4.6 Orthogonal Vectors in R^{*n*}

In this section we show that the geometrical concepts of *distance* and *angle* in *n*-dimensional space can be based on the definition of the *dot product* of two vectors in \mathbb{R}^n . Recall from elementary calculus that the dot product $\mathbf{u} \cdot \mathbf{v}$ of two 3-dimensional vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ is (by definition) the sum

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

of the products of corresponding scalar components of the two vectors.

Similarly, the **dot product** $\mathbf{u} \cdot \mathbf{v}$ of two *n*-dimensional vectors $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$ is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n \tag{1}$$

(with one additional scalar product term for each additional dimension). And just as in \mathbb{R}^3 , it follows readily from the formula in (1) that if \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors in \mathbb{R}^n

and c is a scalar, then

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$$
 (symmetry) (2)

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$$
 (distributivity) (3)

$$(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$$
 (homogeneity) (4)

$$\mathbf{u} \cdot \mathbf{u} \ge 0;$$

$$\mathbf{u} \cdot \mathbf{u} = 0$$
 if and only if

$$\mathbf{u} = \mathbf{0}.$$
 (positivity) (5)

Therefore, the dot product in \mathbf{R}^n is an example of an *inner product*.

DEFINITION Inner Product

An inner product on a vector space V is a function that associates with each pair of vectors \mathbf{u} and \mathbf{v} in V a scalar $\langle \mathbf{u}, \mathbf{v} \rangle$ such that, if \mathbf{u}, \mathbf{v} , and \mathbf{w} are vectors and c is a scalar, then

- (ii) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle;$
- (iii) $\langle c\mathbf{u}, \mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$;

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle;$

(iv) $\langle \mathbf{u}, \mathbf{u} \rangle \ge 0$; $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

The dot product on \mathbb{R}^n is sometimes called the Euclidean inner product, and with this inner product \mathbb{R}^n is sometimes called Euclidean *n*-dimensional space. We can use any of the notations in

$$\mathbf{u} \cdot \mathbf{v} = \langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v}$$

for the dot product of the two $n \times 1$ column vectors **u** and **v**. (Note in the last expression that $\mathbf{u}^T = (u_1, u_2, \dots, u_n)^T$ is the $1 \times n$ row vector with the indicated entries, so the 1×1 matrix product $\mathbf{u}^T \mathbf{v}$ is simply a scalar.) Here we will ordinarily use the notation $\mathbf{u} \cdot \mathbf{v}$.

The **length** $|\mathbf{u}|$ of the vector $\mathbf{u} = (u_1, u_2, \dots, u_n)$ is defined as follows:

$$|\mathbf{u}| = \sqrt{(\mathbf{u} \cdot \mathbf{u})} = \left(u_1^2 + u_2^2 + \dots + u_n^2\right)^{1/2}.$$
 (6)

Note that the case n = 2 is a consequence of the familiar Pythagorean formula in the plane.

Theorem 1 gives one of the most important inequalities in mathematics. Many proofs are known, but none of them seems direct and well motivated.

THEOREM 1 The Cauchy-Schwarz Inequality If u and v are vectors in \mathbb{R}^n , then $|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| |\mathbf{v}|$.

(7)

4.6 Orthogonal Vectors in Rⁿ 271

Proof: If $\mathbf{u} = \mathbf{0}$, then $|\mathbf{u} \cdot \mathbf{v}| = |\mathbf{u}| = 0$, so the inequality is satisfied trivially. If $\mathbf{u} \neq \mathbf{0}$, then we let $a = \mathbf{u} \cdot \mathbf{u}$, $b = 2\mathbf{u} \cdot \mathbf{v}$, and $c = \mathbf{v} \cdot \mathbf{v}$. For any real number x, the distributivity and positivity properties of the dot product then yield

$$0 \le (x\mathbf{u} + \mathbf{v}) \cdot (x\mathbf{u} + \mathbf{v})$$

= $(\mathbf{u} \cdot \mathbf{u})x^2 + 2(\mathbf{u} \cdot \mathbf{v})x + (\mathbf{v} \cdot \mathbf{v}),$

so that

$$0 \leq ax^2 + bx + c.$$

Thus the quadratic equation $ax^2 + bx + c = 0$ either has no real roots or has a repeated real root. Hence the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

implies that the discriminant $b^2 - 4ac$ cannot be positive; that is, $b^2 \leq 4ac$, so

$$4(\mathbf{u}\cdot\mathbf{v})^2 \leq 4(\mathbf{u}\cdot\mathbf{u})(\mathbf{v}\cdot\mathbf{v}).$$

We get the Cauchy-Schwarz inequality in (7) when we take square roots, remembering that the numbers $|\mathbf{u}| = (\mathbf{u} \cdot \mathbf{u})^{1/2}$ and $|\mathbf{v}| = (\mathbf{v} \cdot \mathbf{v})^{1/2}$ are nonnegative.

The Cauchy-Schwarz inequality enables us to *define* the angle θ between the nonzero vectors **u** and **v**. (See Figure 4.6.1.) Division by the positive number $|\mathbf{u}||\mathbf{v}|$ in (7) yields $|\mathbf{u} \cdot \mathbf{v}|/(|\mathbf{u}||\mathbf{v}|) \leq 1$, so

$$-1 \leq \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \leq +1. \tag{8}$$

Hence there is a unique angle θ between 0 and π radians, inclusive (that is, between 0° and 180°), such that

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|}.\tag{9}$$

Thus we obtain the same geometric interpretation

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \cos \theta \tag{10}$$

of the dot product in \mathbf{R}^n as one sees (for 3-dimensional vectors) in elementary calculus textbooks—for instance, see Section 11.2 of Edwards and Penney, *Calculus: Early Transcendentals*, 7th edition (Prentice Hall) 2008.

On the basis of (10) we call the vectors \mathbf{u} and \mathbf{v} orthogonal provided that

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{0}. \tag{11}$$

If **u** and **v** are nonzero vectors this means that $\cos \theta = 0$, so $\theta = \pi/2$ (90°). Note that **u** = **0** satisfies (11) for all **v**, so the zero vector is orthogonal to *every* vector.

Example 1 Find the angle θ_n in \mathbf{R}^n between the x_1 -axis and the line through the origin and the point (1, 1, ..., 1).



FIGURE 4.6.1. The angle θ where the vectors **u** and **v**.

272 Chapter 4 Vector Spaces

Solution We take $\mathbf{u} = (1, 0, 0, \dots, 0)$ on the x_1 -axis and $\mathbf{v} = (1, 1, \dots, 1)$. Then $|\mathbf{u}| = 1$ $|\mathbf{v}| = \sqrt{n}$, and $\mathbf{u} \cdot \mathbf{v} = 1$, so the formula in (9) gives

$$\cos \theta_n = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| |\mathbf{v}|} = \frac{1}{\sqrt{n}}$$

For instance, if

n = 3, then
$$\theta_3 = \cos^{-1}\left(\frac{1}{\sqrt{3}}\right) \approx 0.9553$$
 (55°);
n = 4, then $\theta_4 = \cos^{-1}\left(\frac{1}{\sqrt{4}}\right) \approx 1.0472$ (60°);
n = 5, then $\theta_5 = \cos^{-1}\left(\frac{1}{\sqrt{5}}\right) \approx 1.1071$ (63°);
n = 100, then $\theta_{100} = \cos^{-1}\left(\frac{1}{10}\right) \approx 1.4706$ (84°).

It is interesting to note that θ_n increases as *n* increases. Indeed, θ_n approaches $\cos^{-1}(0) = \pi/2$ (90°) as *n* increases without bound (so that $1/\sqrt{n}$ approaches zero).

In addition to angles, the dot product provides a definition of distance in \mathbb{R} The **distance** $d(\mathbf{u}, \mathbf{v})$ between the points (vectors) $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$ is defined to be

$$d(\mathbf{u}, \mathbf{v}) = |\mathbf{u} - \mathbf{v}|$$

= $[(u_1 - v_1)^2 + (u_2 - v_2)^2 + \dots + (u_n - v_n)^2]^{1/2}$.

Example 2 The distance between the points $\mathbf{u} = (1, -1, -2, 3, 5)$ and $\mathbf{v} = (4, 3, 4, 5, 9)$ in \mathbb{R} is

$$|\mathbf{u} - \mathbf{v}| = \sqrt{3^2 + 4^2 + 6^2 + 2^2 + 4^2} = \sqrt{81} = 9.$$

The *triangle inequality* of Theorem 2 relates the three sides of the triangle shown in Figure 4.6.2.

THEOREM 2 The Triangle Inequality

If **u** and **v** are vectors in \mathbb{R}^n , then

$$|\mathbf{u} + \mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|. \tag{13}$$



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

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



FIGURE 4.6.2. The "triangle" of the triangle inequality.

4.6 Orthogonal Vectors in Rⁿ 273

and therefore

$$||u + v||^2 \leq (|u| + |v|)^2$$
.

We now get (13) when we take square roots.

The vectors **u** and **v** are orthogonal if and only if $\mathbf{u} \cdot \mathbf{v} = 0$, so line (14) in the proof of the triangle inequality yields the fact that the **Pythagorean formula**

$$|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2$$
(15)

holds if and only if the triangle with "adjacent side vectors" \mathbf{u} and \mathbf{v} is a right triangle with hypotenuse vector $\mathbf{u} + \mathbf{v}$ (see Figure 4.6.3).

The following theorem states a simple relationship between orthogonality and linear independence.

THEOREM 3 Orthogonality and Linear Independence

If the nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are mutually orthogonal—that is, each two of them are orthogonal—then they are linearly independent.

Proof: Suppose that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0},$$

where, as usual, c_1, c_2, \ldots, c_k are scalars. When we take the dot product of each side of this equation with \mathbf{v}_i , we find that

$$c_i \mathbf{v}_i \cdot \mathbf{v}_i = c_i |\mathbf{v}_i|^2 = 0.$$

Now $|\mathbf{v}_i| \neq 0$ because \mathbf{v}_i is a nonzero vector. It follows that $c_i = 0$. Thus $c_1 = c_2 = \cdots = c_k = 0$, and therefore the mutually orthogonal nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent.

In particular, any set of *n* mutually orthogonal nonzero vectors in \mathbb{R}^n constitutes a basis for \mathbb{R}^n . Such a basis is called an **orthogonal basis**. For instance, the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ form an orthogonal basis for \mathbb{R}^n .

Orthogonal Complements

Now we want to relate orthogonality to the solution of systems of linear equations. Consider the homogeneous linear system

$$\mathbf{A}\mathbf{x} = \mathbf{0} \tag{16}$$

of *m* equations in *n* unknowns. If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ are the row vectors of the $m \times n$ coefficient matrix **A**, then the system looks like





FIGURE 4.6.3. A right triangle in **R**^{*n*}.

274 Chapter 4 Vector Spaces

Consequently, it is clear that \mathbf{x} is a solution vector of $A\mathbf{x} = \mathbf{0}$ if and only if \mathbf{x} is orthogonal to each row vector of \mathbf{A} . But in the latter event \mathbf{x} is orthogonal to every linear combination of row vectors of \mathbf{A} because

$$\mathbf{x} \cdot (c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_m \mathbf{v}_m)$$

= $c_1 \mathbf{x} \cdot \mathbf{v}_1 + c_2 \mathbf{x} \cdot \mathbf{v}_2 + \dots + c_m \mathbf{x} \cdot \mathbf{v}_m$
= $(c_1)(0) + (c_2)(0) + \dots + (c_m)(0) = 0$

Thus we have shown that the vector \mathbf{x} in \mathbf{R}^n is a solution vector of $\mathbf{A}\mathbf{x} = \mathbf{0}$ if and only if \mathbf{x} is orthogonal to each vector in the row space $\text{Row}(\mathbf{A})$ of the matrix \mathbf{A} . This situation motivates the following definition.

DEFINITION The Orthogonal Complement of a Subspace

The vector **u** is **orthogonal** to the subspace V of \mathbb{R}^n provided that **u** is orthogonal to every vector in V. The **orthogonal complement** V^{\perp} (read "V perp") of V is the set of all those vectors in \mathbb{R}^n that are orthogonal to the subspace V.

If \mathbf{u}_1 and \mathbf{u}_2 are vectors in V^{\perp} , **v** is in V, and c_1 and c_2 are scalars, then

 $(c_1\mathbf{u}_1 + c_2\mathbf{u}_2) \cdot \mathbf{v} = c_1\mathbf{u}_1 \cdot \mathbf{v} + c_2\mathbf{u}_2 \cdot \mathbf{v}$ $= (c_1)(0) + (c_2)(0) = 0.$

Thus any linear combination of vectors in V^{\perp} is orthogonal to every vector in V and hence is a vector in V^{\perp} . Therefore *the orthogonal complement* V^{\perp} of a subspace V is itself a subspace of \mathbb{R}^n . The standard picture of two complementary subspaces V and V^{\perp} consists of an orthogonal line and plane through the origin in \mathbb{R}^3 (see Fig. 4.6.4). The proofs of the remaining parts of Theorem 4 are left to the problems.



FIGURE 4.6.4. Orthogonal complements.

THEOREM 4 Properties of Orthogonal Complements
Let V be a subspace of Rⁿ. Then
1. Its orthogonal complement V[⊥] is also a subspace of Rⁿ;
2. The only vector that lies in both V and V[⊥] is the zero vector;
3. The orthogonal complement of V[⊥] is V—that is, (V[⊥])[⊥] = V;
4. If S is a spanning set for V, then the vector u is in V[⊥] if and only if u is orthogonal to every vector in S.

4.6 Orthogonal Vectors in \mathbb{R}^n 275

In our discussion of the homogeneous linear system Ax = 0 in (16), we showed that a vector space lies in the **null space** Null(A) of A—that is, in the solution space of Ax = 0—if and only if it is orthogonal to each vector in the row space of A. In the language of orthogonal complements, this proves Theorem 5.

THEOREM 5 The Row Space and the Null Space

Let A be an $m \times n$ matrix. Then the row space Row(A) and the null space Null(A) are orthogonal complements in \mathbb{R}^n . That is,

If $V = \operatorname{Row}(\mathbf{A})$, then $V^{\perp} = \operatorname{Null}(\mathbf{A})$. (17)

Now suppose that a subspace V of \mathbb{R}^n is given, with $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ a set of vectors that span V. For instance, these vectors may form a given basis for V. Then the implication in (17) provides the following algorithm for finding a basis for the orthogonal complement V^{\perp} of V.

- 1. Let **A** be the $m \times u$ matrix with row vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$.
- Reduce A to echelon form and use the algorithm of Section 4.4 to find a basis {u₁, u₂, ..., u_k} for the solution space Null(A) of Ax = 0. Because V[⊥] = Null(A), this will be a basis for the orthogonal complement of V.
- **Example 3** Let V be the 1-dimensional subspace of \mathbb{R}^3 spanned by the vector $\mathbf{v}_1 = (1, -3, 5)$. Then

 $\mathbf{A} = \begin{bmatrix} 1 & -3 & 5 \end{bmatrix}$

and our linear system Ax = 0 consists of the single equation

$$x_1 - 3x_2 + 5x_3 = 0.$$

If $x_2 = s$ and $x_3 = t$, then $x_1 = 3s - 5t$. With s = 1 and t = 0 we get the solution vector $\mathbf{u}_1 = (3, 1, 0)$, whereas with s = 0 and t = 1 we get the solution vector $\mathbf{u}_2 = (-5, 0, 1)$. Thus the orthogonal complement V^{\perp} is the 2-dimensional subspace of \mathbf{R}^3 having $\mathbf{u}_1 = (3, 1, 0)$ and $\mathbf{u}_2 = (-5, 0, 1)$ as basis vectors.

Example 4 Let V be the 2-dimensional subspace of \mathbb{R}^5 that has $\mathbf{v}_1 = (1, 2, 1, -3, -3)$ and $\mathbf{v}_2 = (2, 5, 6, -10, -12)$ as basis vectors. The matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & -3 & -3 \\ 2 & 5 & 6 & -10 & -12 \end{bmatrix}$$

with row vectors \mathbf{v}_1 and \mathbf{v}_2 has reduced echelon form

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & -7 & 5 & 9 \\ 0 & 1 & 4 & -4 & -6 \end{bmatrix}.$$

Hence the solution space of Ax = 0 is described parametrically by

$$x_3 = r, \quad x_4 = s, \quad x_5 = t,$$

 $x_2 = -4r + 4s + 6t$
 $x_1 = 7r - 5s - 9t.$

Then the choice

r = 1,	s=0,	t = 0	yields	$\mathbf{u}_1 = (7, -4, 1, 0, 0);$
r = 0,	s = 1,	t = 0	yields	$\mathbf{u}_2 = (-5, 4, 0, 1, 0);$
r = 0,	s = 0,	t = 1	yields	$\mathbf{u}_3 = (-9, 6, 0, 0, 1).$

Thus the orthogonal complement V^{\perp} is the 3-dimensional subspace of \mathbb{R}^5 with basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

Observe that dim $V + \dim V^{\perp} = 3$ in Example 3, but dim $V + \dim V^{\perp} = 5$ in Example 4. It is no coincidence that in each case the dimensions of V and V^{\perp} add up to the dimension n of the Euclidean space containing them. To see why, suppose that V is a subspace of \mathbb{R}^n and let \mathbb{A} be an $m \times n$ matrix whose row vectors span V. Then Equation (12) in Section 4.5 implies that

$$\operatorname{rank}(\mathbf{A}) + \operatorname{dim} \operatorname{Null}(\mathbf{A}) = n$$

But

 $\dim V = \dim \operatorname{Row}(\mathbf{A}) = \operatorname{rank}(\mathbf{A})$

and

$$\dim V^{\perp} = \dim \operatorname{Null}(\mathbf{A})$$

by Theorem 5, so it follows that

$$\dim V + \dim V^{\perp} = n.$$

Moreover, it should be apparent intuitively that if

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$$
 is a basis for V

and

 $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is a basis for V^{\perp} ,

then

 $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is a basis for \mathbf{R}^n .

That is, the union of a basis for V and a basis for V^{\perp} is a basis for \mathbb{R}^n . In Problem 34 of this section we ask you to prove that this is so.

4.6 Problems

In Problems 1-4, determine whether the given vectors are mutually orthogonal.

- 1. $\mathbf{v}_1 = (2, 1, 2, 1), \mathbf{v}_2 = (3, -6, 1, -2), \mathbf{v}_3 = (3, -1, -5, 5)$
- **2.** $\mathbf{v}_1 = (3, -2, 3, -4), \mathbf{v}_2 = (6, 3, 4, 6),$ $\mathbf{v}_3 = (17, -12, -21, 3)$
- 3. $\mathbf{v}_1 = (5, 2, -4, -1), \mathbf{v}_2 = (3, -5, 1, 1), \mathbf{v}_3 = (3, 0, 8, -17)$
- 4. $\mathbf{v}_1 = (1, 2, 3, -2, 1), \mathbf{v}_2 = (3, 2, 3, 6, -4), \mathbf{v}_3 = (6, 2, -4, 1, 4)$

In Problems 5–8, the three vertices A, B, and C of a triangle are given. Prove that each triangle is a right triangle by showing that its sides a, b, and c satisfy the Pythagorean relation $a^2 + b^2 = c^2$.

- **5.** *A*(6, 6, 5, 8), *B*(6, 8, 6, 5), *C*(5, 7, 4, 6)
- 6. A(3, 5, 1, 3), B(4, 2, 6, 4), C(1, 3, 4, 2)
- 7. A(4, 5, 3, 5, -1), B(3, 4, -1, 4, 4), C(1, 3, 1, 3, 1)
- 8. A(2, 8, -3, -1, 2), B(-2, 5, 6, 2, 12), C(-5, 3, 2, -3, 5)
- **9–12.** Find the acute angles (in degrees) of each of the right triangles of Problems 5–8, respectively.

(18)