

SOLUTIONS TO PRACTICE PROBLEMS

1. $\mathbf{a} \cdot \mathbf{b} = 7$, $\mathbf{a} \cdot \mathbf{a} = 5$. Hence $\frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} = \frac{7}{5}$, and $\left(\frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}}\right) \mathbf{a} = \frac{7}{5} \mathbf{a} = \begin{bmatrix} -14/5 \\ 7/5 \end{bmatrix}$.

2. Scale \mathbf{c} , multiplying by 3 to get $\mathbf{y} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$. Compute $\|\mathbf{y}\|^2 = 29$ and $\|\mathbf{y}\| = \sqrt{29}$.

The unit vector in the direction of both \mathbf{c} and \mathbf{y} is $\mathbf{u} = \frac{1}{\|\mathbf{y}\|} \mathbf{y} = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}$.

3. \mathbf{d} is orthogonal to \mathbf{c} , because

$$\mathbf{d} \cdot \mathbf{c} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

4. \mathbf{d} is orthogonal to \mathbf{u} because \mathbf{u} has the form $k\mathbf{c}$ for some k , and

$$\mathbf{d} \cdot \mathbf{u} = \mathbf{d} \cdot (k\mathbf{c}) = k(\mathbf{d} \cdot \mathbf{c}) = k(0) = 0$$

6.2 ORTHOGONAL SETS

A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in \mathbb{R}^n is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$

$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\frac{1}{2}\right) + 1(-2) + 1\left(\frac{7}{2}\right) = 0$$

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\frac{1}{2}\right) + 2(-2) + 1\left(\frac{7}{2}\right) = 0$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Fig. 1; the three line segments there are mutually perpendicular. ■

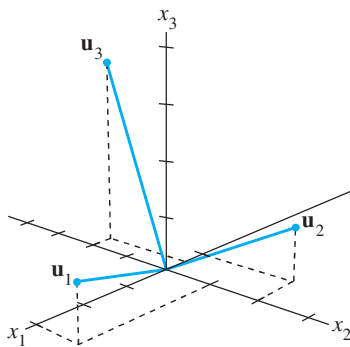


FIGURE 1

THEOREM 4

If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in \mathbb{R}^n , then S is linearly independent and hence is a basis for the subspace spanned by S .

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \dots + c_p\mathbf{u}_p$ for some scalars c_1, \dots, c_p , then

$$\begin{aligned} 0 &= \mathbf{0} \cdot \mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= (c_1\mathbf{u}_1) \cdot \mathbf{u}_1 + (c_2\mathbf{u}_2) \cdot \mathbf{u}_1 + \dots + (c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) + c_2(\mathbf{u}_2 \cdot \mathbf{u}_1) + \dots + c_p(\mathbf{u}_p \cdot \mathbf{u}_1) \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) \end{aligned}$$

because \mathbf{u}_1 is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$. Since \mathbf{u}_1 is nonzero, $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, c_2, \dots, c_p must be zero. Thus S is linearly independent. ■

DEFINITION

An **orthogonal basis** for a subspace W of \mathbb{R}^n is a basis for W that is also an orthogonal set.

The next theorem suggests why an orthogonal basis is much nicer than other bases. The weights in a linear combination can be computed easily.

THEOREM 5

Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of \mathbb{R}^n . For each \mathbf{y} in W , the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \quad (j = 1, \dots, p)$$

PROOF As in the preceding proof, the orthogonality of $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ shows that

$$\mathbf{y} \cdot \mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{u}_1 = c_1(\mathbf{u}_1 \cdot \mathbf{u}_1)$$

Since $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero, the equation above can be solved for c_1 . To find c_j for $j = 2, \dots, p$, compute $\mathbf{y} \cdot \mathbf{u}_j$ and solve for c_j . ■

EXAMPLE 2 The set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in Example 1 is an orthogonal basis for \mathbb{R}^3 .

Express the vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S .

SOLUTION Compute

$$\begin{aligned} \mathbf{y} \cdot \mathbf{u}_1 &= 11, & \mathbf{y} \cdot \mathbf{u}_2 &= -12, & \mathbf{y} \cdot \mathbf{u}_3 &= -33 \\ \mathbf{u}_1 \cdot \mathbf{u}_1 &= 11, & \mathbf{u}_2 \cdot \mathbf{u}_2 &= 6, & \mathbf{u}_3 \cdot \mathbf{u}_3 &= 33/2 \end{aligned}$$

By Theorem 5,

$$\begin{aligned} \mathbf{y} &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 + \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3} \mathbf{u}_3 \\ &= \frac{11}{11} \mathbf{u}_1 + \frac{-12}{6} \mathbf{u}_2 + \frac{-33}{33/2} \mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3 \end{aligned} \quad \blacksquare$$

Notice how easy it is to compute the weights needed to build \mathbf{y} from an orthogonal basis. If the basis were not orthogonal, it would be necessary to solve a system of linear equations in order to find the weights, as in Chapter 1.

We turn next to a construction that will become a key step in many calculations involving orthogonality, and it will lead to a geometric interpretation of Theorem 5.

An Orthogonal Projection

Given a nonzero vector \mathbf{u} in \mathbb{R}^n , consider the problem of decomposing a vector \mathbf{y} in \mathbb{R}^n into the sum of two vectors, one a multiple of \mathbf{u} and the other orthogonal to \mathbf{u} . We wish to write

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \quad (1)$$

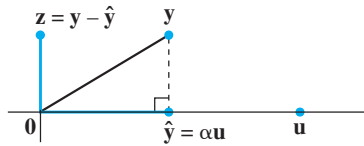


FIGURE 2

Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to \mathbf{u} .

where $\hat{\mathbf{y}} = \alpha \mathbf{u}$ for some scalar α and \mathbf{z} is some vector orthogonal to \mathbf{u} . See Fig. 2. Given any scalar α , let $\mathbf{z} = \mathbf{y} - \alpha \mathbf{u}$, so that (1) is satisfied. Then $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to \mathbf{u} if and only if

$$0 = (\mathbf{y} - \alpha \mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - (\alpha \mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \alpha (\mathbf{u} \cdot \mathbf{u})$$

That is, (1) is satisfied with \mathbf{z} orthogonal to \mathbf{u} if and only if $\alpha = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}$ and $\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$. The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of \mathbf{y} onto \mathbf{u}** , and the vector \mathbf{z} is called the **component of \mathbf{y} orthogonal to \mathbf{u}** .

If c is any nonzero scalar and if \mathbf{u} is replaced by $c\mathbf{u}$ in the definition of $\hat{\mathbf{y}}$, then the orthogonal projection of \mathbf{y} onto $c\mathbf{u}$ is exactly the same as the orthogonal projection of \mathbf{y} onto \mathbf{u} (Exercise 31). Hence this projection is determined by the *subspace* L spanned by \mathbf{u} (the line through \mathbf{u} and $\mathbf{0}$). Sometimes $\hat{\mathbf{y}}$ is denoted by $\text{proj}_L \mathbf{y}$ and is called the **orthogonal projection of \mathbf{y} onto L** . That is,

$$\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} \quad (2)$$

EXAMPLE 3 Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of \mathbf{y} onto \mathbf{u} . Then write \mathbf{y} as the sum of two orthogonal vectors, one in $\text{Span}\{\mathbf{u}\}$ and one orthogonal to \mathbf{u} .

SOLUTION Compute

$$\begin{aligned} \mathbf{y} \cdot \mathbf{u} &= \begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40 \\ \mathbf{u} \cdot \mathbf{u} &= \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20 \end{aligned}$$

The orthogonal projection of \mathbf{y} onto \mathbf{u} is

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} = \frac{40}{20} \mathbf{u} = 2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

and the component of \mathbf{y} orthogonal to \mathbf{u} is

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

The sum of these two vectors is \mathbf{y} . That is,

$$\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$\uparrow \qquad \qquad \uparrow \qquad \qquad \uparrow$
 $\mathbf{y} \qquad \qquad \hat{\mathbf{y}} \qquad \qquad (\mathbf{y} - \hat{\mathbf{y}})$

This decomposition of \mathbf{y} is illustrated in Fig. 3. *Note:* If the calculations above are correct, then $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ will be an orthogonal set. As a check, compute

$$\hat{\mathbf{y}} \cdot (\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

Since the line segment in Fig. 3 between \mathbf{y} and $\hat{\mathbf{y}}$ is perpendicular to L , by construction of $\hat{\mathbf{y}}$, the point identified with $\hat{\mathbf{y}}$ is the closest point of L to \mathbf{y} . (This can be proved from geometry. We will assume this for \mathbb{R}^2 now and prove it for \mathbb{R}^n in Section 6.3.)

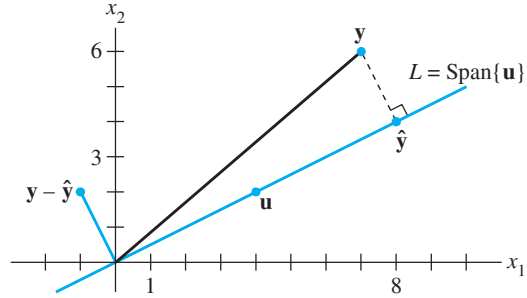


FIGURE 3 The orthogonal projection of \mathbf{y} onto a line L through the origin.

EXAMPLE 4 Find the distance in Fig. 3 from \mathbf{y} to L .

SOLUTION The distance from \mathbf{y} to L is the length of the perpendicular line segment from \mathbf{y} to the orthogonal projection $\hat{\mathbf{y}}$. This length equals the length of $\mathbf{y} - \hat{\mathbf{y}}$. Thus the distance is

$$\|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

A Geometric Interpretation of Theorem 5

The formula for the orthogonal projection $\hat{\mathbf{y}}$ in (2) has the same appearance as each of the terms in Theorem 5. Thus Theorem 5 decomposes a vector \mathbf{y} into a sum of orthogonal projections onto one-dimensional subspaces.

It is easy to visualize the case in which $W = \mathbb{R}^2 = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, with \mathbf{u}_1 and \mathbf{u}_2 orthogonal. Any \mathbf{y} in \mathbb{R}^2 can be written in the form

$$\mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 \quad (3)$$

The first term in (3) is the projection of \mathbf{y} onto the subspace spanned by \mathbf{u}_1 (the line through \mathbf{u}_1 and the origin), and the second term is the projection of \mathbf{y} onto the subspace spanned by \mathbf{u}_2 . Thus (3) expresses \mathbf{y} as the sum of its projections onto the (orthogonal) axes determined by \mathbf{u}_1 and \mathbf{u}_2 . See Fig. 4.

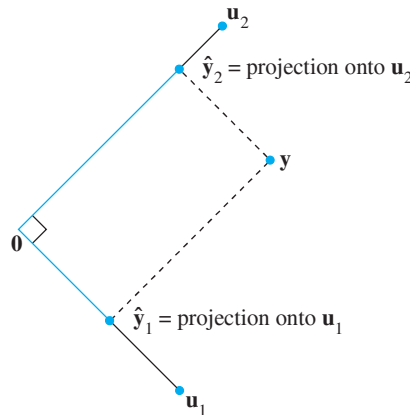


FIGURE 4 A vector decomposed into the sum of two projections.

Theorem 5 decomposes each \mathbf{y} in $\text{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ into the sum of p projections onto one-dimensional subspaces that are mutually orthogonal.

Decomposing a Force into Component Forces

The decomposition in Fig. 4 can occur in physics when some sort of force is applied to an object. Choosing an appropriate coordinate system allows the force to be represented by a vector \mathbf{y} in \mathbb{R}^2 or \mathbb{R}^3 . Often the problem involves some particular direction of interest, which is represented by another vector \mathbf{u} . For instance, if the object is moving in a straight line when the force is applied, the vector \mathbf{u} might point in the direction of movement, as in Fig. 5. A key step in the problem is to decompose the force into a component in the direction of \mathbf{u} and a component orthogonal to \mathbf{u} . The calculations would be analogous to those made in Example 3 above.

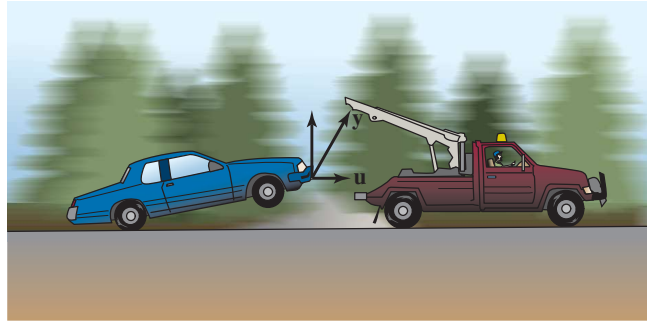


FIGURE 5

Orthonormal Sets

A set $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal basis** for W , since the set is automatically linearly independent, by Theorem 4.

The simplest example of an orthonormal set is the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for \mathbb{R}^n . Any nonempty subset of $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is orthonormal, too. Here is a more complicated example.

EXAMPLE 5 Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of \mathbb{R}^3 , where

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

SOLUTION Compute

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set. Also,

$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$$

$$\mathbf{v}_2 \cdot \mathbf{v}_2 = 1/6 + 4/6 + 1/6 = 1$$

$$\mathbf{v}_3 \cdot \mathbf{v}_3 = 1/66 + 16/66 + 49/66 = 1$$

which shows that $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 are unit vectors. Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set. Since the set is linearly independent, its three vectors form a basis for \mathbb{R}^3 . See Fig. 6. ■

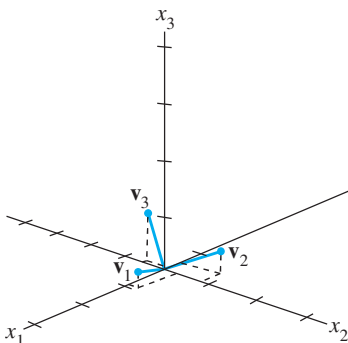


FIGURE 6

When the vectors in an orthogonal set of nonzero vectors are *normalized* to have unit length, the new vectors will still be orthogonal, and hence the new set will be an orthonormal set. See Exercise 32. It is easy to check that the vectors in Fig. 6 (Example 5) are simply the unit vectors in the directions of the vectors in Fig. 1 (Example 1).

Matrices whose columns form an orthonormal set are important in applications and in computer algorithms for matrix computations. Their main properties are given in Theorems 6 and 7.

THEOREM 6

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

PROOF To simplify notation, we suppose that U has only three columns, each a vector in \mathbb{R}^m . The proof of the general case is essentially the same. Let $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3]$ and compute

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3] = \begin{bmatrix} \mathbf{u}_1^T \mathbf{u}_1 & \mathbf{u}_1^T \mathbf{u}_2 & \mathbf{u}_1^T \mathbf{u}_3 \\ \mathbf{u}_2^T \mathbf{u}_1 & \mathbf{u}_2^T \mathbf{u}_2 & \mathbf{u}_2^T \mathbf{u}_3 \\ \mathbf{u}_3^T \mathbf{u}_1 & \mathbf{u}_3^T \mathbf{u}_2 & \mathbf{u}_3^T \mathbf{u}_3 \end{bmatrix} \quad (4)$$

The entries in the matrix at the right are inner products, using transpose notation. The columns of U are orthogonal if and only if

$$\mathbf{u}_1^T \mathbf{u}_2 = \mathbf{u}_2^T \mathbf{u}_1 = 0, \quad \mathbf{u}_1^T \mathbf{u}_3 = \mathbf{u}_3^T \mathbf{u}_1 = 0, \quad \mathbf{u}_2^T \mathbf{u}_3 = \mathbf{u}_3^T \mathbf{u}_2 = 0 \quad (5)$$

The columns of U all have unit length if and only if

$$\mathbf{u}_1^T \mathbf{u}_1 = 1, \quad \mathbf{u}_2^T \mathbf{u}_2 = 1, \quad \mathbf{u}_3^T \mathbf{u}_3 = 1 \quad (6)$$

The theorem follows immediately from (4)–(6). ■

THEOREM 7

Let U be an $m \times n$ matrix with orthonormal columns, and let \mathbf{x} and \mathbf{y} be in \mathbb{R}^n . Then

- a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
- b. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
- c. $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$

Properties (a) and (c) say that the linear mapping $\mathbf{x} \mapsto U\mathbf{x}$ preserves lengths and orthogonality. These properties are crucial for many computer algorithms. See Exercise 25 for the proof of Theorem 7.

EXAMPLE 6 Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Notice that U has orthonormal columns and

$$U^T U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Verify that $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

SOLUTION

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9 + 1 + 1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2 + 9} = \sqrt{11}$$

Theorems 6 and 7 are particularly useful when applied to *square* matrices. An **orthogonal matrix** is a square invertible matrix U such that $U^{-1} = U^T$. By Theorem 6, such a matrix has orthonormal columns.¹ It is easy to see that any *square* matrix with orthonormal columns is an orthogonal matrix. Surprisingly, such a matrix must have orthonormal *rows*, too. See Exercises 27 and 28. Orthogonal matrices will appear frequently in Chapter 7.

EXAMPLE 7 The matrix

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthogonal matrix because it is square and because its columns are orthonormal, by Example 5. Verify that the rows are orthonormal, too!

PRACTICE PROBLEMS

- Let $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal basis for \mathbb{R}^2 .
- Let \mathbf{y} and L be as in Example 3 and Fig. 3. Compute the orthogonal projection $\hat{\mathbf{y}}$ of \mathbf{y} onto L using $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ instead of the \mathbf{u} in Example 3.
- Let U and \mathbf{x} be as in Example 6, and let $\mathbf{y} = \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix}$. Verify that $U\mathbf{x} \cdot U\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$.

6.2 EXERCISES

In Exercises 1–6, determine which sets of vectors are orthogonal.

1. $\begin{bmatrix} -1 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -7 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

5. $\begin{bmatrix} 3 \\ -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 7 \\ 0 \end{bmatrix}$

6. $\begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ -3 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 5 \\ -1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -7 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

4. $\begin{bmatrix} 2 \\ -5 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}$

In Exercises 7–10, show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ or $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for \mathbb{R}^2 or \mathbb{R}^3 , respectively. Then express \mathbf{x} as a linear combination of the \mathbf{u} 's.

7. $\mathbf{u}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

¹A better name might be *orthonormal matrix*, and this term is found in some statistics texts. However, *orthogonal matrix* is the standard term in linear algebra.