

MATH-224 I: Linear Algebra



Norms and inner products

Reference: Linear Algebra and Its Application, David C Lay, 4th ed.

INNER PRODUCT, LENGTH, AND ORTHOGONALITY

Geometric concepts of length, distance, and perpendicularity, which are well known for \mathbb{R}^2 and \mathbb{R}^3 , are defined here for \mathbb{R}^n . These concepts provide powerful geometric tools for solving many applied problems, including the least-squares problems mentioned above. All three notions are defined in terms of the inner product of two vectors.

The Inner Product

If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , then we regard \mathbf{u} and \mathbf{v} as $n \times 1$ matrices. The transpose \mathbf{u}^T is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T \mathbf{v}$ is a 1×1 matrix, which we write as a single real number (a scalar) without brackets. The number $\mathbf{u}^T \mathbf{v}$ is called the **inner product** of \mathbf{u} and \mathbf{v} , and often it is written as $\mathbf{u} \cdot \mathbf{v}$. This inner product, mentioned in the exercises for Section 2.1, is also referred to as a **dot product**. If

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then the inner product of \mathbf{u} and \mathbf{v} is

$$[u_1 \ u_2 \ \cdots \ u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

EXAMPLE 1 Compute $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{u}$ for $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

SOLUTION

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = [2 \quad -5 \quad -1] \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v} \cdot \mathbf{u} = \mathbf{v}^T \mathbf{u} = [3 \quad 2 \quad -3] \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = (3)(2) + (2)(-5) + (-3)(-1) = -1 \quad \blacksquare$$

Let \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n , and let c be a scalar. Then

- a. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- b. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- c. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
- d. $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Properties (b) and (c) can be combined several times to produce the following useful rule:

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

The Length of a Vector

If \mathbf{v} is in \mathbb{R}^n , with entries v_1, \dots, v_n , then the square root of $\mathbf{v} \cdot \mathbf{v}$ is defined because $\mathbf{v} \cdot \mathbf{v}$ is nonnegative.

The **length** (or **norm**) of \mathbf{v} is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$$

Suppose \mathbf{v} is in \mathbb{R}^2 , say, $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. If we identify \mathbf{v} with a geometric point in the plane, as usual, then $\|\mathbf{v}\|$ coincides with the standard notion of the length of the line segment from the origin to \mathbf{v} . This follows from the Pythagorean Theorem applied to a triangle such as the one in Fig. 1.

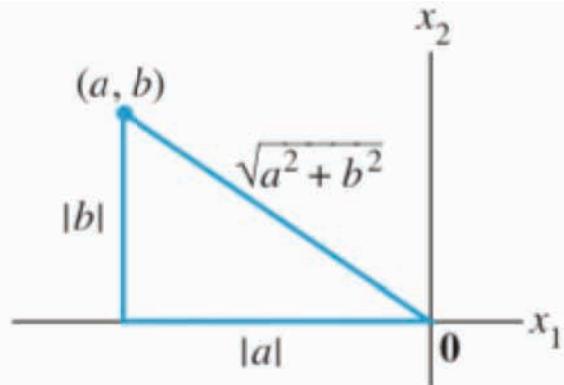


FIGURE 1

Interpretation of $\|\mathbf{v}\|$ as length.

A similar calculation with the diagonal of a rectangular box shows that the definition of length of a vector \mathbf{v} in \mathbb{R}^3 coincides with the usual notion of length.

For any scalar c , the length of $c\mathbf{v}$ is $|c|$ times the length of \mathbf{v} . That is,

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$$

(To see this, compute $\|c\mathbf{v}\|^2 = (c\mathbf{v}) \cdot (c\mathbf{v}) = c^2 \mathbf{v} \cdot \mathbf{v} = c^2 \|\mathbf{v}\|^2$ and take square roots.)

A vector whose length is 1 is called a **unit vector**. If we *divide* a nonzero vector \mathbf{v} by its length—that is, multiply by $1/\|\mathbf{v}\|$ —we obtain a unit vector \mathbf{u} because the length of \mathbf{u} is $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$. The process of creating \mathbf{u} from \mathbf{v} is sometimes called **normalizing** \mathbf{v} , and we say that \mathbf{u} is *in the same direction* as \mathbf{v} .

Several examples that follow use the space-saving notation for (column) vectors.

EXAMPLE 2 Let $\mathbf{v} = (1, -2, 2, 0)$. Find a unit vector \mathbf{u} in the same direction as \mathbf{v} .

SOLUTION First, compute the length of \mathbf{v} :

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9$$

$$\|\mathbf{v}\| = \sqrt{9} = 3$$

Then, multiply \mathbf{v} by $1/\|\mathbf{v}\|$ to obtain

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|} \mathbf{v} = \frac{1}{3} \mathbf{v} = \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

To check that $\|\mathbf{u}\| = 1$, it suffices to show that $\|\mathbf{u}\|^2 = 1$.

$$\begin{aligned}\|\mathbf{u}\|^2 &= \mathbf{u} \cdot \mathbf{u} = \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + (0)^2 \\ &= \frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0 = 1\end{aligned}$$



EXAMPLE 3 Let W be the subspace of \mathbb{R}^2 spanned by $\mathbf{x} = (\frac{2}{3}, 1)$. Find a unit vector \mathbf{z} that is a basis for W .

SOLUTION W consists of all multiples of \mathbf{x} , as in Fig. 2(a). Any nonzero vector in W is a basis for W . To simplify the calculation, “scale” \mathbf{x} to eliminate fractions. That is, multiply \mathbf{x} by 3 to get

$$\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Now compute $\|\mathbf{y}\|^2 = 2^2 + 3^2 = 13$, $\|\mathbf{y}\| = \sqrt{13}$, and normalize \mathbf{y} to get

$$\mathbf{z} = \frac{1}{\sqrt{13}} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

See Fig. 2(b). Another unit vector is $(-2/\sqrt{13}, -3/\sqrt{13})$.

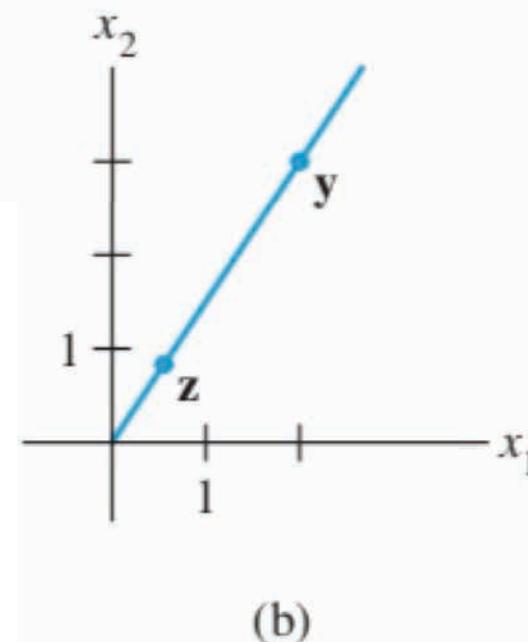
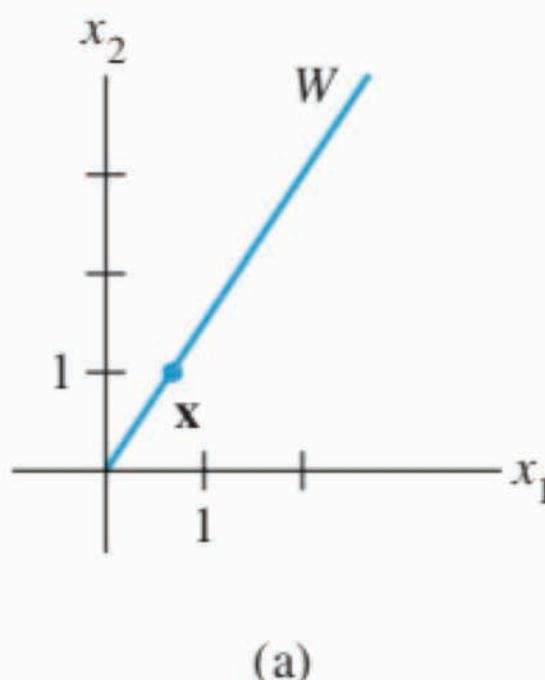
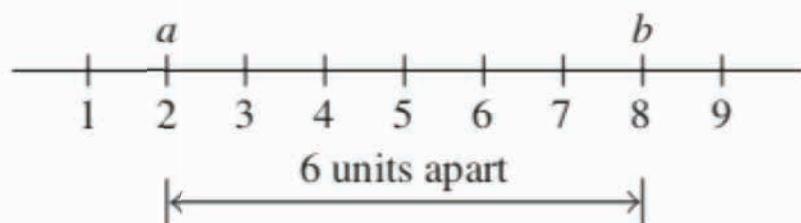


FIGURE 2

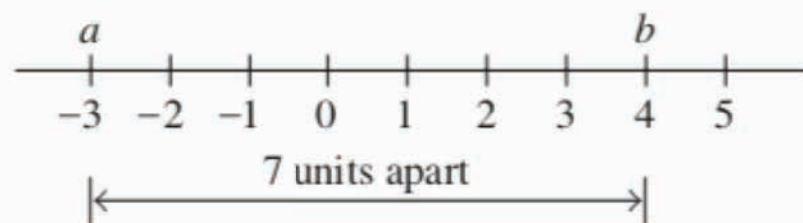
Normalizing a vector to produce a unit vector.

Distance in \mathbb{R}^n

We are ready now to describe how close one vector is to another. Recall that if a and b are real numbers, the distance on the number line between a and b is the number $|a - b|$. Two examples are shown in Fig. 3. This definition of distance in \mathbb{R} has a direct analogue in \mathbb{R}^n .



$$|2 - 8| = |-6| = 6 \quad \text{or} \quad |8 - 2| = |6| = 6$$



$$|(-3) - 4| = |-7| = 7 \quad \text{or} \quad |4 - (-3)| = |7| = 7$$

FIGURE 3 Distances in \mathbb{R} .

For \mathbf{u} and \mathbf{v} in \mathbb{R}^n , the **distance between \mathbf{u} and \mathbf{v}** , written as $\text{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

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

EXAMPLE 4 Compute the distance between the vectors $\mathbf{u} = (7, 1)$ and $\mathbf{v} = (3, 2)$.

SOLUTION Calculate

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

The vectors \mathbf{u} , \mathbf{v} , and $\mathbf{u} - \mathbf{v}$ are shown in Fig. 4. When the vector $\mathbf{u} - \mathbf{v}$ is added to \mathbf{v} , the result is \mathbf{u} . Notice that the parallelogram in Fig. 4 shows that the distance from \mathbf{u} to \mathbf{v} is the same as the distance from $\mathbf{u} - \mathbf{v}$ to $\mathbf{0}$. ■

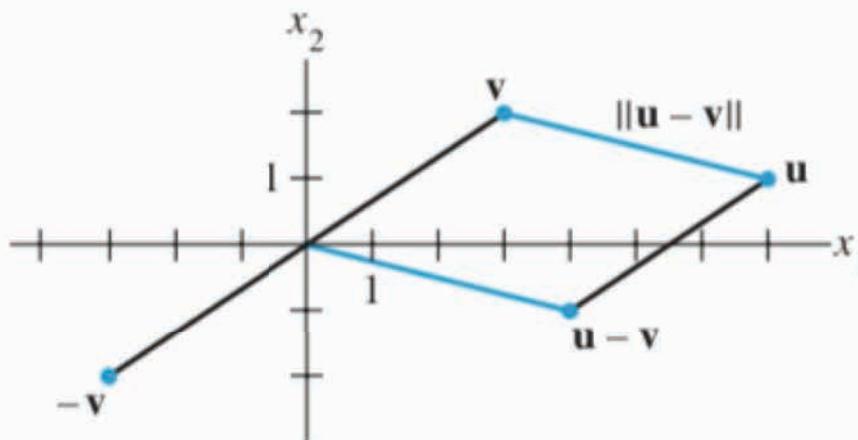


FIGURE 4 The distance between \mathbf{u} and \mathbf{v} is the length of $\mathbf{u} - \mathbf{v}$.

EXAMPLE 5 If $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$, then

$$\begin{aligned}\text{dist}(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}\end{aligned}$$

Orthogonal Vectors

The rest of this chapter depends on the fact that the concept of perpendicular lines in ordinary Euclidean geometry has an analogue in \mathbb{R}^n .

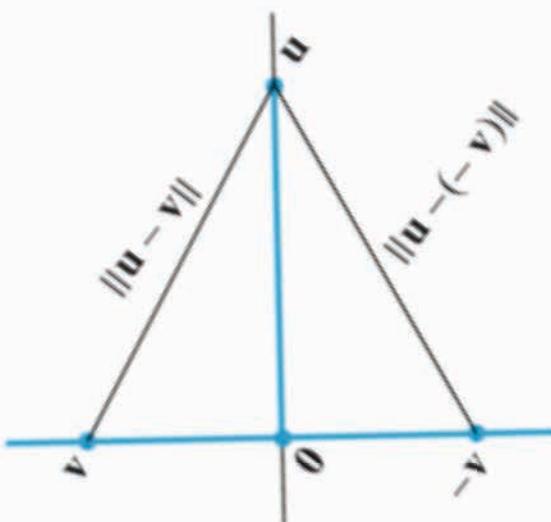
Consider \mathbb{R}^2 or \mathbb{R}^3 and two lines through the origin determined by vectors \mathbf{u} and \mathbf{v} . The two lines shown in Fig. 5 are geometrically perpendicular if and only if the distance from \mathbf{u} to \mathbf{v} is the same as the distance from \mathbf{u} to $-\mathbf{v}$. This is the same as requiring the squares of the distances to be the same. Now

$$\begin{aligned}[\text{dist}(\mathbf{u}, -\mathbf{v})]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\ &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) && \text{Theorem 1(b)} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} && \text{Theorem 1(a), (b)} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} && \text{Theorem 1(a)}\end{aligned} \tag{1}$$

The same calculations with \mathbf{v} and $-\mathbf{v}$ interchanged show that

$$\begin{aligned} [\text{dist}(\mathbf{u}, \mathbf{v})]^2 &= \|\mathbf{u}\|^2 + \|-\mathbf{v}\|^2 + 2\mathbf{u} \cdot (-\mathbf{v}) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v} \end{aligned}$$

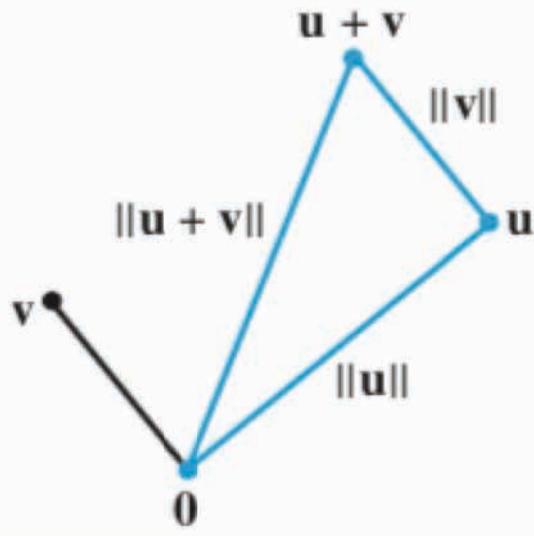
The two squared distances are equal if and only if $2\mathbf{u} \cdot \mathbf{v} = -2\mathbf{u} \cdot \mathbf{v}$, which happens if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.



Two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

The Pythagorean Theorem

Two vectors \mathbf{u} and \mathbf{v} are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.



Orthogonal Complements

To provide practice using inner products, we introduce a concept here that will be of use in Section 6.3 and elsewhere in the chapter. If a vector \mathbf{z} is orthogonal to every vector in a subspace W of \mathbb{R}^n , then \mathbf{z} is said to be **orthogonal to W** . The set of all vectors \mathbf{z} that are orthogonal to W is called the **orthogonal complement** of W and is denoted by W^\perp (and read as “ W perpendicular” or simply “ W perp”).

EXAMPLE 6 Let W be a plane through the origin in \mathbb{R}^3 , and let L be the line through the origin and perpendicular to W . If \mathbf{z} and \mathbf{w} are nonzero, \mathbf{z} is on L , and \mathbf{w} is in W , then the line segment from $\mathbf{0}$ to \mathbf{z} is perpendicular to the line segment from $\mathbf{0}$ to \mathbf{w} ; that is, $\mathbf{z} \cdot \mathbf{w} = 0$. See Fig. 7. So each vector on L is orthogonal to every \mathbf{w} in W . In fact, L consists of *all* vectors that are orthogonal to the \mathbf{w} 's in W , and W consists of all vectors orthogonal to the \mathbf{z} 's in L . That is,

$$L = W^\perp \quad \text{and} \quad W = L^\perp$$

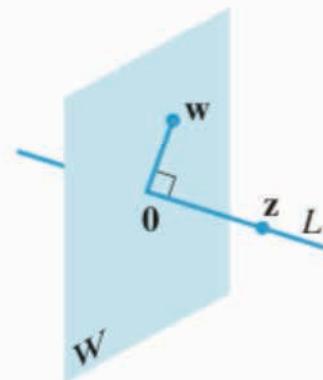


FIGURE 7

A plane and line through $\mathbf{0}$ as orthogonal complements.

1. A vector \mathbf{x} is in W^\perp if and only if \mathbf{x} is orthogonal to every vector in a set that spans W .
2. W^\perp is a subspace of \mathbb{R}^n .

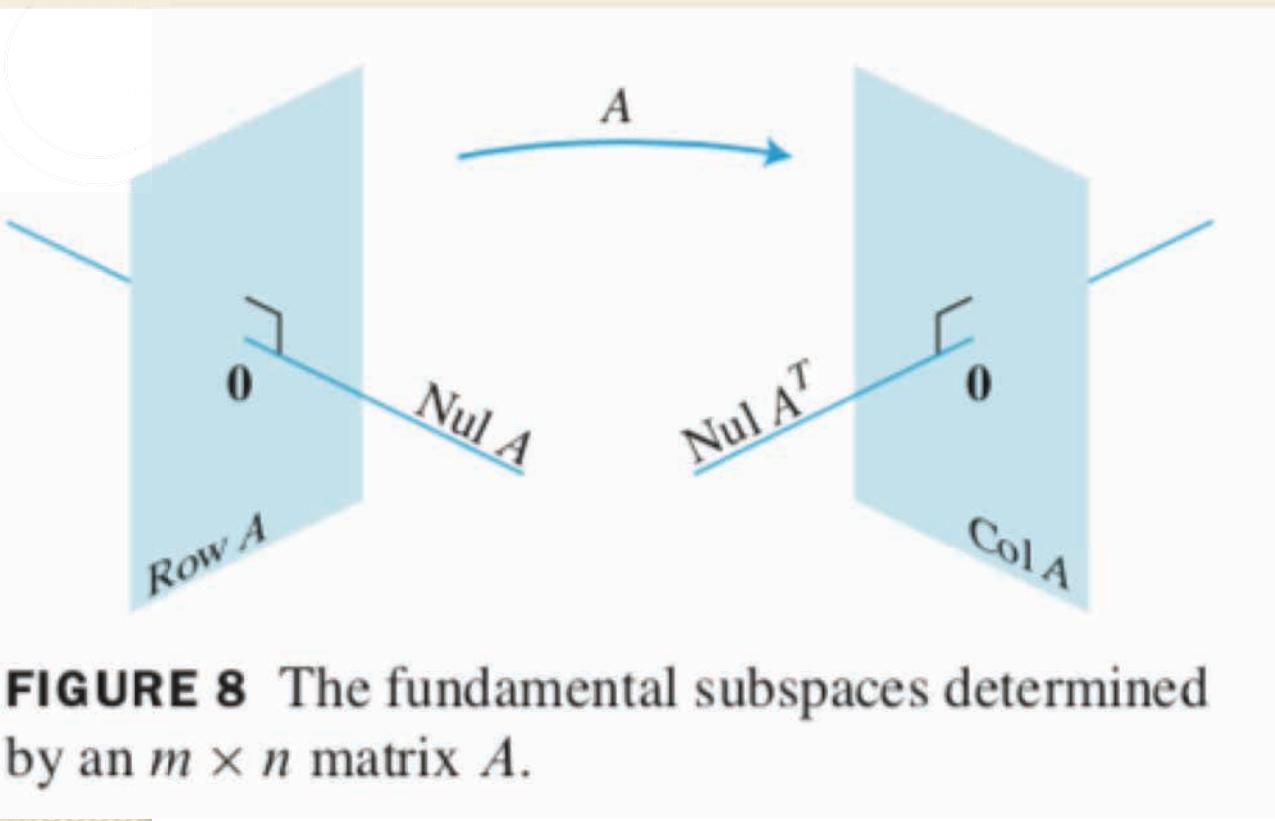


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A .

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

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

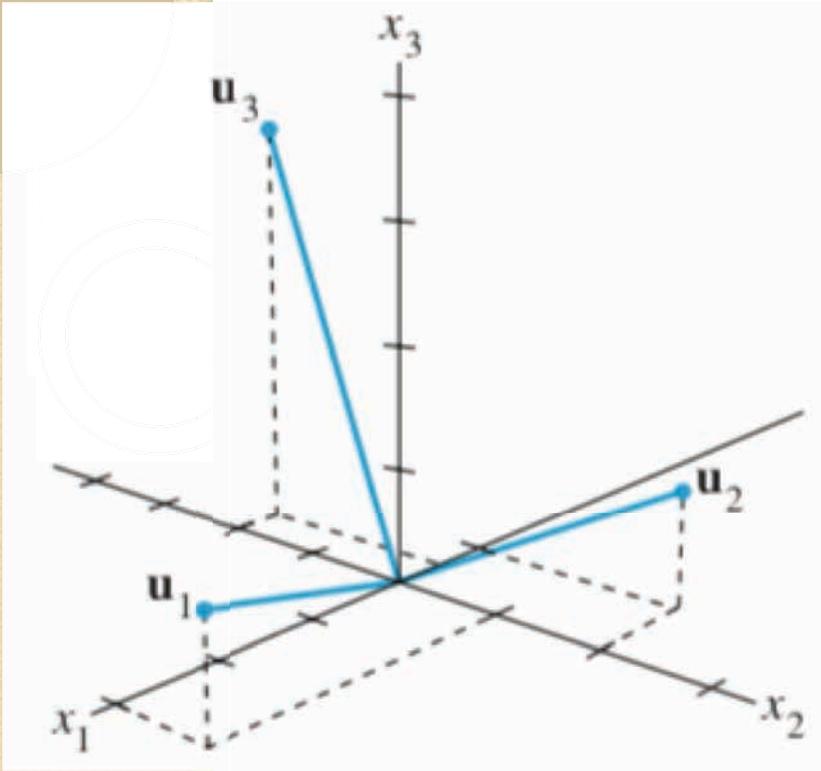
SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

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

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

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

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. ■



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. ■

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

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

$$\mathbf{y} \cdot \mathbf{u}_1 = 11, \quad \mathbf{y} \cdot \mathbf{u}_2 = -12, \quad \mathbf{y} \cdot \mathbf{u}_3 = -33$$

$$\mathbf{u}_1 \cdot \mathbf{u}_1 = 11, \quad \mathbf{u}_2 \cdot \mathbf{u}_2 = 6, \quad \mathbf{u}_3 \cdot \mathbf{u}_3 = 33/2$$

By Theorem 5,

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



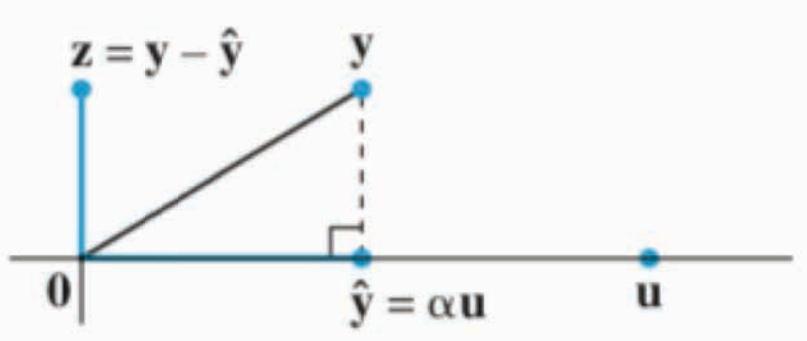
An Orthogonal Projection

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

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

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

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

$$\mathbf{u} \cdot \mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

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,

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

$\begin{matrix} \uparrow & \uparrow & \uparrow \\ \mathbf{y} & \hat{\mathbf{y}} & (\mathbf{y} - \hat{\mathbf{y}}) \end{matrix}$

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

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

■

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

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

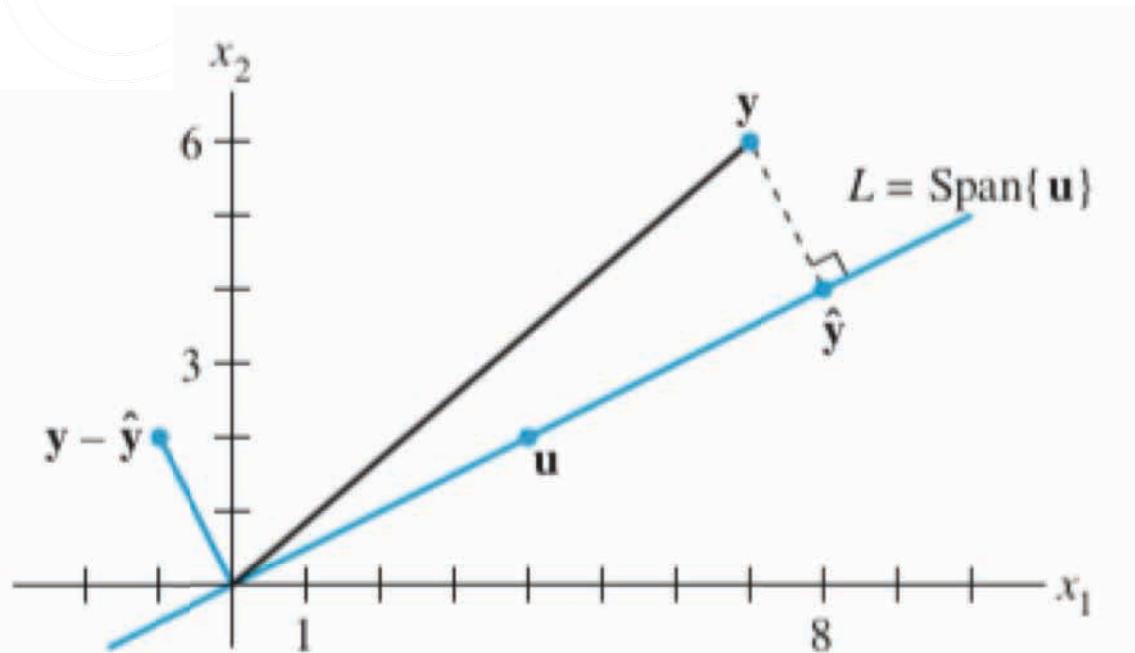


FIGURE 3 The orthogonal projection of \mathbf{y} onto a line L through the origin.

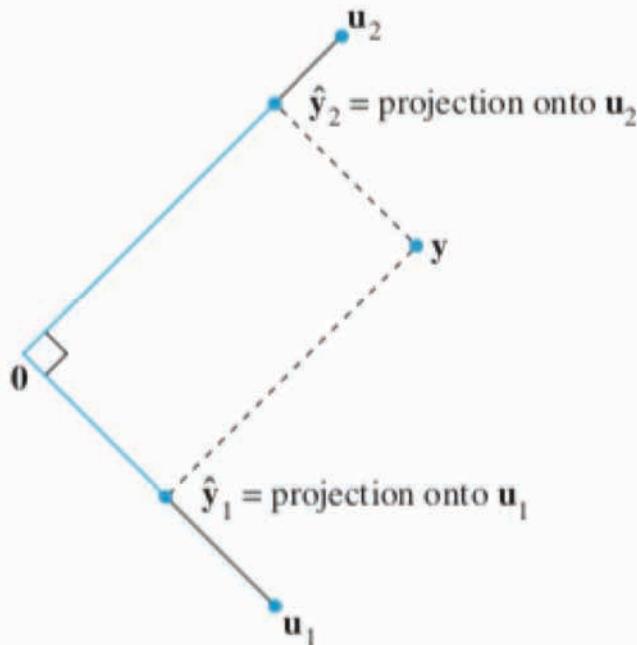
A Geometric Interpretation of Theorem 5

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.



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,

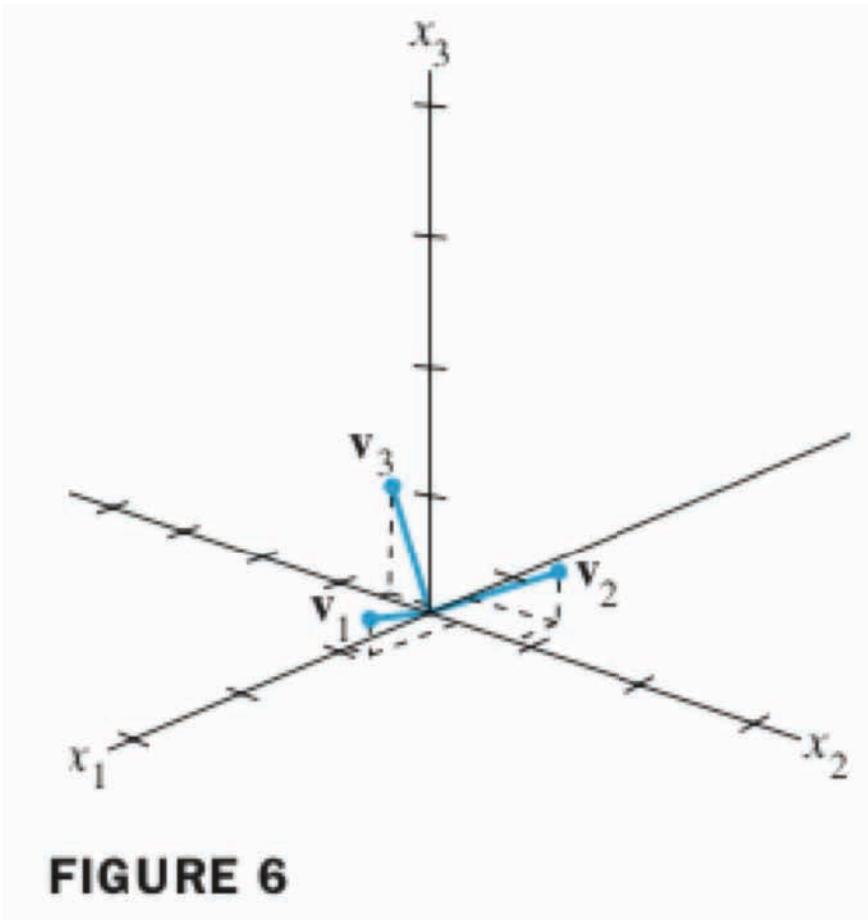
$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$$

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

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

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. ■



ORTHOGONAL PROJECTIONS

The orthogonal projection of a point in \mathbb{R}^2 onto a line through the origin has an important analogue in \mathbb{R}^n . Given a vector \mathbf{y} and a subspace W in \mathbb{R}^n , there is a vector $\hat{\mathbf{y}}$ in W such that (1) $\hat{\mathbf{y}}$ is the unique vector in W for which $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W , and (2) $\hat{\mathbf{y}}$ is the unique vector in W closest to \mathbf{y} . See Fig. 1. These two properties of $\hat{\mathbf{y}}$ provide the key to finding least-squares solutions of linear systems, mentioned in the introductory example for this chapter. The full story will be told in Section 6.5.

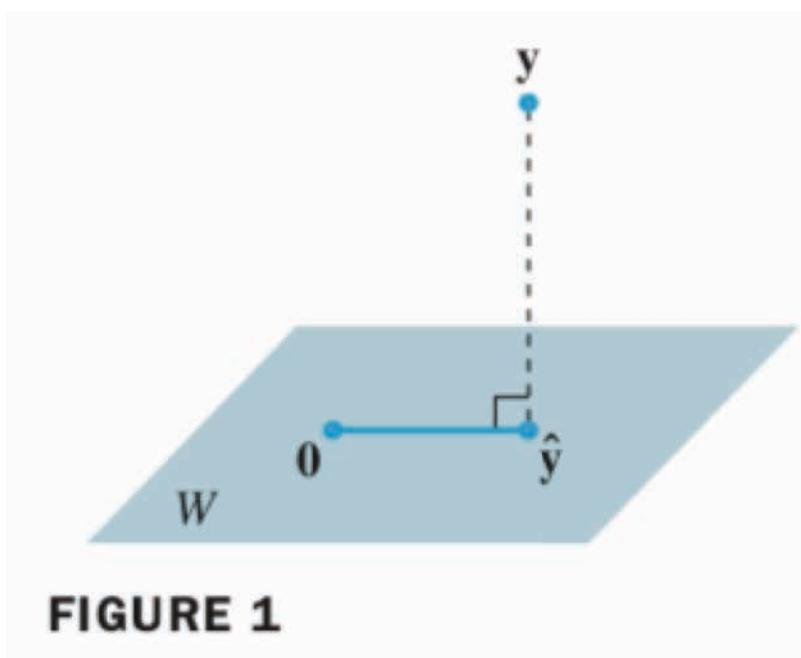


FIGURE 1

To prepare for the first theorem, observe that whenever a vector \mathbf{y} is written as a linear combination of vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ in \mathbb{R}^n , the terms in the sum for \mathbf{y} can be grouped into two parts so that \mathbf{y} can be written as

$$\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$$

where \mathbf{z}_1 is a linear combination of some of the \mathbf{u}_i and \mathbf{z}_2 is a linear combination of the rest of the \mathbf{u}_i . This idea is particularly useful when $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthogonal basis. Recall from Section 6.1 that W^\perp denotes the set of all vectors orthogonal to a subspace W .

EXAMPLE 1 Let $\{\mathbf{u}_1, \dots, \mathbf{u}_5\}$ be an orthogonal basis for \mathbb{R}^5 and let

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Consider the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, and write \mathbf{y} as the sum of a vector \mathbf{z}_1 in W and a vector \mathbf{z}_2 in W^\perp .

SOLUTION Write

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$

where $\mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ is in $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$

and $\mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5$ is in $\text{Span}\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

To show that \mathbf{z}_2 is in W^\perp , it suffices to show that \mathbf{z}_2 is orthogonal to the vectors in the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W . (See Section 6.1.) Using properties of the inner product, compute

$$\begin{aligned}\mathbf{z}_2 \cdot \mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5) \cdot \mathbf{u}_1 \\ &= c_3\mathbf{u}_3 \cdot \mathbf{u}_1 + c_4\mathbf{u}_4 \cdot \mathbf{u}_1 + c_5\mathbf{u}_5 \cdot \mathbf{u}_1 \\ &= 0\end{aligned}$$

because \mathbf{u}_1 is orthogonal to \mathbf{u}_3 , \mathbf{u}_4 , and \mathbf{u}_5 . A similar calculation shows that $\mathbf{z}_2 \cdot \mathbf{u}_2 = 0$. Thus \mathbf{z}_2 is in W^\perp . ■

The Orthogonal Decomposition Theorem

Let W be a subspace of \mathbb{R}^n . Then each \mathbf{y} in \mathbb{R}^n can be written uniquely in the form

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

where $\hat{\mathbf{y}}$ is in W and \mathbf{z} is in W^\perp . In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W , then

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \mathbf{u}_p \quad (2)$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of \mathbf{y} onto W** and often is written as $\text{proj}_W \mathbf{y}$. See Fig. 2. When W is a one-dimensional subspace, the formula for $\hat{\mathbf{y}}$ matches the formula given in Section 6.2.

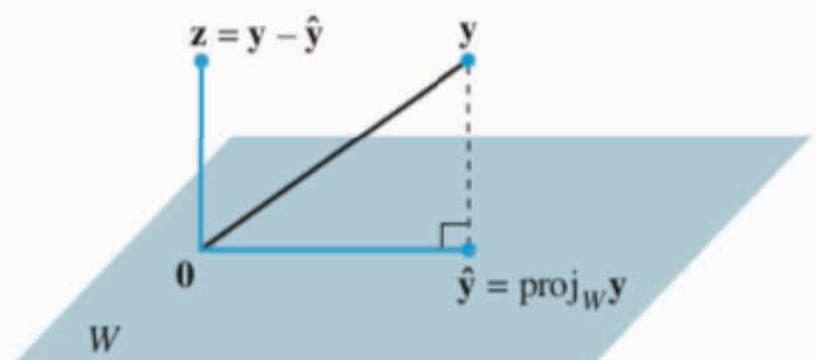


FIGURE 2 The orthogonal projection of \mathbf{y} onto W .

PROOF Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be any orthogonal basis for W , and define $\hat{\mathbf{y}}$ by (2).¹ Then $\hat{\mathbf{y}}$ is in W because $\hat{\mathbf{y}}$ is a linear combination of the basis $\mathbf{u}_1, \dots, \mathbf{u}_p$. Let $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Since \mathbf{u}_1 is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$, it follows from (2) that

$$\begin{aligned}\mathbf{z} \cdot \mathbf{u}_1 &= (\mathbf{y} - \hat{\mathbf{y}}) \cdot \mathbf{u}_1 = \mathbf{y} \cdot \mathbf{u}_1 - \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 \cdot \mathbf{u}_1 = 0 - \dots - 0 \\ &= \mathbf{y} \cdot \mathbf{u}_1 - \mathbf{y} \cdot \mathbf{u}_1 = 0\end{aligned}$$

Thus \mathbf{z} is orthogonal to \mathbf{u}_1 . Similarly, \mathbf{z} is orthogonal to each \mathbf{u}_j in the basis for W . Hence \mathbf{z} is orthogonal to every vector in W . That is, \mathbf{z} is in W^\perp .

To show that the decomposition in (1) is unique, suppose \mathbf{y} can also be written as $\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$, with $\hat{\mathbf{y}}_1$ in W and \mathbf{z}_1 in W^\perp . Then $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ (since both sides equal \mathbf{y}), and so

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$$

This equality shows that the vector $\mathbf{v} = \hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ is in W and in W^\perp (because \mathbf{z}_1 and \mathbf{z} are both in W^\perp , and W^\perp is a subspace). Hence $\mathbf{v} \cdot \mathbf{v} = 0$, which shows that $\mathbf{v} = \mathbf{0}$. This proves that $\hat{\mathbf{y}} = \hat{\mathbf{y}}_1$ and also $\mathbf{z}_1 = \mathbf{z}$. ■

EXAMPLE 2 Let $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$

is an orthogonal basis for $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write \mathbf{y} as the sum of a vector in W and a vector orthogonal to W .

SOLUTION The orthogonal projection of \mathbf{y} onto W is

$$\begin{aligned}\hat{\mathbf{y}} &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 \\ &= \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}\end{aligned}$$

Also

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Theorem 8 ensures that $\mathbf{y} - \hat{\mathbf{y}}$ is in W^\perp . To check the calculations, however, it is a good idea to verify that $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to both \mathbf{u}_1 and \mathbf{u}_2 and hence to all of W . The desired decomposition of \mathbf{y} is

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

■

A Geometric Interpretation of the Orthogonal Projection

When W is a one-dimensional subspace, the formula (2) for $\text{proj}_W \mathbf{y}$ contains just one term. Thus, when $\dim W > 1$, each term in (2) is itself an orthogonal projection of \mathbf{y} onto a one-dimensional subspace spanned by one of the \mathbf{u} 's in the basis for W . Figure 3 illustrates this when W is a subspace of \mathbb{R}^3 spanned by \mathbf{u}_1 and \mathbf{u}_2 . Here $\hat{\mathbf{y}}_1$ and $\hat{\mathbf{y}}_2$ denote the projections of \mathbf{y} onto the lines spanned by \mathbf{u}_1 and \mathbf{u}_2 , respectively. The orthogonal projection $\hat{\mathbf{y}}$ of \mathbf{y} onto W is the sum of the projections of \mathbf{y} onto one-dimensional subspaces that are orthogonal to each other. The vector $\hat{\mathbf{y}}$ in Fig. 3 corresponds to the vector \mathbf{y} in Fig. 4 of Section 6.2, because now it is $\hat{\mathbf{y}}$ that is in W .

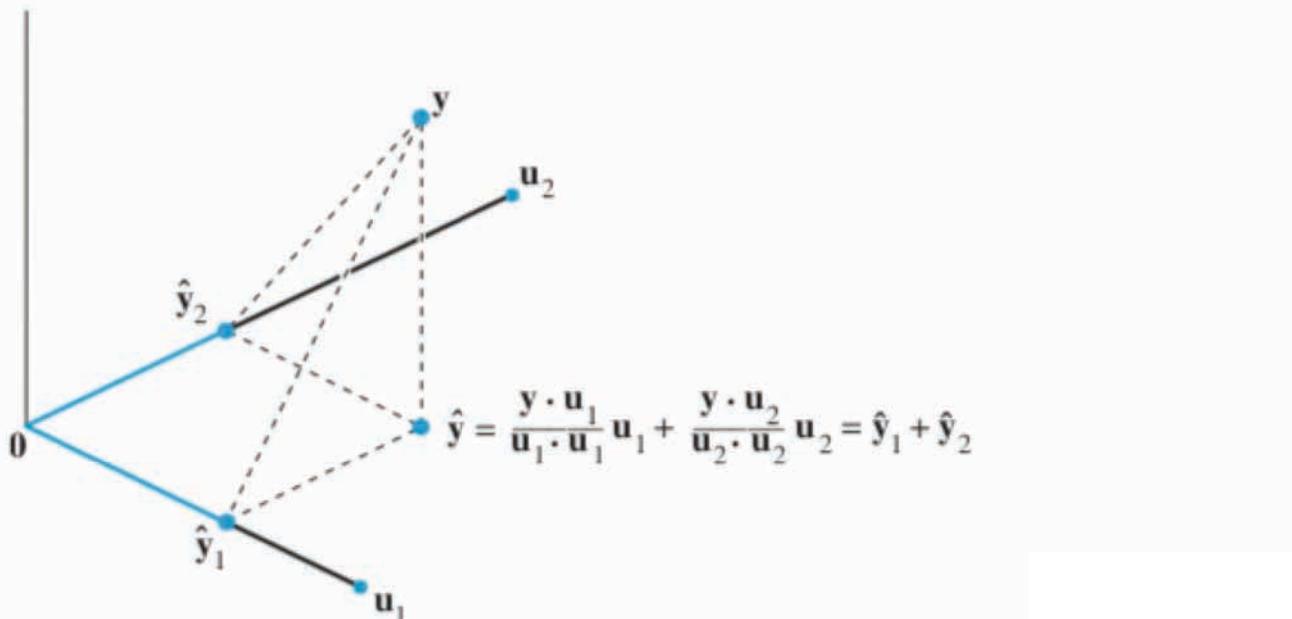


FIGURE 3 The orthogonal projection of \mathbf{y} is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

THE GRAM–SCHMIDT PROCESS

The Gram–Schmidt process is a simple algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of \mathbb{R}^n . The first two examples of the process are aimed at hand calculation.

EXAMPLE 1 Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for W .

SOLUTION The subspace W is shown in Fig. 1, along with \mathbf{x}_1 , \mathbf{x}_2 , and the projection \mathbf{p} of \mathbf{x}_2 onto \mathbf{x}_1 . The component of \mathbf{x}_2 orthogonal to \mathbf{x}_1 is $\mathbf{x}_2 - \mathbf{p}$, which is in W because it is formed from \mathbf{x}_2 and a multiple of \mathbf{x}_1 . Let $\mathbf{v}_1 = \mathbf{x}_1$ and

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1} \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal set of nonzero vectors in W . Since $\dim W = 2$, the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for W . ■

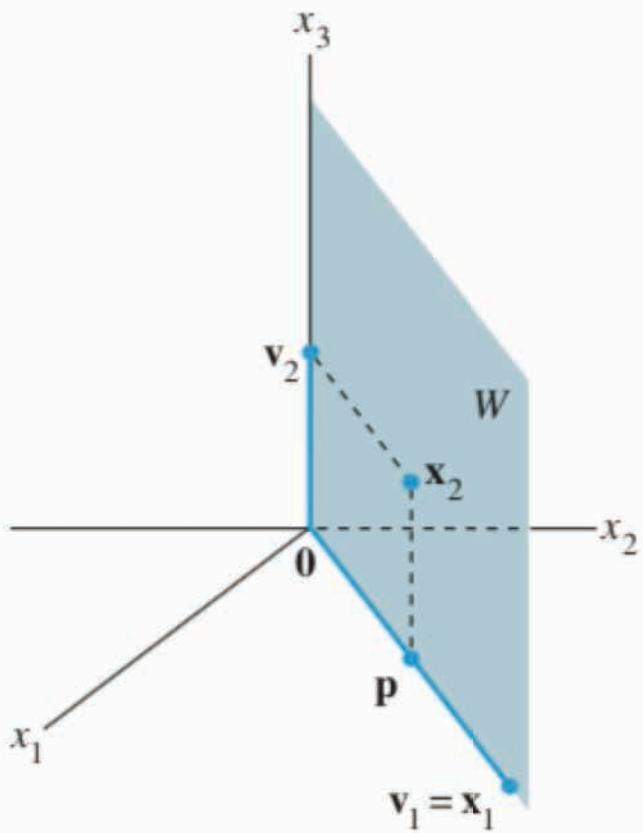


FIGURE 1

Construction of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

EXAMPLE 2 Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of \mathbb{R}^4 . Construct an orthogonal basis for W .

SOLUTION

Step 1. Let $\mathbf{v}_1 = \mathbf{x}_1$ and $W_1 = \text{Span}\{\mathbf{x}_1\} = \text{Span}\{\mathbf{v}_1\}$.

Step 2. Let \mathbf{v}_2 be the vector produced by subtracting from \mathbf{x}_2 its projection onto the subspace W_1 . That is, let

$$\begin{aligned}\mathbf{v}_2 &= \mathbf{x}_2 - \text{proj}_{W_1} \mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 \quad \text{Since } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}\end{aligned}$$

Step 3. Let \mathbf{v}_3 be the vector produced by subtracting from \mathbf{x}_3 its projection onto the subspace W_2 . Use the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}'_2\}$ to compute this projection onto W_2 :

$$\text{proj}_{W_2} \mathbf{x}_3 = \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{x}_3 \cdot \mathbf{v}'_2}{\mathbf{v}'_2 \cdot \mathbf{v}'_2} \mathbf{v}'_2 = \frac{2}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12} \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Then \mathbf{v}_3 is the component of \mathbf{x}_3 orthogonal to W_2 , namely,

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proj}_{W_2} \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

See Fig. 2 for a diagram of this construction. Observe that \mathbf{v}_3 is in W , because \mathbf{x}_3 and $\text{proj}_{W_2} \mathbf{x}_3$ are both in W . Thus $\{\mathbf{v}_1, \mathbf{v}'_2, \mathbf{v}_3\}$ is an orthogonal set of nonzero vectors and hence a linearly independent set in W . Note that W is three-dimensional since it was defined by a basis of three vectors. Hence, by the Basis Theorem in Section 4.5, $\{\mathbf{v}_1, \mathbf{v}'_2, \mathbf{v}_3\}$ is an orthogonal basis for W . ■

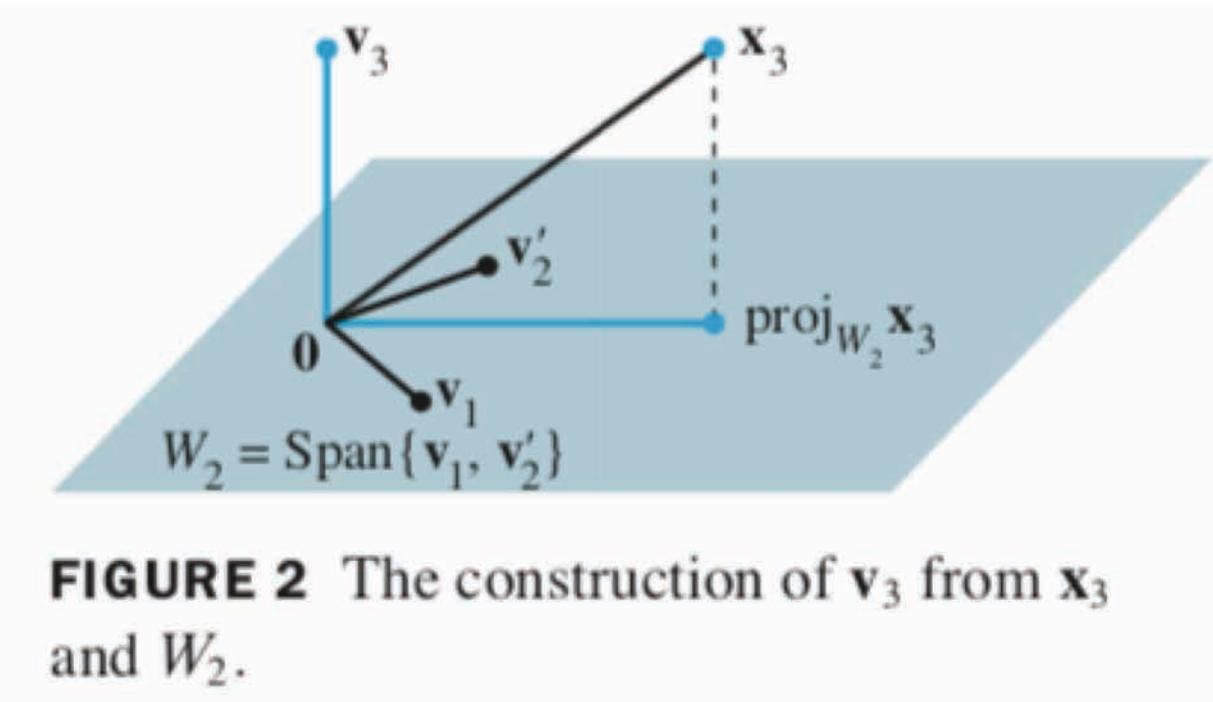


FIGURE 2 The construction of v_3 from x_3 and W_2 .

The Gram–Schmidt Process

Given a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ for a nonzero subspace W of \mathbb{R}^n , define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2$$

⋮

$$\mathbf{v}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}} \mathbf{v}_{p-1}$$

Then $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is an orthogonal basis for W . In addition

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \quad \text{for } 1 \leq k \leq p \tag{1}$$