, then

$$T(f) = \big(f(x_1), f(x_2), f(x_3)\big) = (0, 3, 15)$$

The evaluation transformation in (2) is linear, for if k is any scalar, and if f and g are any functions in V, then

$$\begin{aligned} T(kf) &= \big((kf)(x_1), (kf)(x_2), \ldots, (kf)(x_n)\big) \\ &= \big(kf(x_1), kf(x_2), \ldots, kf(x_n)\big) \\ &= k\big(f(x_1), f(x_2), \ldots, f(x_n)\big) = kT(f) \end{aligned}$$

and

$$\begin{aligned} T(f+g) &= \big((f+g)(x_1), (f+g)(x_2), \ldots, (f+g)(x_n)\big) \\ &= \big(f(x_1) + g(x_1), f(x_2) + g(x_2), \ldots, f(x_n) + g(x_n)\big) \\ &= \big(f(x_1), f(x_2), \ldots, f(x_n)\big) + \big(g(x_1), g(x_2), \ldots, g(x_n)\big) \\ &= T(f) + T(g) \end{aligned}$$ ◀

Finding Linear Transformations from Images of Basis Vectors

We saw in Formula (15) of Section 1.8 that if $T_A: R^n \to R^m$ is multiplication by A, and if $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard basis vectors for R^n, then A can be expressed as

$$A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \cdots \mid T(\mathbf{e}_n)]$$

It follows from this that the image of any vector $\mathbf{v} = (c_1, c_2, \ldots, c_n)$ in R^n under multiplication by A can be expressed as

$$T_A(\mathbf{v}) = c_1 T_A(\mathbf{e}_1) + c_2 T_A(\mathbf{e}_2) + \cdots + c_n T_A(\mathbf{e}_n)$$

This formula tells us that for a matrix transformation the image of any vector is expressible as a linear combination of the images of the standard basis vectors. This is a special case of the following more general result.

THEOREM 8.1.2 *Let $T: V \to W$ be a linear transformation, where V is finite-dimensional. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for V, then the image of any vector $\mathbf{v}$ in V can be expressed as*

$$T(\mathbf{v}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n) \quad (3)$$

where $c_1, c_2, \ldots, c_n$ are the coefficients required to express $\mathbf{v}$ as a linear combination of the vectors in the basis S.

Proof Express $\mathbf{v}$ as $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ and use the linearity of T. ◀

▶ EXAMPLE 10 Computing with Images of Basis Vectors

Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for R^3, where

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = (1, 1, 0), \quad \mathbf{v}_3 = (1, 0, 0)$$

Let $T: R^3 \to R^2$ be the linear transformation for which

$$T(\mathbf{v}_1) = (1, 0), \quad T(\mathbf{v}_2) = (2, -1), \quad T(\mathbf{v}_3) = (4, 3)$$

Find a formula for $T(x_1, x_2, x_3)$, and then use that formula to compute $T(2, -3, 5)$.

Solution We first need to express $\mathbf{x} = (x_1, x_2, x_3)$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. If we write

$$(x_1, x_2, x_3) = c_1(1, 1, 1) + c_2(1, 1, 0) + c_3(1, 0, 0)$$

then on equating corresponding components, we obtain

$$\begin{aligned} c_1 + c_2 + c_3 &= x_1 \\ c_1 + c_2 &= x_2 \\ c_1 &= x_3 \end{aligned}$$

which yields $c_1 = x_3$, $c_2 = x_2 - x_3$, $c_3 = x_1 - x_2$, so

$$\begin{aligned} (x_1, x_2, x_3) &= x_3(1, 1, 1) + (x_2 - x_3)(1, 1, 0) + (x_1 - x_2)(1, 0, 0) \\ &= x_3\mathbf{v}_1 + (x_2 - x_3)\mathbf{v}_2 + (x_1 - x_2)\mathbf{v}_3 \end{aligned}$$

Thus

$$\begin{aligned} T(x_1, x_2, x_3) &= x_3T(\mathbf{v}_1) + (x_2 - x_3)T(\mathbf{v}_2) + (x_1 - x_2)T(\mathbf{v}_3) \\ &= x_3(1, 0) + (x_2 - x_3)(2, -1) + (x_1 - x_2)(4, 3) \\ &= (4x_1 - 2x_2 - x_3, 3x_1 - 4x_2 + x_3) \end{aligned}$$

From this formula we obtain

$$T(2, -3, 5) = (9, 23)$$

CALCULUS REQUIRED

▶ EXAMPLE 11 A Linear Transformation from $C^1(-\infty, \infty)$ to $F(-\infty, \infty)$

Let $V = C^1(-\infty, \infty)$ be the vector space of functions with continuous first derivatives on $(-\infty, \infty)$, and let $W = F(-\infty, \infty)$ be the vector space of all real-valued functions defined on $(-\infty, \infty)$. Let $D: V \to W$ be the transformation that maps a function $\mathbf{f} = f(x)$ into its derivative—that is,

$$D(\mathbf{f}) = f'(x)$$

From the properties of differentiation, we have

$$D(\mathbf{f} + \mathbf{g}) = D(\mathbf{f}) + D(\mathbf{g}) \quad \text{and} \quad D(k\mathbf{f}) = kD(\mathbf{f})$$

Thus, D is a linear transformation.

CALCULUS REQUIRED

▶ EXAMPLE 12 An Integral Transformation

Let $V = C(-\infty, \infty)$ be the vector space of continuous functions on the interval $(-\infty, \infty)$, let $W = C^1(-\infty, \infty)$ be the vector space of functions with continuous first derivatives on $(-\infty, \infty)$, and let $J: V \to W$ be the transformation that maps a function f in V into

$$J(f) = \int_0^x f(t)\,dt$$

For example, if $f(x) = x^2$, then

$$J(f) = \int_0^x t^2\,dt = \frac{t^3}{3}\bigg]_0^x = \frac{x^3}{3}$$

The transformation $J: V \to W$ is linear, for if k is any constant, and if f and g are any functions in V, then properties of the integral imply that

$$J(kf) = \int_0^x kf(t)\,dt = k\int_0^x f(t)\,dt = kJ(f)$$

$$J(f+g) = \int_0^x (f(t)+g(t))\,dt = \int_0^x f(t)\,dt + \int_0^x g(t)\,dt = J(f) + J(g) \blacktriangleleft$$

Kernel and Range

Recall that if A is an $m \times n$ matrix, then the null space of A consists of all vectors $\mathbf{x}$ in R^n such that $A\mathbf{x} = \mathbf{0}$, and by Theorem 4.7.1 the column space of A consists of all vectors $\mathbf{b}$ in R^m for which there is at least one vector $\mathbf{x}$ in R^n such that $A\mathbf{x} = \mathbf{b}$. From the viewpoint of matrix transformations, the null space of A consists of all vectors in R^n that multiplication by A maps into $\mathbf{0}$, and the column space of A consists of all vectors in R^m that are images of at least one vector in R^n under multiplication by A. The following definition extends these ideas to general linear transformations.

DEFINITION 2 If $T: V \to W$ is a linear transformation, then the set of vectors in V that T maps into $\mathbf{0}$ is called the ***kernel*** of T and is denoted by $\ker(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the ***range*** of T and is denoted by $R(T)$.

▶ EXAMPLE 13 Kernel and Range of a Matrix Transformation

If $T_A: R^n \to R^m$ is multiplication by the $m \times n$ matrix A, then, as discussed above, the kernel of T_A is the null space of A, and the range of T_A is the column space of A. $\blacktriangleleft$

▶ EXAMPLE 14 Kernel and Range of the Zero Transformation

Let $T: V \to W$ be the zero transformation. Since T maps *every* vector in V into $\mathbf{0}$, it follows that $\ker(T) = V$. Moreover, since $\mathbf{0}$ is the *only* image under T of vectors in V, it follows that $R(T) = \{\mathbf{0}\}$. $\blacktriangleleft$

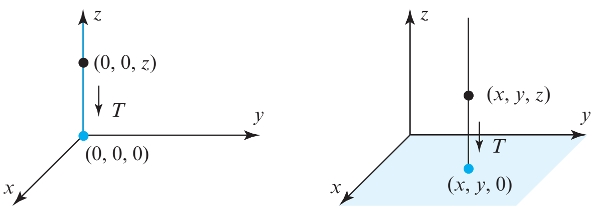
▶ EXAMPLE 15 Kernel and Range of the Identity Operator

Let $I: V \to V$ be the identity operator. Since $I(\mathbf{v}) = \mathbf{v}$ for all vectors in V, *every* vector in V is the image of some vector (namely, itself); thus $R(I) = V$. Since the *only* vector that I maps into $\mathbf{0}$ is $\mathbf{0}$, it follows that $\ker(I) = \{\mathbf{0}\}$. $\blacktriangleleft$

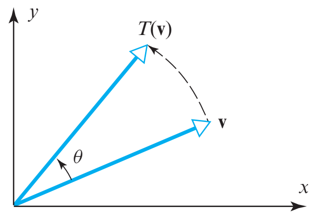
▶ EXAMPLE 16 Kernel and Range of an Orthogonal Projection

Let $T: R^3 \to R^3$ be the orthogonal projection onto the xy-plane. As illustrated in Figure 8.1.2a, the points that T maps into $\mathbf{0} = (0, 0, 0)$ are precisely those on the z-axis, so

$\ker(T)$ is the set of points of the form $(0, 0, z)$. As illustrated in Figure 8.1.2b, T maps the points in R^3 to the xy-plane, where each point in that plane is the image of each point on the vertical line above it. Thus, $R(T)$ is the set of points of the form $(x, y, 0)$.



▶ Figure 8.1.2 (a) $\ker(T)$ is the z-axis. (b) $R(T)$ is the entire xy-plane



▲ Figure 8.1.3

▶ EXAMPLE 17 Kernel and Range of a Rotation

Let $T: R^2 \to R^2$ be the linear operator that rotates each vector in the xy-plane through the angle θ (Figure 8.1.3). Since *every* vector in the xy-plane can be obtained by rotating some vector through the angle θ, it follows that $R(T) = R^2$. Moreover, the only vector that rotates into $\mathbf{0}$ is $\mathbf{0}$, so $\ker(T) = \{\mathbf{0}\}$.

CALCULUS REQUIRED

▶ EXAMPLE 18 Kernel of a Differentiation Transformation

Let $V = C^1(-\infty, \infty)$ be the vector space of functions with continuous first derivatives on $(-\infty, \infty)$, let $W = F(-\infty, \infty)$ be the vector space of all real-valued functions defined on $(-\infty, \infty)$, and let $D: V \to W$ be the differentiation transformation $D(\mathbf{f}) = f'(x)$. The kernel of D is the set of functions in V with derivative zero. From calculus, this is the set of constant functions on $(-\infty, \infty)$. ◀

Properties of Kernel and Range

In all of the preceding examples, $\ker(T)$ and $R(T)$ turned out to be *subspaces*. In Examples 14, 15, and 17 they were either the zero subspace or the entire vector space. In Example 16 the kernel was a line through the origin, and the range was a plane through the origin, both of which are subspaces of R^3. All of this is a consequence of the following general theorem.

THEOREM 8.1.3 *If $T: V \to W$ is a linear transformation, then:*

(a) *The kernel of T is a subspace of V.*

(b) *The range of T is a subspace of W.*

Proof (a) To show that $\ker(T)$ is a subspace, we must show that it contains at least one vector and is closed under addition and scalar multiplication. By part (a) of Theorem 8.1.1, the vector $\mathbf{0}$ is in $\ker(T)$, so the kernel contains at least one vector. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be vectors in $\ker(T)$, and let k be any scalar. Then

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

so $\mathbf{v}_1 + \mathbf{v}_2$ is in $\ker(T)$. Also,

$$T(k\mathbf{v}_1) = kT(\mathbf{v}_1) = k\mathbf{0} = \mathbf{0}$$

so $k\mathbf{v}_1$ is in $\ker(T)$.

Proof (b) To show that $R(T)$ is a subspace of W, we must show that it contains at least one vector and is closed under addition and scalar multiplication. However, it contains at least the zero vector of W since $T(\mathbf{0}) = \mathbf{0}$ by part (a) of Theorem 8.1.1. To prove that it is closed under addition and scalar multiplication, we must show that if $\mathbf{w}_1$ and $\mathbf{w}_2$ are vectors in $R(T)$, and if k is any scalar, then there exist vectors $\mathbf{a}$ and $\mathbf{b}$ in V for which

$$T(\mathbf{a}) = \mathbf{w}_1 + \mathbf{w}_2 \quad \text{and} \quad T(\mathbf{b}) = k\mathbf{w}_1 \tag{4}$$

But the fact that $\mathbf{w}_1$ and $\mathbf{w}_2$ are in $R(T)$ tells us there exist vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in V such that

$$T(\mathbf{v}_1) = \mathbf{w}_1 \quad \text{and} \quad T(\mathbf{v}_2) = \mathbf{w}_2$$

The following computations complete the proof by showing that the vectors $\mathbf{a} = \mathbf{v}_1 + \mathbf{v}_2$ and $\mathbf{b} = k\mathbf{v}_1$ satisfy the equations in (4):

$$T(\mathbf{a}) = T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2$$
$$T(\mathbf{b}) = T(k\mathbf{v}_1) = kT(\mathbf{v}_1) = k\mathbf{w}_1 \quad ◀$$

CALCULUS REQUIRED

▶ EXAMPLE 19 Application to Differential Equations

Differential equations of the form

$$y'' + \omega^2 y = 0 \qquad (\omega \text{ a positive constant}) \tag{5}$$

arise in the study of vibrations. The set of all solutions of this equation on the interval $(-\infty, \infty)$ is the kernel of the linear transformation $D\colon C^2(-\infty, \infty) \to C(-\infty, \infty)$, given by

$$D(y) = y'' + \omega^2 y$$

It is proved in standard textbooks on differential equations that the kernel is a two-dimensional subspace of $C^2(-\infty, \infty)$, so that if we can find two linearly independent solutions of (5), then all other solutions can be expressed as linear combinations of those two. We leave it for you to confirm by differentiating that

$$y_1 = \cos \omega x \quad \text{and} \quad y_2 = \sin \omega x$$

are solutions of (5). These functions are linearly independent since neither is a scalar multiple of the other, and thus

$$y = c_1 \cos \omega x + c_2 \sin \omega x \tag{6}$$

is a "general solution" of (5) in the sense that every choice of c_1 and c_2 produces a solution, and every solution is of this form. ◀

Rank and Nullity of Linear Transformations

In Definition 1 of Section 4.8 we defined the notions of *rank* and *nullity* for an $m \times n$ matrix, and in Theorem 4.8.2, which we called the *Dimension Theorem for Matrices*, we proved that the sum of the rank and nullity is n. We will show next that this result is a special case of a more general result about linear transformations. We start with the following definition.

DEFINITION 3 Let $T\colon V \to W$ be a linear transformation. If the range of T is finite-dimensional, then its dimension is called the ***rank of T***; and if the kernel of T is finite-dimensional, then its dimension is called the ***nullity of T***. The rank of T is denoted by $\text{rank}(T)$ and the nullity of T by $\text{nullity}(T)$.

The following theorem, whose proof is optional, generalizes Theorem 4.8.2.

THEOREM 8.1.4 Dimension Theorem for Linear Transformations

If $T: V \to W$ is a linear transformation from a finite-dimensional vector space V to a vector space W, then the range of T is finite-dimensional, and

$$\text{rank}(T) + \text{nullity}(T) = \dim(V) \tag{7}$$

In the special case where A is an $m \times n$ matrix and $T_A: R^n \to R^m$ is multiplication by A, the kernel of T_A is the null space of A, and the range of T_A is the column space of A. Thus, it follows from Theorem 8.1.4 that

$$\text{rank}(T_A) + \text{nullity}(T_A) = n$$

OPTIONAL

Proof of Theorem 8.1.4 Assume that V is n-dimensional. We must show that

$$\dim(R(T)) + \dim(\ker(T)) = n$$

We will give the proof for the case where $1 \le \dim(\ker(T)) < n$. The cases where $\dim(\ker(T)) = 0$ and $\dim(\ker(T)) = n$ are left as exercises. Assume $\dim(\ker(T)) = r$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_r$ be a basis for the kernel. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent, Theorem 4.5.5(b) states that there are $n - r$ vectors, $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$, such that the extended set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ is a basis for V. To complete the proof, we will show that the $n - r$ vectors in the set $S = \{T(\mathbf{v}_{r+1}), \ldots, T(\mathbf{v}_n)\}$ form a basis for the range of T. It will then follow that

$$\dim(R(T)) + \dim(\ker(T)) = (n - r) + r = n$$

First we show that S spans the range of T. If $\mathbf{b}$ is any vector in the range of T, then $\mathbf{b} = T(\mathbf{v})$ for some vector $\mathbf{v}$ in V. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ is a basis for V, the vector $\mathbf{v}$ can be written in the form

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} + \cdots + c_n\mathbf{v}_n$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_r$ lie in the kernel of T, we have $T(\mathbf{v}_1) = \cdots = T(\mathbf{v}_r) = \mathbf{0}$, so

$$\mathbf{b} = T(\mathbf{v}) = c_{r+1}T(\mathbf{v}_{r+1}) + \cdots + c_nT(\mathbf{v}_n)$$

Thus S spans the range of T.

Finally, we show that S is a linearly independent set and consequently forms a basis for the range of T. Suppose that some linear combination of the vectors in S is zero; that is,

$$k_{r+1}T(\mathbf{v}_{r+1}) + \cdots + k_nT(\mathbf{v}_n) = \mathbf{0} \tag{8}$$

We must show that $k_{r+1} = \cdots = k_n = 0$. Since T is linear, (8) can be rewritten as

$$T(k_{r+1}\mathbf{v}_{r+1} + \cdots + k_n\mathbf{v}_n) = \mathbf{0}$$

which says that $k_{r+1}\mathbf{v}_{r+1} + \cdots + k_n\mathbf{v}_n$ is in the kernel of T. This vector can therefore be written as a linear combination of the basis vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, say

$$k_{r+1}\mathbf{v}_{r+1} + \cdots + k_n\mathbf{v}_n = k_1\mathbf{v}_1 + \cdots + k_r\mathbf{v}_r$$

Thus,

$$k_1\mathbf{v}_1 + \cdots + k_r\mathbf{v}_r - k_{r+1}\mathbf{v}_{r+1} - \cdots - k_n\mathbf{v}_n = \mathbf{0}$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, all of the k's are zero; in particular, $k_{r+1} = \cdots = k_n = 0$, which completes the proof. ◀

Exercise Set 8.1

In Exercises 1–2, suppose that T is a mapping whose domain is the vector space M_{22}. In each part, determine whether T is a linear transformation, and if so, find its kernel.

1. (a) $T(A) = A^2$ (b) $T(A) = \text{tr}(A)$

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

2. (a) $T(A) = (A)_{11}$ (b) $T(A) = 0_{2\times 2}$

(c) $T(A) = cA$

In Exercises 3–9, determine whether the mapping T is a linear transformation, and if so, find its kernel.

3. $T: R^3 \to R$, where $T(\mathbf{u}) = \|\mathbf{u}\|$.

4. $T: R^3 \to R^3$, where $\mathbf{v}_0$ is a fixed vector in R^3 and $T(\mathbf{u}) = \mathbf{u} \times \mathbf{v}_0$.

5. $T: M_{22} \to M_{23}$, where B is a fixed 2×3 matrix and $T(A) = AB$.

6. $T: M_{22} \to R$, where

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

(b) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a^2 + b^2$

7. $T: P_2 \to P_2$, where

(a) $T(a_0 + a_1x + a_2x^2) = a_0 + a_1(x+1) + a_2(x+1)^2$

(b) $T(a_0 + a_1x + a_2x^2)$
$= (a_0 + 1) + (a_1 + 1)x + (a_2 + 1)x^2$

8. $T: F(-\infty, \infty) \to F(-\infty, \infty)$, where

(a) $T(f(x)) = 1 + f(x)$ (b) $T(f(x)) = f(x+1)$

9. $T: R^\infty \to R^\infty$, where
$T(a_0, a_1, a_2, \ldots, a_n, \ldots) = (0, a_0, a_1, a_2, \ldots, a_n, \ldots)$

10. Let $T: P_2 \to P_3$ be the linear transformation defined by $T(p(x)) = xp(x)$. Which of the following are in $\ker(T)$?

(a) x^2 (b) 0 (c) $1 + x$ (d) $-x$

11. Let $T: P_2 \to P_3$ be the linear transformation in Exercise 10. Which of the following are in $R(T)$?

(a) $x + x^2$ (b) $1 + x$ (c) $3 - x^2$ (d) $-x$

12. Let V be any vector space, and let $T: V \to V$ be defined by $T(\mathbf{v}) = 3\mathbf{v}$.

(a) What is the kernel of T?

(b) What is the range of T?

13. In each part, use the given information to find the nullity of the linear transformation T.

(a) $T: R^5 \to P_5$ has rank 3.

(b) $T: P_4 \to P_3$ has rank 1.

(c) The range of $T: M_{mn} \to R^3$ is R^3.

(d) $T: M_{22} \to M_{22}$ has rank 3.

14. In each part, use the given information to find the rank of the linear transformation T.

(a) $T: R^7 \to M_{32}$ has nullity 2.

(b) $T: P_3 \to R$ has nullity 1.

(c) The null space of $T: P_5 \to P_5$ is P_5.

(d) $T: P_n \to M_{mn}$ has nullity 3.

15. Let $T: M_{22} \to M_{22}$ be the dilation operator with factor $k = 3$.

(a) Find $T\left(\begin{bmatrix} 1 & 2 \\ -4 & 3 \end{bmatrix}\right)$.

(b) Find the rank and nullity of T.

16. Let $T: P_2 \to P_2$ be the contraction operator with factor $k = 1/4$.

(a) Find $T(1 + 4x + 8x^2)$.

(b) Find the rank and nullity of T.

17. Let $T: P_2 \to R^3$ be the evaluation transformation at the sequence of points $-1, 0, 1$. Find

(a) $T(x^2)$ (b) $\ker(T)$ (c) $R(T)$

18. Let V be the subspace of $C[0, 2\pi]$ spanned by the vectors 1, $\sin x$, and $\cos x$, and let $T: V \to R^3$ be the evaluation transformation at the sequence of points $0, \pi, 2\pi$. Find

(a) $T(1 + \sin x + \cos x)$ (b) $\ker(T)$

(c) $R(T)$

19. Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ for R^2, where $\mathbf{v}_1 = (1, 1)$ and $\mathbf{v}_2 = (1, 0)$, and let $T: R^2 \to R^2$ be the linear operator for which

$$T(\mathbf{v}_1) = (1, -2) \quad \text{and} \quad T(\mathbf{v}_2) = (-4, 1)$$

Find a formula for $T(x_1, x_2)$, and use that formula to find $T(5, -3)$.

20. Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ for R^2, where $\mathbf{v}_1 = (-2, 1)$ and $\mathbf{v}_2 = (1, 3)$, and let $T: R^2 \to R^3$ be the linear transformation such that

$$T(\mathbf{v}_1) = (-1, 2, 0) \quad \text{and} \quad T(\mathbf{v}_2) = (0, -3, 5)$$

Find a formula for $T(x_1, x_2)$, and use that formula to find $T(2, -3)$.

21. Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for R^3, where $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (1, 1, 0)$, and $\mathbf{v}_3 = (1, 0, 0)$, and let $T: R^3 \to R^3$ be the linear operator for which

$$T(\mathbf{v}_1) = (2, -1, 4), \quad T(\mathbf{v}_2) = (3, 0, 1),$$
$$T(\mathbf{v}_3) = (-1, 5, 1)$$

Find a formula for $T(x_1, x_2, x_3)$, and use that formula to find $T(2, 4, -1)$.

22. Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for R^3, where $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (2, 9, 0)$, and $\mathbf{v}_3 = (3, 3, 4)$, and let $T: R^3 \to R^2$ be the linear transformation for which

$$T(\mathbf{v}_1) = (1, 0), \quad T(\mathbf{v}_2) = (-1, 1), \quad T(\mathbf{v}_3) = (0, 1)$$

Find a formula for $T(x_1, x_2, x_3)$, and use that formula to find $T(7, 13, 7)$.

23. Let $T: P_3 \to P_2$ be the mapping defined by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = 5a_0 + a_3x^2$$

(a) Show that T is linear.

(b) Find a basis for the kernel of T.

(c) Find a basis for the range of T.

24. Let $T: P_2 \to P_2$ be the mapping defined by

$$T(a_0 + a_1x + a_2x^2) = 3a_0 + a_1x + (a_0 + a_1)x^2$$

(a) Show that T is linear.

(b) Find a basis for the kernel of T.

(c) Find a basis for the range of T.

25. (a) ***(Calculus required)*** Let $D: P_3 \to P_2$ be the differentiation transformation $D(\mathbf{p}) = p'(x)$. What is the kernel of D?

(b) ***(Calculus required)*** Let $J: P_1 \to R$ be the integration transformation $J(\mathbf{p}) = \int_{-1}^{1} p(x)\,dx$. What is the kernel of J?

26. ***(Calculus required)*** Let $V = C[a, b]$ be the vector space of continuous functions on $[a, b]$, and let $T: V \to V$ be the transformation defined by

$$T(\mathbf{f}) = 5f(x) + 3\int_a^x f(t)\,dt$$

Is T a linear operator?

27. ***(Calculus required)*** Let V be the vector space of real-valued functions with continuous derivatives of all orders on the interval $(-\infty, \infty)$, and let $W = F(-\infty, \infty)$ be the vector space of real-valued functions defined on $(-\infty, \infty)$.

(a) Find a linear transformation $T: V \to W$ whose kernel is P_3.

(b) Find a linear transformation $T: V \to W$ whose kernel is P_n.

28. For a positive integer $n > 1$, let $T: M_{nn} \to R$ be the linear transformation defined by $T(A) = \text{tr}(A)$, where A is an $n \times n$ matrix with real entries. Determine the dimension of $\ker(T)$.

29. (a) Let $T: V \to R^3$ be a linear transformation from a vector space V to R^3. Geometrically, what are the possibilities for the range of T?

(b) Let $T: R^3 \to W$ be a linear transformation from R^3 to a vector space W. Geometrically, what are the possibilities for the kernel of T?

30. In each part, determine whether the mapping $T: P_n \to P_n$ is linear.

(a) $T(p(x)) = p(x + 1)$

(b) $T(p(x)) = p(x) + 1$

31. Let $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ be vectors in a vector space V, and let $T: V \to R^3$ be a linear transformation for which

$$T(\mathbf{v}_1) = (1, -1, 2), \quad T(\mathbf{v}_2) = (0, 3, 2), \quad T(\mathbf{v}_3) = (-3, 1, 2)$$

Find $T(2\mathbf{v}_1 - 3\mathbf{v}_2 + 4\mathbf{v}_3)$.

Working with Proofs

32. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V, and let $T: V \to W$ be a linear transformation. Prove that if

$$T(\mathbf{v}_1) = T(\mathbf{v}_2) = \cdots = T(\mathbf{v}_n) = \mathbf{0}$$

then T is the zero transformation.

33. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V, and let $T: V \to V$ be a linear operator. Prove that if

$$T(\mathbf{v}_1) = \mathbf{v}_1, \quad T(\mathbf{v}_2) = \mathbf{v}_2, \ldots, \quad T(\mathbf{v}_n) = \mathbf{v}_n$$

then T is the identity transformation on V.

34. Prove: If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V and $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ are vectors in a vector space W, not necessarily distinct, then there exists a linear transformation $T: V \to W$ such that

$$T(\mathbf{v}_1) = \mathbf{w}_1, \quad T(\mathbf{v}_2) = \mathbf{w}_2, \ldots, \quad T(\mathbf{v}_n) = \mathbf{w}_n$$

True-False Exercises

TF. In parts (a)–(i) determine whether the statement is true or false, and justify your answer.

(a) If $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$ for all vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in V and all scalars c_1 and c_2, then T is a linear transformation.

(b) If $\mathbf{v}$ is a nonzero vector in V, then there is exactly one linear transformation $T: V \to W$ such that $T(-\mathbf{v}) = -T(\mathbf{v})$.

(c) There is exactly one linear transformation $T: V \to W$ for which $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u} - \mathbf{v})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V.

(d) If $\mathbf{v}_0$ is a nonzero vector in V, then the formula $T(\mathbf{v}) = \mathbf{v}_0 + \mathbf{v}$ defines a linear operator on V.

(e) The kernel of a linear transformation is a vector space.

(f) The range of a linear transformation is a vector space.

(g) If $T: P_6 \to M_{22}$ is a linear transformation, then the nullity of T is 3.

(h) The function $T: M_{22} \to R$ defined by $T(A) = \det A$ is a linear transformation.

(i) The linear transformation $T: M_{22} \to M_{22}$ defined by

$$T(A) = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} A$$

has rank 1.