6.2 Angle and Orthogonality in Inner Product Spaces

In Section 3.2 we defined the notion of "angle" between vectors in R^n. In this section we will extend this idea to general vector spaces. This will enable us to extend the notion of orthogonality as well, thereby setting the groundwork for a variety of new applications.

Recall from Formula (20) of Section 3.2 that the angle θ between two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n is

$$\theta = \cos^{-1}\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right) \tag{1}$$

We were assured that this formula was valid because it followed from the Cauchy–Schwarz inequality (Theorem 3.2.4) that

$$-1 \le \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \le 1 \tag{2}$$

as required for the inverse cosine to be defined. The following generalization of the Cauchy–Schwarz inequality will enable us to define the angle between two vectors in *any* real inner product space.

THEOREM 6.2.1 Cauchy–Schwarz Inequality

If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in a real inner product space* V, *then*

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\|\|\mathbf{v}\| \tag{3}$$

Proof We warn you [illegible]ance that the proof presented here depends on a clever trick that is not easy to motivate.

In the case where $\mathbf{u} = \mathbf{0}$ the two sides of (3) are equal since $\langle \mathbf{u}, \mathbf{v}\rangle$ and $\|\mathbf{u}\|$ are both zero. Thus, we need only consider the case where $\mathbf{u} \neq \mathbf{0}$. Making this assumption, let

$$a = \langle \mathbf{u}, \mathbf{u}\rangle, \quad b = 2\langle \mathbf{u}, \mathbf{v}\rangle, \quad c = \langle \mathbf{v}, \mathbf{v}\rangle$$

and let t be any real number. Since the positivity axiom states that the inner product of any vector with itself is nonnegative, it follows that

$$\begin{aligned} 0 \le \langle t\mathbf{u}+\mathbf{v}, t\mathbf{u}+\mathbf{v}\rangle &= \langle \mathbf{u},\mathbf{u}\rangle t^2 + 2\langle \mathbf{u},\mathbf{v}\rangle t + \langle \mathbf{v},\mathbf{v}\rangle \\ &= at^2+bt+c \end{aligned}$$

This inequality implies that the quadratic polynomial at^2+bt+c has either no real roots or a repeated real root. Therefore, its discriminant must satisfy the inequality $b^2-4ac \le 0$. Expressing the coefficients a, b, and c in terms of the vectors $\mathbf{u}$ and $\mathbf{v}$ gives $4\langle \mathbf{u},\mathbf{v}\rangle^2 - 4\langle \mathbf{u},\mathbf{u}\rangle\langle \mathbf{v},\mathbf{v}\rangle \le 0$ or, equivalently,

$$\langle \mathbf{u},\mathbf{v}\rangle^2 \le \langle \mathbf{u},\mathbf{u}\rangle\langle \mathbf{v},\mathbf{v}\rangle$$

Taking square roots of both sides and using the fact that $\langle \mathbf{u},\mathbf{u}\rangle$ and $\langle \mathbf{v},\mathbf{v}\rangle$ are nonnegative yields

$$|\langle \mathbf{u},\mathbf{v}\rangle| \le \langle \mathbf{u},\mathbf{u}\rangle^{1/2}\langle \mathbf{v},\mathbf{v}\rangle^{1/2} \quad \text{or equivalently} \quad |\langle \mathbf{u},\mathbf{v}\rangle| \le \|\mathbf{u}\|\|\mathbf{v}\|$$

which completes the proof. ■

Note

The following two alternative forms of the Cauchy–Schwarz inequality are useful to know:

$$\langle \mathbf{u},\mathbf{v}\rangle^2 \le \langle \mathbf{u},\mathbf{u}\rangle\langle \mathbf{v},\mathbf{v}\rangle \tag{4}$$

$$\langle \mathbf{u},\mathbf{v}\rangle^2 \le \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \tag{5}$$

The first of these formulas was obtained in the proof of Theorem 6.2.1, and the second is a variation of the first.

Properties of Length and Distance in General Inner Product Spaces

In Section 3.2 we used the dot product to extend the notions of length and distance to R^n, and we showed that various basic geometry theorems remained valid (see Theorems 3.2.5, 3.2.6, and 3.2.7). By making only minor adjustments to the proofs of those theorems, one can show that they remain valid in any real inner product space. For example, here is the generalization of Theorem 3.2.5 (the triangle inequalities).

THEOREM 6.2.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in a real inner product space* V, *and if* k *is any scalar, then:*

(*a*) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ [Triangle inequality for vectors]

(*b*) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ [Triangle inequality for distances]

Proof **(*a*)**

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle \\
&= \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&\leq \langle \mathbf{u}, \mathbf{u}\rangle + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \langle \mathbf{v}, \mathbf{v}\rangle && \text{[Property of absolute value]} \\
&\leq \langle \mathbf{u}, \mathbf{u}\rangle + 2\|\mathbf{u}\|\|\mathbf{v}\| + \langle \mathbf{v}, \mathbf{v}\rangle && \text{[By (3)]} \\
&= \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

Taking square roots gives $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Proof **(*b*)** Identical to the proof of part (*b*) of Theorem 3.2.5.

Orthogonality

Although Example 1 is a useful mathematical exercise, there is only an occasional need to compute angles in vector spaces other than R^2 and R^3. A problem of more interest in general vector spaces is ascertaining whether the angle between vectors is $\pi/2$. You should be able to see from Formula (8) that if $\mathbf{u}$ and $\mathbf{v}$ are *nonzero* vectors, then the angle between them is $\theta = \pi/2$ if and only if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. Accordingly, we make the following definition, which is a generalization of Definition 1 in Section 3.3 and is applicable even if one or both of the vectors is zero.

Orthagonality

DEFINITION 1 Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space V called ***orthogonal*** if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

As the following example shows, orthogonality depends on the inner product in the sense that for different inner products two vectors can be orthogonal with respect to one but not the other.

▶ EXAMPLE 2 Orthogonality Depends on the Inner Product

The vectors $\mathbf{u} = (1, 1)$ and $\mathbf{v} = (1, -1)$ are orthogonal with respect to the Euclidean inner product on R^2 since

$$\mathbf{u} \cdot \mathbf{v} = (1)(1) + (1)(-1) = 0$$

However, they are not orthogonal with respect to the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$ since

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3(1)(1) + 2(1)(-1) = 1 \neq 0$$

▶ EXAMPLE 3 Orthogonal Vectors in M_{22}

If M_{22} has the inner product of Example 6 in the preceding section, then the matrices

$$U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

are orthogonal since

$$\langle U, V \rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0$$

CALCULUS REQUIRED

▶ EXAMPLE 4 Orthogonal Vectors in P_2

Let P_2 have the inner product

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} p(x)q(x)\,dx$$

and let $\mathbf{p} = x$ and $\mathbf{q} = x^2$. Then

$$\|\mathbf{p}\| = \langle \mathbf{p}, \mathbf{p} \rangle^{1/2} = \left[\int_{-1}^{1} xx\,dx\right]^{1/2} = \left[\int_{-1}^{1} x^2\,dx\right]^{1/2} = \sqrt{\frac{2}{3}}$$

$$\|\mathbf{q}\| = \langle \mathbf{q}, \mathbf{q} \rangle^{1/2} = \left[\int_{-1}^{1} x^2x^2\,dx\right]^{1/2} = \left[\int_{-1}^{1} x^4\,dx\right]^{1/2} = \sqrt{\frac{2}{5}}$$

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} xx^2\,dx = \int_{-1}^{1} x^3\,dx = 0$$

Because $\langle \mathbf{p}, \mathbf{q} \rangle = 0$, the vectors $\mathbf{p} = x$ and $\mathbf{q} = x^2$ are orthogonal relative to the given inner product. ◀

In Theorem 3.3.3 we proved the Theorem of Pythagoras for vectors in Euclidean n-space. The following theorem extends this result to vectors in any real inner product space.

THEOREM 6.2.3 Generalized Theorem of Pythagoras

If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in a real inner product space, then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

$\langle v+u, v+u \rangle = \langle v, v+u \rangle + \langle u, v+u \rangle$

$= \langle v, v \rangle + \langle v, u \rangle + \langle u, v \rangle + \langle u, u \rangle$

Proof The orthogonality of **u** and **v** implies that $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, so

$$\|\mathbf{u}+\mathbf{v}\|^2 = \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v} \rangle = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2$$
$$= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

CALCULUS REQUIRED

EXAMPLE 5 Theorem of Pythagoras in P_2

In Example 4 we showed that $\mathbf{p} = x$ and $\mathbf{q} = x^2$ are orthogonal with respect to the inner product

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} p(x)q(x)\,dx$$

on P_2. It follows from Theorem 6.2.3 that

$$\|\mathbf{p}+\mathbf{q}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{q}\|^2$$

Thus, from the computations in Example 4, we have

$$\|\mathbf{p}+\mathbf{q}\|^2 = \left(\sqrt{\frac{2}{3}}\right)^2 + \left(\sqrt{\frac{2}{5}}\right)^2 = \frac{2}{3} + \frac{2}{5} = \frac{16}{15}$$

We can check this result by direct integration:

$$\|\mathbf{p}+\mathbf{q}\|^2 = \langle \mathbf{p}+\mathbf{q}, \mathbf{p}+\mathbf{q} \rangle = \int_{-1}^{1} (x+x^2)(x+x^2)\,dx$$
$$= \int_{-1}^{1} x^2\,dx + 2\int_{-1}^{1} x^3\,dx + \int_{-1}^{1} x^4\,dx = \frac{2}{3} + 0 + \frac{2}{5} = \frac{16}{15} \blacktriangleleft$$

Orthogonal Complements

In Section 4.8 we defined the notion of an *orthogonal complement* for subspaces of R^n, and we used that definition to establish a geometric link between the fundamental spaces of a matrix. The following definition extends that idea to general inner product spaces.

Orthagonal Complement

DEFINITION 2 If W is a subspace of a real inner product space V, then the set of all vectors in V that are orthogonal to every vector in W is called the ***orthogonal complement*** of W and is denoted by the symbol $W^\perp$.

In Theorem 4.8.6 we stated three properties of orthogonal complements in R^n. The following theorem generalizes parts (*a*) and (*b*) of that theorem to general real inner product spaces.

Theorem 6.2.4

THEOREM 6.2.4 *If W is a subspace of a real inner product space V, then:*

(*a*) *$W^\perp$ is a subspace of V.*

(*b*) $W \cap W^\perp = \{\mathbf{0}\}$.

Proof (a) The set $W^\perp$ contains at least the zero vector, since $\langle \mathbf{0}, \mathbf{w} \rangle = 0$ for every vector **w** in W. Thus, it remains to show that $W^\perp$ is closed under addition and scalar multiplication. To do this, suppose that **u** and **v** are vectors in $W^\perp$, so that for every vector **w** in W we have $\langle \mathbf{u}, \mathbf{w} \rangle = 0$ and $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. It follows from the additivity and homogeneity axioms of inner products that

$$\langle \mathbf{u}+\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle = 0 + 0 = 0$$
$$\langle k\mathbf{u}, \mathbf{w} \rangle = k\langle \mathbf{u}, \mathbf{w} \rangle = k(0) = 0$$

which proves that $\mathbf{u}+\mathbf{v}$ and $k\mathbf{u}$ are in $W^\perp$.

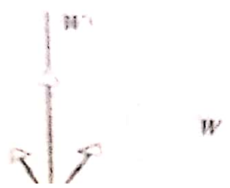
Proof (b) If $\mathbf{v}$ is any vector in both W and $W^{\perp}$, then $\mathbf{v}$ is orthogonal to itself; that is, $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. It follows from the positivity axiom for inner products that $\mathbf{v} = \mathbf{0}$. ◂

The next theorem, which we state without proof, generalizes part (c) of Theorem 4.8.6. Note, however, that this theorem applies only to finite-dimensional inner product spaces, whereas Theorem 4.8.6 does not have this restriction.

Theorem 6.2.5 implies that in a finite-dimensional inner product space orthogonal complements occur in pairs, each being orthogonal to the other (Figure 6.2.2).

THEOREM 6.2.5 *If W is a subspace of a real finite-dimensional inner product space V, then the orthogonal complement of $W^{\perp}$ is W; that is,*

$$(W^{\perp})^{\perp} = W$$



In our study of the fundamental spaces of a matrix in Section 4.8 we showed that the row space and null space of a matrix are orthogonal complements with respect to the Euclidean inner product on R^n (Theorem 4.8.7). The following example takes advantage of that fact.

6.3 Gram–Schmidt Process; QR-Decomposition

In many problems involving vector spaces, the problem solver is free to choose any basis for the vector space that seems appropriate. In inner product spaces, the solution of a problem can often be simplified by choosing a basis in which the vectors are orthogonal to one another. In this section we will show how such bases can be obtained.

Orthogonal and Orthonormal Sets

Orthagonal and orthanormal Sets

Recall from Section 6.2 that two vectors in an inner product space are said to be *orthogonal* if their inner product is zero. The following definition extends the notion of orthogonality to *sets* of vectors in an inner product space.

DEFINITION 1 A set of two or more vectors in a real inner product space is said to be ***orthogonal*** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is said to be ***orthonormal.***

EXAMPLE 1 An Orthogonal Set in R^3

Let

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

and assume that R^3 has the Euclidean inner product. It follows that the set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthogonal since $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$. ◄

Note

It frequently happens that one has found a set of orthogonal vectors in an inner product space but what is actually needed is a set of *orthonormal* vectors. A simple way to convert an orthogonal set of nonzero vectors into an orthonormal set is to multiply each vector $\mathbf{v}$ in the orthogonal set by the reciprocal of its length to create a vector of norm 1 (called a ***unit vector***). To see why this works, suppose that $\mathbf{v}$ is a nonzero vector in an inner product space, and let

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} \tag{1}$$

Then it follows from Theorem 6.1.1(b) with $k = \|\mathbf{v}\|$ that

$$\|\mathbf{u}\| = \left\| \frac{1}{\|\mathbf{v}\|}\mathbf{v} \right\| = \left| \frac{1}{\|\mathbf{v}\|} \right| \|\mathbf{v}\| = \frac{1}{\|\mathbf{v}\|}\|\mathbf{v}\| = 1$$

This process of multiplying a vector $\mathbf{v}$ by the reciprocal of its length is called ***normalizing*** $\mathbf{v}$. We leave it as an exercise to show that normalizing the vectors in an orthogonal set of nonzero vectors preserves the orthogonality of the vectors and produces an orthonormal set.

Note that Formula (1) is identical to Formula (4) of Section 3.2, but whereas Formula (4) was valid only for vectors in R^n with the Euclidean inner product, Formula (1) is valid in general inner product spaces.

EXAMPLE 2 Constructing an Orthonormal Set

The Euclidean norms of the vectors in Example 1 are

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

Consequently, normalizing $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ yields

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (0, 1, 0), \quad \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

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

We leave it for you to verify that the set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthonormal by showing that

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_3 \rangle = \langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0 \quad \text{and} \quad \|\mathbf{u}_1\| = \|\mathbf{u}_2\| = \|\mathbf{u}_3\| = 1 \blacktriangleleft$$

In R^2 any two nonzero perpendicular vectors are linearly independent because neither is a scalar multiple of the other; and in R^3 any three nonzero mutually perpendicular vectors are linearly independent because no one lies in the plane of the other two (and hence is not expressible as a linear combination of the other two). The following theorem generalizes these observations.

THEOREM 6.3.1 *If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space, then S is linearly independent.*

Proof Assume that

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n = \mathbf{0} \tag{2}$$

To demonstrate that $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent, we must prove that $k_1 = k_2 = \cdots = k_n = 0$.

For each $\mathbf{v}_i$ in S, it follows from (2) that

$$\langle k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n, \mathbf{v}_i \rangle = \langle \mathbf{0}, \mathbf{v}_i \rangle = 0$$

or, equivalently,

$$k_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + k_2\langle \mathbf{v}_2, \mathbf{v}_i \rangle + \cdots + k_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle = 0$$

From the orthogonality of S it follows that $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$ when $j \neq i$, so this equation reduces to

$$k_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 0$$

Since the vectors in S are assumed to be nonzero, it follows from the positivity axiom for inner products that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle \neq 0$. Thus, the preceding equation implies that each k_i in Equation (2) is zero, which is what we wanted to prove. ◀

Since an orthonormal set is orthogonal, and since its vectors are nonzero (norm 1), it follows from Theorem 6.3.1 that every *orthonormal* set is linearly independent.

Note

In an inner product space, a basis consisting of orthonormal vectors is called an ***orthonormal basis***, and a basis consisting of orthogonal vectors is called an ***orthogonal basis***. A familiar example of an orthonormal basis is the standard basis for R^n with the Euclidean inner product:

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

▶ EXAMPLE 3 An Orthonormal Basis for P_n

Recall from Example 7 of Section 6.1 that the standard inner product of the polynomials

$$\mathbf{p} = a_0 + a_1x + \cdots + a_nx^n \quad \text{and} \quad \mathbf{q} = b_0 + b_1x + \cdots + b_nx^n$$

is

$$\langle \mathbf{p}, \mathbf{q} \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$$

and the norm of $\mathbf{p}$ relative to this inner product is

$$\|\mathbf{p}\| = \sqrt{\langle \mathbf{p}, \mathbf{p} \rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2}$$

You should be able to see from these formulas that the standard basis

$$S = \{1, x, x^2, \ldots, x^n\}$$

is orthonormal with respect to this inner product.

EXAMPLE 4 An Orthonormal Basis

In Example 2 we showed that the vectors

$$\mathbf{u}_1 = (0, 1, 0), \quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad \text{and} \quad \mathbf{u}_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

form an orthonormal set with respect to the Euclidean inner product on R^3. By Theorem 6.3.1, these vectors form a linearly independent set, and since R^3 is three-dimensional, it follows from Theorem 4.5.4 that $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for R^3. ◀

oordinates Relative to Orthonormal Bases

One way to express a vector $\mathbf{u}$ as a linear combination of basis vectors

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

is to convert the vector equation

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

to a linear system and solve for the coefficients $c_1, c_2, \ldots, c_n$. However, if the basis happens to be orthogonal or orthonormal, then the following theorem shows that the coefficients can be obtained more simply by computing appropriate inner products.

THEOREM 6.3.2

(*a*) *If* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *is an orthogonal basis for an inner product space* V, *and if* $\mathbf{u}$ *is any vector in* V, *then*

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n\rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n \tag{3}$$

(*b*) *If* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *is an orthonormal basis for an inner product space* V, *and if* $\mathbf{u}$ *is any vector in* V, *then*

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle\mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle\mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n\rangle\mathbf{v}_n \tag{4}$$

▶ EXAMPLE An Orthonormal Basis from an Orthogonal Basis

(a) Show that the vectors

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

form an orthogonal basis for R^3 with the Euclidean inner product, and use that basis to find an orthonormal basis by normalizing each vector.

(b) Express the vector $\mathbf{u} = (1, 2, 4)$ as a linear combination of the orthonormal basis vectors obtained in part (a).

Solution (a) The given vectors form an orthogonal set since

$$\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 0, \quad \langle \mathbf{w}_1, \mathbf{w}_3 \rangle = 0, \quad \langle \mathbf{w}_2, \mathbf{w}_3 \rangle = 0$$

It follows from Theorem 6.3.1 that these vectors are linearly independent and hence form a basis for R^3 by Theorem 4.5.4. We leave it for you to calculate the norms of $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$ and then obtain the orthonormal basis

$$\mathbf{v}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = (0, 1, 0), \quad \mathbf{v}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

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

Solution (b) It follows from Formula (4) that

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle \mathbf{v}_2 + \langle \mathbf{u}, \mathbf{v}_3\rangle \mathbf{v}_3$$

We leave it for you to confirm that

$$\langle \mathbf{u}, \mathbf{v}_1\rangle = (1, 2, 4) \cdot (0, 1, 0) = 2$$

$$\langle \mathbf{u}, \mathbf{v}_2\rangle = (1, 2, 4) \cdot \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = \frac{5}{\sqrt{2}}$$

$$\langle \mathbf{u}, \mathbf{v}_3\rangle = (1, 2, 4) \cdot \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = \frac{3}{\sqrt{2}}$$

and hence that

$$(1, 2, 4) = 2(0, 1, 0) + \frac{5}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + \frac{3}{\sqrt{2}}\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \blacktriangleleft$$

Orthogonal Projections

Many applied problems are best solved by working with orthogonal or orthonormal basis vectors. Such bases are typically found by starting with some simple basis (say a standard basis) and then converting that basis into an orthogonal or orthonormal basis. To explain exactly how that is done will require some preliminary ideas about orthogonal projections.

In Section 3.3 we proved a result called the *Projection Theorem* (see Theorem 3.3.2) that dealt with the problem of decomposing a vector $\mathbf{u}$ in R^n into a sum of two terms, $\mathbf{w}_1$ and $\mathbf{w}_2$, in which $\mathbf{w}_1$ is the orthogonal projection of $\mathbf{u}$ on some nonzero vector a and $\mathbf{w}_2$ is orthogonal to $\mathbf{w}_1$ (Figure 3.3.2). That result is a special case of the following more general theorem, which we will state without proof.

THEOREM 6.3.3 Projection Theorem

If W is a finite-dimensional subspace of an inner product space V, then every vector $\mathbf{u}$ in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \tag{8}$$

where $\mathbf{w}_1$ is in W and $\mathbf{w}_2$ is in $W^{\perp}$.

The vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in Formula (8) are commonly denoted by

$$\mathbf{w}_1 = \text{proj}_W \mathbf{u} \quad \text{and} \quad \mathbf{w}_2 = \text{proj}_{W^\perp} \mathbf{u} \tag{9}$$

These are called the ***orthogonal projection of* u *on* W** and the ***orthogonal projection of* u *on* $W^{\perp}$**, respectively. The vector $\mathbf{w}_2$ is also called the ***component of* u *orthogonal to* W**. Using the notation in (9), Formula (8) can be expressed as

$$\mathbf{u} = \text{proj}_W \mathbf{u} + \text{proj}_{W^\perp} \mathbf{u} \tag{10}$$

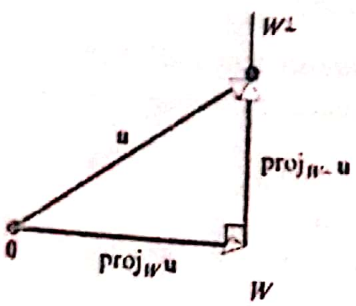


Figure 6.3.1

(Figure 6.3.1). Moreover, since $\text{proj}_{W^\perp}\mathbf{u} = \mathbf{u} - \text{proj}_W \mathbf{u}$, we can also express Formula (10) as

$$\mathbf{u} = \text{proj}_W \mathbf{u} + (\mathbf{u} - \text{proj}_W \mathbf{u}) \tag{11}$$