Section 7.9 Orthonormal Bases and the Gram–Schmidt Process

ORTHOGONAL AND ORTHONORMAL BASES

a set of vectors in R^n is said to be *orthogonal* if each pair of distinct vectors in the set is orthogonal, and it is said to be *orthonormal* if it is orthogonal and each vector has length 1. In this section we will be concerned with ***orthogonal bases*** and ***orthonormal bases*** for subspaces of R^n. Here are some examples.

EXAMPLE 2

The Standard Basis for R^n is an Orthonormal Basis

Recall from Example 2 of Section 7.1 that the vectors

$$\mathbf{e}_1 = (1, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, \dots, 0), \dots, \quad \mathbf{e}_n = (0, 0, \dots, 1)$$

form the standard basis for R^n. This is an orthonormal basis, since these are unit vectors and $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ if $i \neq j$. ■

EXAMPLE 1 Converting an Orthogonal Basis to an Orthonormal Basis

Show that the vectors

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

form an orthogonal basis for R^3, and convert it into an orthonormal basis by normalizing each vector.

Solution We showed in Example 3 of Section 7.1 that these vectors are linearly independent, so they must form a basis for R^3 by Theorem 7.2.6. We leave it for you to confirm that this is an orthogonal basis by showing that

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0, \quad \mathbf{v}_1 \cdot \mathbf{v}_3 = 0, \quad \mathbf{v}_2 \cdot \mathbf{v}_3 = 0$$

To convert the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ to an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$, we first compute $\|\mathbf{v}_1\| = 2$, $\|\mathbf{v}_2\| = 3\sqrt{2}$, and $\|\mathbf{v}_3\| = 4\sqrt{2}$, and then normalize to obtain

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (0, 1, 0), \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right), \quad \mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(-\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right) \qquad \blacksquare$$

Theorem 7.9.1 *An orthogonal set of nonzero vectors in R^n is linearly independent.*

Proof Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal set of nonzero vectors in R^n. We must show that the only scalars that satisfy the vector equation

$$t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k = \mathbf{0} \tag{1}$$

are $t_1 = 0, t_2 = 0, \ldots, t_k = 0$. To do this, let $\mathbf{v}_j$ be any vector in S; then (1) implies that

$$(t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k) \cdot \mathbf{v}_j = \mathbf{0} \cdot \mathbf{v}_j = 0$$

which we can rewrite as

$$t_1(\mathbf{v}_1 \cdot \mathbf{v}_j) + t_2(\mathbf{v}_2 \cdot \mathbf{v}_j) + \cdots + t_k(\mathbf{v}_k \cdot \mathbf{v}_j) = 0 \tag{2}$$

But each pair of distinct vectors in S is orthogonal, so all of the dot products in this equation are zero, with the possible exception of $\mathbf{v}_j \cdot \mathbf{v}_j$. Thus, (2) can be simplified to

$$t_j(\mathbf{v}_j \cdot \mathbf{v}_j) = 0 \tag{3}$$

Since we have assumed that each vector in S is nonzero, this is true of $\mathbf{v}_j$, so it follows that $\mathbf{v}_j \cdot \mathbf{v}_j = \|\mathbf{v}_j\|^2 \neq 0$. Thus, (3) implies that $t_j = 0$, and since the choice of j is arbitrary, the proof is complete. ■

EXAMPLE 3

An Orthonormal Basis for R^3

Show that the vectors

$$\mathbf{v}_1 = \left(\tfrac{3}{7}, -\tfrac{6}{7}, \tfrac{2}{7}\right), \quad \mathbf{v}_2 = \left(\tfrac{2}{7}, \tfrac{3}{7}, \tfrac{6}{7}\right), \quad \mathbf{v}_3 = \left(\tfrac{6}{7}, \tfrac{2}{7}, -\tfrac{3}{7}\right)$$

form an orthonormal basis for R^3.

Solution The vectors are orthonormal, since

$$\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1 \quad \text{and} \quad \mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{v}_2 \cdot \mathbf{v}_3 = 0$$

and hence they are linearly independent by Theorem 7.9.1. Since we have three linearly independent vectors in R^3, they must form a basis for R^3. ■

ORTHOGONAL PROJECTIONS USING ORTHONORMAL BASES

Orthonormal bases are important because they simplify many formulas and numerical calculations. For example, we know from Theorem 7.7.5 that if W is a nonzero subspace of R^n, and if $\mathbf{x}$ is a vector in R^n that is expressed in column form, then

$$\text{proj}_W \mathbf{x} = M(M^T M)^{-1} M^T \mathbf{x} \tag{4}$$

for any matrix M whose column vectors form a basis for W. In particular, if the column vectors of M are orthonormal, then $M^T M = I$, so (4) simplifies to

$$\text{proj}_W \mathbf{x} = M M^T \mathbf{x} \tag{5}$$

and Formula (27) of Section 7.7 for the standard matrix of this orthogonal projection simplifies to

$$P = M M^T \tag{6}$$

Thus, using an orthonormal basis for W eliminates the matrix inversions in the projection formulas, and reduces the calculation of an orthogonal projection to matrix multiplication.

EXAMPLE 4

Standard Matrix for a Projection Using an Orthonormal Basis

Find the standard matrix P for the orthogonal projection of R^3 onto the plane through the origin that is spanned by the orthonormal vectors $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$.

Solution Writing the vectors in column form and applying Formula (6) shows that the standard matrix for the projection is

$$P = MM^T = \begin{bmatrix} 0 & -\frac{4}{5} \\ 1 & 0 \\ 0 & \frac{3}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -\frac{4}{5} & 0 & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} \frac{16}{25} & 0 & -\frac{12}{25} \\ 0 & 1 & 0 \\ -\frac{12}{25} & 0 & \frac{9}{25} \end{bmatrix}$$

■

Theorem 7.9.2

(*a*) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthonormal basis for a subspace* W *of* R^n, *then the orthogonal projection of a vector* $\mathbf{x}$ *in* R^n *onto* W *can be expressed as*

$$\operatorname{proj}_W \mathbf{x} = (\mathbf{x} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{x} \cdot \mathbf{v}_k)\mathbf{v}_k \tag{7}$$

(*b*) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthogonal basis for a subspace* W *of* R^n, *then the orthogonal projection of a vector* $\mathbf{x}$ *in* R^n *onto* W *can be expressed as*

$$\operatorname{proj}_W \mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{x} \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k \tag{8}$$

***Proof* (*a*)** If we let

$$M = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_k]$$

then it follows from Formula (5) that

$$\operatorname{proj}_W \mathbf{x} = MM^T\mathbf{x}$$

$$\boxed{\operatorname{proj}_W \mathbf{x} = MM^T\mathbf{x} \qquad (5)}$$

Since the row vectors of M^T are the transposes of the column vectors of M, it follows from the row-column rule for matrix multiplication (Theorem 3.1.7) that

$$M^T\mathbf{x} = \begin{bmatrix} \mathbf{v}_1^T\mathbf{x} \\ \mathbf{v}_2^T\mathbf{x} \\ \vdots \\ \mathbf{v}_k^T\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{x}\cdot\mathbf{v}_1 \\ \mathbf{x}\cdot\mathbf{v}_2 \\ \vdots \\ \mathbf{x}\cdot\mathbf{v}_k \end{bmatrix} \tag{9}$$

Thus, it follows from (5) and (9) that

$$\operatorname{proj}_W\mathbf{x} = M(M^T\mathbf{x}) = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_k]\begin{bmatrix} \mathbf{x}\cdot\mathbf{v}_1 \\ \mathbf{x}\cdot\mathbf{v}_2 \\ \vdots \\ \mathbf{x}\cdot\mathbf{v}_k \end{bmatrix} = (\mathbf{x}\cdot\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x}\cdot\mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{x}\cdot\mathbf{v}_k)\mathbf{v}_k$$

which proves (7).

Proof (b) Formula (8) can be derived by normalizing the orthogonal basis to obtain an orthonormal basis and applying (7). We leave the details for the exercises. ■

EXAMPLE 5 An Orthogonal Projection Using an Orthonormal Basis

Find the orthogonal projection of $\mathbf{x} = (1, 1, 1)$ onto the plane W in R^3 that is spanned by the orthonormal vectors $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$.

Solution One way to compute the orthogonal projection is to write $\mathbf{x}$ in column form and use the standard matrix P for the projection that was computed in Example 4. This yields

$$\operatorname{proj}_W \mathbf{x} = P\mathbf{x} = \begin{bmatrix} \frac{16}{25} & 0 & -\frac{12}{25} \\ 0 & 1 & 0 \\ -\frac{12}{25} & 0 & \frac{9}{25} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{4}{25} \\ 1 \\ -\frac{3}{25} \end{bmatrix}$$

which we can write in comma-delimited form as $\operatorname{proj}_W \mathbf{x} = \left(\frac{4}{25}, 1, -\frac{3}{25}\right)$.

Alternative Solution A second method for computing the orthogonal projection is to use Formula (7). This yields

$$\operatorname{proj}_W \mathbf{x} = (\mathbf{x} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x} \cdot \mathbf{v}_2)\mathbf{v}_2 = (1)(0, 1, 0) + \left(-\frac{1}{5}\right)\left(-\frac{4}{5}, 0, \frac{3}{5}\right) = \left(\frac{4}{25}, 1, -\frac{3}{25}\right)$$

which agrees with the result obtained using the standard matrix. ■

EXAMPLE 6 An Orthogonal Projection Using an Orthogonal Basis

Find the orthogonal projection of $\mathbf{x} = (-5, 3, 1)$ onto the plane W in R^3 that is spanned by the orthogonal vectors

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

Solution We could normalize the basis vectors and apply Formula (7) to the resulting orthonormal basis for W, but let us apply (8) directly. We have

$$\|\mathbf{v}_1\|^2 = 0^2 + 1^2 + (-1)^2 = 2 \quad \text{and} \quad \|\mathbf{v}_2\|^2 = 1^2 + 2^2 + 2^2 = 9$$

so it follows from (8) that

$$\text{proj}_W \mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{x} \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = \tfrac{2}{2}(0, 1, -1) + \tfrac{3}{9}(1, 2, 2) = \left(\tfrac{1}{3}, \tfrac{5}{3}, -\tfrac{1}{3}\right)$$ ■

TRACE AND ORTHOGONAL PROJECTIONS

Theorem 7.9.3 *If P is the standard matrix for an orthogonal projection of R^n onto a subspace of R^n, then* $\operatorname{tr}(P) = \operatorname{rank}(P)$.

Proof Suppose that P is the standard matrix for an orthogonal projection of R^n onto a k-dimensional subspace W. If we let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthonormal basis for W, then it follows from Formula (6) and Theorem 3.8.1 that

$$P = MM^T = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_k]\begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_k^T \end{bmatrix} = \mathbf{v}_1\mathbf{v}_1^T + \mathbf{v}_2\mathbf{v}_2^T + \cdots + \mathbf{v}_k\mathbf{v}_k^T \tag{10}$$

Definition 3.1.10 If A is a square matrix, then the ***trace*** of A, denoted by $\operatorname{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A.

Using this result, the additive property of the trace, and Formula (27) of Section 3.1, we obtain

$$\boxed{\operatorname{tr}(\mathbf{u}\mathbf{v}^T) = \operatorname{tr}(\mathbf{v}\mathbf{u}^T) = \mathbf{u} \cdot \mathbf{v}}$$

$$\begin{aligned}
\operatorname{tr}(P) &= \operatorname{tr}\left(\mathbf{v}_1\mathbf{v}_1^T\right) + \operatorname{tr}\left(\mathbf{v}_2\mathbf{v}_2^T\right) + \cdots + \operatorname{tr}\left(\mathbf{v}_k\mathbf{v}_k^T\right) \\
&= (\mathbf{v}_1 \cdot \mathbf{v}_1) + (\mathbf{v}_2 \cdot \mathbf{v}_2) + \cdots + (\mathbf{v}_k \cdot \mathbf{v}_k) \\
&= \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 + \cdots + \|\mathbf{v}_k\|^2 \\
&= 1 + 1 + \cdots + 1 = k = \dim(W)
\end{aligned}$$

But the range of a matrix transformation is the column space of the matrix, so it follows from this computation that $\operatorname{tr}(P) = \dim(\operatorname{col}(P)) = \operatorname{rank}(P)$. ■

The subspace W is the range of P and the range of P is its column space.
As the column space and the row space have the same dimension, the proof is complete.

Theorem 3.8.1 (*Column-Row Rule*) *If A has size $m \times s$ and B has size $s \times n$, and if these matrices are partitioned into column and row vectors as*

$$A = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \cdots \quad \mathbf{c}_s] \quad \text{and} \quad B = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_s \end{bmatrix}$$

then

$$AB = \mathbf{c}_1\mathbf{r}_1 + \mathbf{c}_2\mathbf{r}_2 + \cdots + \mathbf{c}_s\mathbf{r}_s \tag{2}$$

EXAMPLE 7

Using the Trace to Find the Rank of an Orthogonal Projection

We showed in Example 4 that the standard matrix P for the orthogonal projection onto the plane spanned by the vectors $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$ is

$$P = \begin{bmatrix} \frac{16}{25} & 0 & -\frac{12}{25} \\ 0 & 1 & 0 \\ -\frac{12}{25} & 0 & \frac{9}{25} \end{bmatrix}$$

Since the plane is two-dimensional, the matrix P must have rank 2, which is confirmed by the computation

$$\operatorname{rank}(P) = \operatorname{tr}(P) = \frac{16}{25} + 1 + \frac{9}{25} = 2 \qquad \blacksquare$$

LINEAR COMBINATIONS OF ORTHONORMAL BASIS VECTORS

If W is a subspace of R^n, and if $\mathbf{w}$ is a vector in W, then $\text{proj}_W \mathbf{w} = \mathbf{w}$. Thus, we have the following special case of Theorem 7.9.2.

Theorem 7.9.4

(a) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthonormal basis for a subspace* W *of* R^n*, and if* $\mathbf{w}$ *is a vector in* W*, then*

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{w} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{w} \cdot \mathbf{v}_k)\mathbf{v}_k \tag{11}$$

(b) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthogonal basis for a subspace* W *of* R^n*, and if* $\mathbf{w}$ *is a vector in* W*, then*

$$\mathbf{w} = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{w} \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{w} \cdot \mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k \tag{12}$$

(b) *If* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthogonal basis for a subspace* W *of* R^n, *and if* $\mathbf{w}$ *is a vector in* W, *then*

$$\mathbf{w} = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{w} \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{w} \cdot \mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k \qquad (12)$$

Recalling Formula (5) of Section 7.7, observe that the terms on the right side of (12) are the orthogonal projections onto the lines spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, so that Formula (12) decomposes each vector $\mathbf{w}$ in the k-dimensional subspace W into a sum of k projections onto one-dimensional subspaces. Figure 7.9.1 illustrates this idea in the case where W is R^2.

$$\boxed{\text{proj}_{\mathbf{a}}\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \qquad (5)}$$

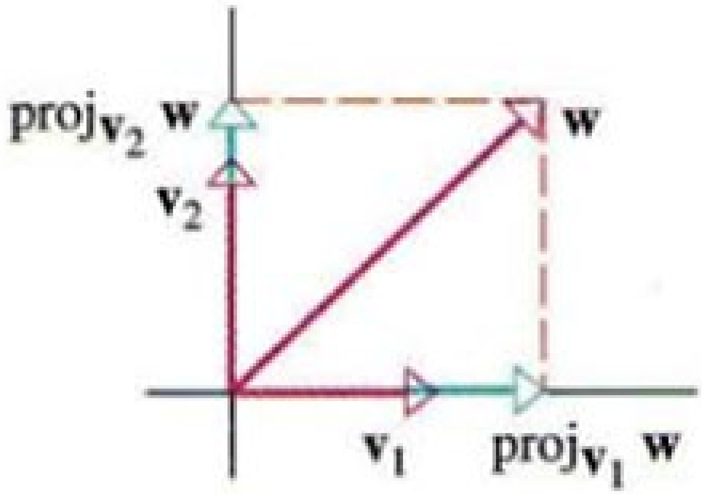

Figure 7.9.1

EXAMPLE 8 Linear Combinations of Orthonormal Basis Vectors

Express the vector $\mathbf{w} = (1, 1, 1)$ in R^3 as a linear combination of the orthonormal basis vectors

$$\mathbf{v}_1 = \left(\tfrac{3}{7}, -\tfrac{6}{7}, \tfrac{2}{7}\right), \quad \mathbf{v}_2 = \left(\tfrac{2}{7}, \tfrac{3}{7}, \tfrac{6}{7}\right), \quad \mathbf{v}_3 = \left(\tfrac{6}{7}, \tfrac{2}{7}, -\tfrac{3}{7}\right)$$

Solution We showed in Example 3 that the given vectors form an orthonormal basis for R^3. Thus, $\mathbf{w}$ can be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, and $\mathbf{v}_3$ using Formula (11). We leave it for you to confirm that

$$\mathbf{w} \cdot \mathbf{v}_1 = -\tfrac{1}{7}, \quad \mathbf{w} \cdot \mathbf{v}_2 = \tfrac{11}{7}, \quad \mathbf{w} \cdot \mathbf{v}_3 = \tfrac{5}{7}$$

Thus, it follows from Formula (11) that

$$\mathbf{w} = -\tfrac{1}{7}\mathbf{v}_1 + \tfrac{11}{7}\mathbf{v}_2 + \tfrac{5}{7}\mathbf{v}_3$$

or expressed in component form,

$$\mathbf{w} = -\tfrac{1}{7}\left(\tfrac{3}{7}, -\tfrac{6}{7}, \tfrac{2}{7}\right) + \tfrac{11}{7}\left(\tfrac{2}{7}, \tfrac{3}{7}, \tfrac{6}{7}\right) + \tfrac{5}{7}\left(\tfrac{6}{7}, \tfrac{2}{7}, -\tfrac{3}{7}\right) \qquad \blacksquare$$

FINDING ORTHOGONAL AND ORTHONORMAL BASES

The following theorem, which is the main result in this section, shows that every nonzero subspace of R^n has an orthonormal basis. The proof of this theorem is especially important because it provides a method for converting any basis for a subspace of R^n into an orthogonal basis.

Theorem 7.9.5 *Every nonzero subspace of R^n has an orthonormal basis.*

Proof Let W be a nonzero subspace of R^n, and let $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ be *any* basis for W. To prove that W has an orthonormal basis, it suffices to show that W has an *orthogonal* basis, since such a basis can then be converted into an orthonormal basis by normalizing the vectors.

The following sequence of steps will produce an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ for W:

Step 1. Let $\mathbf{v}_1 = \mathbf{w}_1$.

Step 2. As illustrated in Figure 7.9.2, we can obtain a vector $\mathbf{v}_2$ that is orthogonal to $\mathbf{v}_1$ by computing the component of $\mathbf{w}_2$ that is orthogonal to the subspace W_1 spanned by $\mathbf{v}_1$. By applying Formula (8) in Theorem 7.9.2 and Formula (24) of Section 7.7, we can express this component as

$$\mathbf{v}_2 = \mathbf{w}_2 - \operatorname{proj}_{W_1}\mathbf{w}_2 = \mathbf{w}_2 - \frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \tag{13}$$

Of course, if $\mathbf{v}_2 = \mathbf{0}$, then $\mathbf{v}_2$ is not a basis vector. But this cannot happen, since it would then follow from (13) and Step 1 that

$$\mathbf{w}_2 = \frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{w}_1$$

which states that $\mathbf{w}_2$ is a scalar multiple of $\mathbf{w}_1$, contradicting the linear independence of the basis vectors $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$.

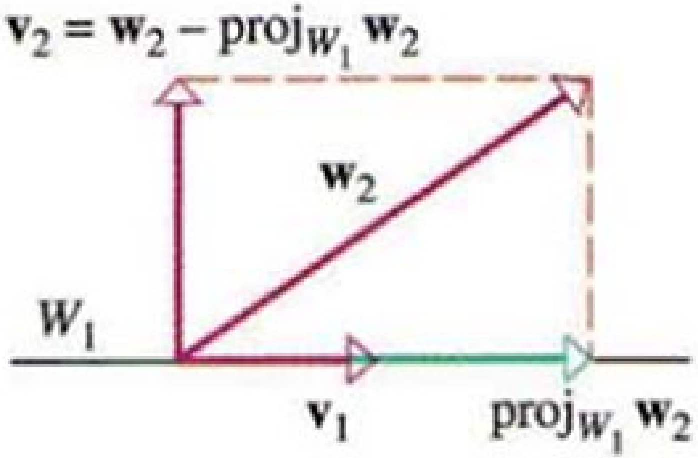

Figure 7.9.2

Step 3. To obtain a vector $\mathbf{v}_3$ that is orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$, we will compute the component of $\mathbf{w}_3$ that is orthogonal to the subspace W_2 that is spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$ (Figure 7.9.3). By applying Formula (8), we can express this component as

$$\mathbf{v}_3 = \mathbf{w}_3 - \text{proj}_{W_2}\mathbf{w}_3 = \mathbf{w}_3 - \frac{\mathbf{w}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{w}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

As in Step 2, the linear independence of $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ ensures that $\mathbf{v}_3 \neq \mathbf{0}$. We leave the details as an exercise.

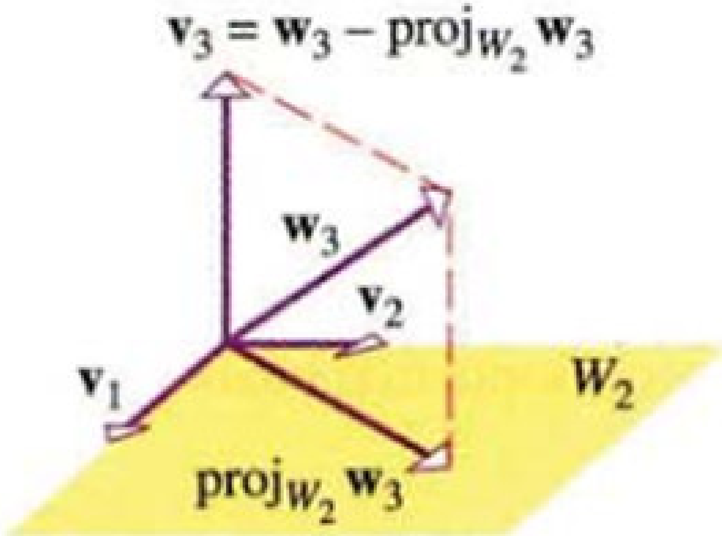

Figure 7.9.3

Step 4. To obtain a vector $\mathbf{v}_4$ that is orthogonal to $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, we will compute the component of $\mathbf{w}_4$ that is orthogonal to the subspace W_3 spanned by $\mathbf{v}_1, \mathbf{v}_2$, and $\mathbf{v}_3$. By applying Formula (8) again, we can express this component as

$$\mathbf{v}_4 = \mathbf{w}_4 - \operatorname{proj}_{W_3}\mathbf{w}_4 = \mathbf{w}_4 - \frac{\mathbf{w}_4 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{w}_4 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \frac{\mathbf{w}_4 \cdot \mathbf{v}_3}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$$

Steps 5 to k. Continuing in this way produces an orthogonal set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ after k steps. Since W is k-dimensional, this set is an orthogonal basis for W, which completes the proof. ■

The proof of this theorem provides an algorithm, called the ***Gram–Schmidt orthogonalization process***, for converting an arbitrary basis for a subspace of R^n into an orthogonal basis for the subspace.

If the resulting orthogonal vectors are normalized to produce an orthonormal basis for the subspace, then the algorithm is called the ***Gram–Schmidt process***.

EXAMPLE 9 The vectors $\mathbf{w}_1 = (1, 1, 1)$, $\mathbf{w}_2 = (0, 1, 1)$, and $\mathbf{w}_3 = (0, 0, 1)$ form a basis for R^3 (verify). Use the Gram–Schmidt orthogonalization process to transform this basis into an orthogonal basis, and then normalize the orthogonal basis vectors to obtain an orthonormal basis for R^3.

Solution Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ denote the orthogonal basis produced by the Gram–Schmidt orthogonalization process, and let $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ denote the orthonormal basis that results from normalizing $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. To find the orthogonal basis we follow the steps in the proof of Theorem 7.9.5:

Step 1. Let $\mathbf{v}_1 = \mathbf{w}_1 = (1, 1, 1)$.

Step 2. Let $\mathbf{v}_2 = \mathbf{w}_2 - \text{proj}_{W_1} \mathbf{w}_2 = \mathbf{w}_2 - \dfrac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

$$= (0, 1, 1) - \tfrac{2}{3}(1, 1, 1) = \left(-\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right)$$

Step 3. Let $\mathbf{v}_3 = \mathbf{w}_3 - \text{proj}_{W_2}\mathbf{w}_3 = \mathbf{w}_3 - \frac{\mathbf{w}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{w}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

$$= (0, 0, 1) - \tfrac{1}{3}(1, 1, 1) - \tfrac{1/3}{2/3}\left(-\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right) = \left(0, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

Thus, the vectors

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

form an orthogonal basis for R^3. The norms of these vectors are

$$\|\mathbf{v}_1\| = \sqrt{3}, \quad \|\mathbf{v}_2\| = \tfrac{\sqrt{6}}{3}, \quad \|\mathbf{v}_3\| = \tfrac{1}{\sqrt{2}}$$

so an orthonormal basis for R^3 is given by

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right), \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\tfrac{2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\right)$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(0, -\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right) \quad \blacksquare$$

an orthonormal basis for R^3 is given by

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \qquad \blacksquare$$

REMARK In this example we first found an orthogonal basis and then normalized at the end to produce the orthonormal basis.

Alternatively, we could have normalized each orthogonal basis vector as soon as it was calculated, thereby generating the orthonormal basis step by step.

For hand calculation, it is usually better to do the normalization at the end, since this tends to postpone the introduction of bothersome square roots.

EXAMPLE 10 Orthonormal Basis for a Plane in R^3

Use the Gram–Schmidt process to construct an orthonormal basis for the plane $x + y + z = 0$ in R^3.

Solution First we will find a basis for the plane, and then we will apply the Gram–Schmidt process to that basis. Any two nonzero vectors in the plane that are not scalar multiples of one another will serve as a basis.

One way to find such vectors is to use the method of Example 7 in Section 1.3 to write the plane in the parametric form

$$x = -t_1 - t_2, \quad y = t_1, \quad z = t_2$$

The parameter values $t_1 = 1$, $t_2 = 0$ and $t_1 = 0$, $t_2 = 1$ produce the vectors

$$\mathbf{w}_1 = (-1, 1, 0) \quad \text{and} \quad \mathbf{w}_2 = (-1, 0, 1)$$

in the given plane. Now we are ready to apply the Gram–Schmidt process.

First we construct the orthogonal basis vectors

$$\mathbf{v}_1 = \mathbf{w}_1 = (-1, 1, 0)$$

$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = (-1, 0, 1) - \tfrac{1}{2}(-1, 1, 0) = \left(-\tfrac{1}{2}, -\tfrac{1}{2}, 1\right)$$

and then normalize these to obtain the orthonormal basis vectors

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)$$

$$\mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \tfrac{2}{\sqrt{6}}\left(-\tfrac{1}{2}, -\tfrac{1}{2}, 1\right) = \left(-\tfrac{1}{\sqrt{6}}, -\tfrac{1}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}\right) \qquad \blacksquare$$

A PROPERTY OF THE GRAM–SCHMIDT PROCESS

In the exercises we will ask you to show that the vector $\mathbf{v}_j$ that is produced at the jth step of the Gram–Schmidt process is expressible as a linear combination of $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_j$.

Thus, not only does the Gram–Schmidt process produce an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ for the subspace W spanned by $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$, but it also creates the basis in such a way that at each intermediate stage the vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j\}$ form an orthogonal basis for the subspace of R^n spanned by $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_j\}$.

Moreover, since $\mathbf{v}_j$ is constructed to be orthogonal to $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{j-1}\}$, the vector $\mathbf{v}_j$ must also be orthogonal to $\text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{j-1}\}$, since the two subspaces are the same.

In summary, we have the following theorem.

Theorem 7.9.6 *If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is a basis for a nonzero subspace of R^n, and if $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is the corresponding orthogonal basis produced by the Gram–Schmidt process, then:*

(*a*) $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j\}$ *is an orthogonal basis for* $\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_j\}$ *at the* jth *step.*

(*b*) $\mathbf{v}_j$ *is orthogonal to* $\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{j-1}\}$ *at the* jth *step* $(j \geq 2)$.

REMARK This theorem remains true if the orthogonal basis vectors are normalized at each step, rather than at the end of the process; that is, $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_j\}$ is an *orthonormal* basis for $\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_j\}$ and $\mathbf{q}_j$ is orthogonal to $\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{j-1}\}$.

EXTENDING ORTHONORMAL SETS TO ORTHONORMAL BASES

Recall from part (*b*) of Theorem 7.2.2 that a linearly independent set in a nonzero subspace W of R^n can be enlarged to a basis for W. The following theorem is an analog of that result for orthogonal bases and orthonormal bases.

Theorem 7.9.7 *If W is a nonzero subspace of R^n, then:*

(*a*) *Every orthogonal set of nonzero vectors in W can be enlarged to an orthogonal basis for W.*

(*b*) *Every orthonormal set in W can be enlarged to an orthonormal basis for W.*

We will prove part (*b*); the proof of part (*a*) is similar.

Proof (b) Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ is an orthonormal set of vectors in W. Part (*b*) of Theorem 7.2.2 tells us that we can enlarge S to some basis $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s, \mathbf{w}_{s+1}, \ldots, \mathbf{w}_k\}$ for W. If we now apply the Gram–Schmidt process to the set S', then the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ will not be altered since they are already orthonormal, and the resulting set

$$S'' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s, \mathbf{v}_{s+1}, \ldots, \mathbf{v}_k\}$$

will be an orthonormal basis for W. ■