9. A matrix of the form

$$\begin{bmatrix} 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & e & 0 & f \\ g & 0 & h & 0 \end{bmatrix}$$

cannot be invertible.

10. A matrix of the form

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

such that $ae - bd = 0$ cannot be invertible.

4.11 Inner Product Spaces

We now extend the familiar idea of a dot product for geometric vectors to an arbitrary vector space V. This enables us to associate a magnitude with each vector in V and also to define the angle between two vectors in V. The major reason that we want to do this is that, as we will see in the next section, it enables us to construct orthogonal bases in a vector space, and the use of such a basis often simplifies the representation of vectors. We begin with a brief review of the dot product.

Let $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ be two arbitrary vectors in $\mathbb{R}^3$, and consider the corresponding geometric vectors

$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}, \qquad \mathbf{y} = y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}.$$

The dot product of $\mathbf{x}$ and $\mathbf{y}$ can be defined in terms of the components of these vectors as

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3. \tag{4.11.1}$$

An equivalent geometric definition of the dot product is

$$\mathbf{x} \cdot \mathbf{y} = ||\mathbf{x}||\ ||\mathbf{y}|| \cos\theta, \tag{4.11.2}$$

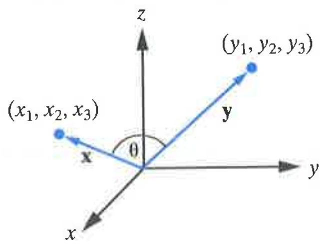

Figure 4.11.1: Defining the dot product in $\mathbb{R}^3$.

where $||\mathbf{x}||$, $||\mathbf{y}||$ denote the lengths of $\mathbf{x}$ and $\mathbf{y}$ respectively, and $0 \le \theta \le \pi$ is the angle between them. (See Figure 4.11.1.)

Taking $\mathbf{y} = \mathbf{x}$ in Equations (4.11.1) and (4.11.2) yields

$$||\mathbf{x}||^2 = \mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 + x_3^2,$$

so that the length of a geometric vector is given in terms of the dot product by

$$||\mathbf{x}|| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Furthermore, from Equation (4.11.2), the angle between any two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ is

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||\ ||\mathbf{y}||}, \tag{4.11.3}$$

which implies that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal (perpendicular) if and only if

$$\mathbf{x} \cdot \mathbf{y} = 0.$$

In a general vector space, we do not have a geometrical picture to guide us in defining the dot product, hence our definitions must be purely algebraic. We begin by considering the vector space $\mathbb{R}^n$, since there is a natural way to extend Equation (4.11.1) in this case. Before proceeding, we note that from now on we will use the standard terms *inner product* and *norm* in place of dot product and length, respectively.

9. A matrix of the form

$$\begin{bmatrix} 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & e & 0 & f \\ g & 0 & h & 0 \end{bmatrix}$$

cannot be invertible.

10. A matrix of the form

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

such that $ae - bd = 0$ cannot be invertible.

4.11 Inner Product Spaces

We now extend the familiar idea of a dot product for geometric vectors to an arbitrary vector space V. This enables us to associate a magnitude with each vector in V and also to define the angle between two vectors in V. The major reason that we want to do this is that, as we will see in the next section, it enables us to construct orthogonal bases in a vector space, and the use of such a basis often simplifies the representation of vectors. We begin with a brief review of the dot product.

Let $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ be two arbitrary vectors in $\mathbb{R}^3$, and consider the corresponding geometric vectors

$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}, \qquad \mathbf{y} = y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}.$$

The dot product of $\mathbf{x}$ and $\mathbf{y}$ can be defined in terms of the components of these vectors as

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3. \tag{4.11.1}$$

An equivalent geometric definition of the dot product is

$$\mathbf{x} \cdot \mathbf{y} = ||\mathbf{x}||\ ||\mathbf{y}|| \cos\theta, \tag{4.11.2}$$

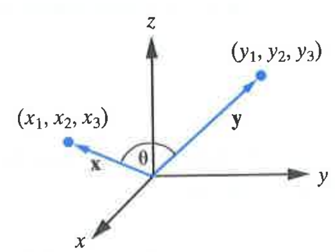

Figure 4.11.1: Defining the dot product in $\mathbb{R}^3$.

where $||\mathbf{x}||$, $||\mathbf{y}||$ denote the lengths of $\mathbf{x}$ and $\mathbf{y}$ respectively, and $0 \le \theta \le \pi$ is the angle between them. (See Figure 4.11.1.)

Taking $\mathbf{y} = \mathbf{x}$ in Equations (4.11.1) and (4.11.2) yields

$$||\mathbf{x}||^2 = \mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 + x_3^2,$$

so that the length of a geometric vector is given in terms of the dot product by

$$||\mathbf{x}|| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Furthermore, from Equation (4.11.2), the angle between any two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ is

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||\ ||\mathbf{y}||}, \tag{4.11.3}$$

which implies that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal (perpendicular) if and only if

$$\mathbf{x} \cdot \mathbf{y} = 0.$$

In a general vector space, we do not have a geometrical picture to guide us in defining the dot product, hence our definitions must be purely algebraic. We begin by considering the vector space $\mathbb{R}^n$, since there is a natural way to extend Equation (4.11.1) in this case. Before proceeding, we note that from now on we will use the standard terms *inner product* and *norm* in place of dot product and length, respectively.

DEFINITION 4.11.1

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be vectors in $\mathbb{R}^n$. We define the **standard inner product in** $\mathbb{R}^n$, denoted $\langle \mathbf{x}, \mathbf{y} \rangle$, by

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

The **norm** of $\mathbf{x}$ is

$$||\mathbf{x}|| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Example 4.11.2 If $\mathbf{x} = (1, -1, 0, 2, 4)$ and $\mathbf{y} = (2, 1, 1, 3, 0)$ in $\mathbb{R}^5$, then

$$\langle \mathbf{x}, \mathbf{y} \rangle = (1)(2) + (-1)(1) + (0)(1) + (2)(3) + (4)(0) = 7,$$
$$||\mathbf{x}|| = \sqrt{1^2 + (-1)^2 + 0^2 + 2^2 + 4^2} = \sqrt{22},$$
$$||\mathbf{y}|| = \sqrt{2^2 + 1^2 + 1^2 + 3^2 + 0^2} = \sqrt{15}.$$

□

Basic Properties of the Standard Inner Product in $\mathbb{R}^n$

In the case of $\mathbb{R}^n$, the definition of the standard inner product was a natural extension of the familiar dot product in $\mathbb{R}^3$. To generalize this definition further to an arbitrary vector space, we isolate the most important properties of the standard inner product in $\mathbb{R}^n$ and use them as the defining criteria for a general notion of an inner product. Let us examine the inner product in $\mathbb{R}^n$ more closely. We view it as a mapping that associates with any two vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$ the real number

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

This mapping has the following properties:

For all $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ in $\mathbb{R}^n$ and all real numbers k,

1. $\langle \mathbf{x}, \mathbf{x} \rangle \ge 0$. Furthermore, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
2. $\langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$.
3. $\langle k\mathbf{x}, \mathbf{y} \rangle = k\langle \mathbf{x}, \mathbf{y} \rangle$.
4. $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.

These properties are easily established using Definition 4.11.1. For example, to prove property 1, we proceed as follows. From Definition 4.11.1,

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Since this is a sum of squares of real numbers, it is necessarily nonnegative. Further, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $x_1 = x_2 = \cdots = x_n = 0$—that is, if and only if $\mathbf{x} = \mathbf{0}$. Similarly, for property 2, we have

$$\langle \mathbf{y}, \mathbf{x} \rangle = y_1x_1 + y_2x_2 + \cdots + y_nx_n = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \langle \mathbf{x}, \mathbf{y} \rangle.$$

We leave the verification of properties 3 and 4 for the reader.

Definition of a Real Inner Product Space

We now use properties 1–4 as the basic defining properties of an inner product in a real vector space.

DEFINITION 4.11.3

Let V be a real vector space. A mapping that associates with each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a real number, denoted $\langle \mathbf{u}, \mathbf{v} \rangle$, is called an **inner product** in V, provided it satisfies the following properties. For all $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V, and all real numbers k,

1. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$. Furthermore, $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.
2. $\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$.
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$.
4. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$.

The **norm** of $\mathbf{u}$ is defined in terms of an inner product by

$$||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

A real vector space together with an inner product defined in it is called a **real inner product space**.

Remarks

1. Observe that $||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$ takes a well-defined nonnegative real value, since property 1 of an inner product guarantees that the norm evaluates the square root of a nonnegative real number.
2. It follows from the discussion above that $\mathbb{R}^n$ together with the inner product defined in Definition 4.11.1 is an example of a real inner product space.

One of the fundamental inner products arises in the vector space $C^0[a, b]$ of all real-valued functions that are *continuous* on the interval $[a, b]$. In this vector space, we define the mapping $\langle f, g \rangle$ by

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx, \tag{4.11.4}$$

for all f and g in $C^0[a, b]$. We establish that this mapping defines an inner product in $C^0[a, b]$ by verifying properties 1–4 of Definition 4.11.3. If f is in $C^0[a, b]$, then

$$\langle f, f \rangle = \int_a^b [f(x)]^2\,dx.$$

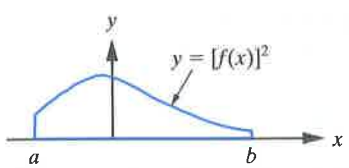

Figure 4.11.2: $\langle f, f \rangle$ gives the area between the graph of $y = [f(x)]^2$ and the x-axis, lying over the interval $[a, b]$.

Since the integrand, $[f(x)]^2$, is a nonnegative continuous function, it follows that $\langle f, f \rangle$ measures the area between the graph $y = [f(x)]^2$ and the x-axis on the interval $[a, b]$. (See Figure 4.11.2.)

Consequently, $\langle f, f \rangle \geq 0$. Furthermore, $\langle f, f \rangle = 0$ if and only if there is zero area between the graph $y = [f(x)]^2$ and the x-axis—that is, if and only if

$$[f(x)]^2 = 0 \qquad \text{for all } x \text{ in } [a, b].$$

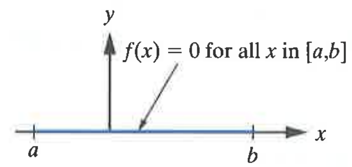

Figure 4.11.3: $\langle f, f \rangle = 0$ if and only if f is the zero function.

Hence, $\langle f, f \rangle = 0$ if and only if $f(x) = 0$, for all x in $[a, b]$, so f must be the zero function. (See Figure 4.11.3.) Consequently, property 1 of Definition 4.11.3 is satisfied. Now let f, g, and h be in $C^0[a, b]$, and let k be an arbitrary real number. Then

$$\langle g, f \rangle = \int_a^b g(x)f(x)\,dx = \int_a^b f(x)g(x)\,dx = \langle f, g \rangle.$$

Hence, property 2 of Definition 4.11.3 is satisfied.

For property 3, we have

$$\langle kf, g \rangle = \int_a^b (kf)(x)g(x)\,dx = \int_a^b kf(x)g(x)\,dx = k\int_a^b f(x)g(x)\,dx = k\langle f, g \rangle,$$

as needed. Finally,

$$\begin{aligned}\langle f + g, h \rangle &= \int_a^b (f + g)(x)h(x)\,dx = \int_a^b [f(x) + g(x)]h(x)\,dx \\ &= \int_a^b f(x)h(x)\,dx + \int_a^b g(x)h(x)\,dx = \langle f, h \rangle + \langle g, h \rangle,\end{aligned}$$

so that property (4) of Definition 4.11.3 is satisfied. We can now conclude that Equation (4.11.4) does define an inner product in the vector space $C^0[a, b]$.

Example 4.11.4 Use Equation (4.11.4) to determine the inner product of the following functions in $C^0[0, 1]$:

$$f(x) = 8x, \qquad g(x) = x^2 - 1.$$

Also find $||f||$ and $||g||$.

Solution: From Equation (4.11.4),

$$\langle f, g \rangle = \int_0^1 8x(x^2 - 1)\,dx = \Big[2x^4 - 4x^2\Big]_0^1 = -2.$$

Moreover, we have

$$||f|| = \sqrt{\int_0^1 64x^2\,dx} = \frac{8}{\sqrt{3}}$$

and

$$||g|| = \sqrt{\int_0^1 (x^2 - 1)^2\,dx} = \sqrt{\int_0^1 (x^4 - 2x^2 + 1)\,dx} = \sqrt{\frac{8}{15}}.$$ □

We have already seen that the norm concept generalizes the length of a geometric vector. Our next goal is to show how an inner product enables us to define the angle between two vectors in an abstract vector space. The key result is the *Cauchy-Schwarz inequality* established in the next theorem.

Theorem 4.11.5 **(Cauchy-Schwarz Inequality)**

Let $\mathbf{u}$ and $\mathbf{v}$ be arbitrary vectors in a real inner product space V. Then

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}||\ ||\mathbf{v}||. \tag{4.11.5}$$

Proof Let k be an arbitrary real number. For the vector $\mathbf{u} + k\mathbf{v}$, we have

$$0 \le ||\mathbf{u} + k\mathbf{v}||^2 = \langle \mathbf{u} + k\mathbf{v}, \mathbf{u} + k\mathbf{v} \rangle. \tag{4.11.6}$$

But, using the properties of a real inner product,

$$\begin{aligned}
\langle \mathbf{u} + k\mathbf{v}, \mathbf{u} + k\mathbf{v} \rangle &= \langle \mathbf{u}, \mathbf{u} + k\mathbf{v} \rangle + \langle k\mathbf{v}, \mathbf{u} + k\mathbf{v} \rangle \\
&= \langle \mathbf{u} + k\mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{u} + k\mathbf{v}, k\mathbf{v} \rangle \\
&= \langle \mathbf{u}, \mathbf{u} \rangle + \langle k\mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{u}, k\mathbf{v} \rangle + \langle k\mathbf{v}, k\mathbf{v} \rangle \\
&= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle k\mathbf{v}, \mathbf{u} \rangle + k\langle \mathbf{v}, k\mathbf{v} \rangle \\
&= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle k\mathbf{v}, \mathbf{u} \rangle + k\langle k\mathbf{v}, \mathbf{v} \rangle \\
&= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle k\mathbf{v}, \mathbf{u} \rangle + k^2\langle \mathbf{v}, \mathbf{v} \rangle \\
&= ||\mathbf{u}||^2 + 2k\langle \mathbf{v}, \mathbf{u} \rangle + k^2||\mathbf{v}||^2.
\end{aligned}$$

Consequently, (4.11.6) implies that

$$||\mathbf{v}||^2 k^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle k + ||\mathbf{u}||^2 \ge 0. \tag{4.11.7}$$

The left-hand side of this inequality defines the quadratic expression

$$P(k) = ||\mathbf{v}||^2 k^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle k + ||\mathbf{u}||^2.$$

The discriminant of this quadratic is

$$\Delta = 4(\langle \mathbf{u}, \mathbf{v} \rangle)^2 - 4||\mathbf{u}||^2||\mathbf{v}||^2.$$

If $\Delta > 0$, then $P(k)$ has two real and distinct roots. This would imply that the graph of P crosses the k-axis and, therefore, P would assume negative values, contrary to (4.11.7). Consequently, we must have $\Delta \le 0$. That is,

$$4(\langle \mathbf{u}, \mathbf{v} \rangle)^2 - 4||\mathbf{u}||^2||\mathbf{v}||^2 \le 0,$$

or equivalently,

$$(\langle \mathbf{u}, \mathbf{v} \rangle)^2 \le ||\mathbf{u}||^2||\mathbf{v}||^2.$$

Hence,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \le ||\mathbf{u}|| \; ||\mathbf{v}||.$$ ■

If $\mathbf{u}$ and $\mathbf{v}$ are arbitrary vectors in a real inner product space V, then $\langle \mathbf{u}, \mathbf{v} \rangle$ is a real number, and so (4.11.5) can be written in the equivalent form

$$-||\mathbf{u}|| \; ||\mathbf{v}|| \le \langle \mathbf{u}, \mathbf{v} \rangle \le ||\mathbf{u}|| \; ||\mathbf{v}||.$$

Consequently, provided that $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, we have

$$-1 \le \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{u}|| \; ||\mathbf{v}||} \le 1.$$

Thus, each pair of nonzero vectors in a *real* inner product space V determines a unique angle θ by

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{u}|| \; ||\mathbf{v}||}, \qquad 0 \le \theta \le \pi. \tag{4.11.8}$$

We call θ the angle between $\mathbf{u}$ and $\mathbf{v}$. In the case when $\mathbf{u}$ and $\mathbf{v}$ are geometric vectors, the formula (4.11.8) coincides with Equation (4.11.3).

Example 4.11.6 Determine the angle between the vectors $\mathbf{u} = (1, -1, 2, 3)$ and $\mathbf{v} = (-2, 1, 2, -2)$ in $\mathbb{R}^4$.

Solution: Using the standard inner product in $\mathbb{R}^4$ yields

$$\langle \mathbf{u}, \mathbf{v} \rangle = -5, \qquad ||\mathbf{u}|| = \sqrt{15}, \qquad ||\mathbf{v}|| = \sqrt{13},$$

so that the angle between $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\cos\theta = -\frac{5}{\sqrt{15}\sqrt{13}} = -\frac{\sqrt{195}}{39}, \qquad 0 \le \theta \le \pi.$$

Hence,

$$\theta = \arccos\left(-\frac{\sqrt{195}}{39}\right) \approx 1.937 \text{ radians} \approx 110^\circ \, 58'.$$ □

Example 4.11.7 Use the inner product (4.11.4) to determine the angle between the functions $f_1(x) = \sin 2x$ and $f_2(x) = \cos 2x$ on the interval $[-\pi, \pi]$.

Solution: Using the inner product (4.11.4), we have

$$\langle f_1, f_2 \rangle = \int_{-\pi}^{\pi} \sin 2x \, \cos 2x \, dx = \frac{1}{2}\int_{-\pi}^{\pi} \sin 4x \, dx = \frac{1}{8}(-\cos 4x)\Big|_{-\pi}^{\pi} = 0.$$

Consequently, the angle between the two functions satisfies

$$\cos\theta = 0, \qquad 0 \le \theta \le \pi,$$

which implies that $\theta = \pi/2$. We say that the functions are *orthogonal* on the interval $[-\pi, \pi]$, relative to the inner product (4.11.4). In the next section we will have much more to say about orthogonality of vectors. □

Complex Inner Products[9]

The preceding discussion has been concerned with real vector spaces. In order to generalize the definition of an inner product to a complex vector space, we first consider the case of $\mathbb{C}^n$. By analogy with Definition 4.11.1, one might think that the natural inner product in $\mathbb{C}^n$ would be obtained by summing the products of corresponding components of vectors in $\mathbb{C}^n$ in exactly the same manner as in the standard inner product for $\mathbb{R}^n$. However, one reason for introducing an inner product is to obtain a concept of "length" of a vector. In order for a quantity to be considered a reasonable measure of length, we would want it to be a nonnegative real number that vanishes if and only if the vector itself is the zero vector (property 1 of a real inner product). But, if we apply the inner product in $\mathbb{R}^n$ given in Definition 4.11.1 to vectors in $\mathbb{C}^n$, then, since the components of vectors in $\mathbb{C}^n$ are complex numbers, it follows that the resulting norm of a vector in

[9] In the remainder of the text, the only complex inner product that we will require is the standard inner product in $\mathbb{C}^n$, and this is needed only in Section 5.10.

$\mathbb{C}^n$ would be a complex number also. Furthermore, applying the $\mathbb{R}^2$ inner product to, for example, the vector $\mathbf{u} = (1-i, 1+i)$, we obtain

$$||\mathbf{u}||^2 = (1-i)^2 + (1+i)^2 = 0,$$

which means that a nonzero vector would have zero "length." To rectify this situation, we must define an inner product in $\mathbb{C}^n$ more carefully. We take advantage of complex conjugation to do this, as the definition shows.

DEFINITION 4.11.8

If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in $\mathbb{C}^n$, we define the **standard inner product in** $\mathbb{C}^n$ by[10]

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1\overline{v}_1 + u_2\overline{v}_2 + \cdots + u_n\overline{v}_n.$$

The **norm** of **u** is defined to be the *real number*

$$||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{|u_1|^2 + |u_2|^2 + \cdots + |u_n|^2}.$$

The preceding inner product is a mapping that associates with the two vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{C}^n$ the *scalar*

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1\overline{v}_1 + u_2\overline{v}_2 + \cdots + u_n\overline{v}_n.$$

In general, $\langle \mathbf{u}, \mathbf{v} \rangle$ will be nonreal (i.e., it will have a nonzero imaginary part). The key point to notice is that the norm of **u** is always a *real* number, even though the separate components of **u** are complex numbers.

Example 4.11.9 If $\mathbf{u} = (1+2i, 2-3i)$ and $\mathbf{v} = (2-i, 3+4i)$, find $\langle \mathbf{u}, \mathbf{v} \rangle$ and $||\mathbf{u}||$.

Solution: Using Definition 4.11.8,

$$\langle \mathbf{u}, \mathbf{v} \rangle = (1+2i)(2+i) + (2-3i)(3-4i) = 5i - 6 - 17i = -6 - 12i,$$

$$||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{(1+2i)(1-2i) + (2-3i)(2+3i)} = \sqrt{5+13} = 3\sqrt{2}. \quad \square$$

The standard inner product in $\mathbb{C}^n$ satisfies properties (1), (3), and (4), but not property (2). We now derive the appropriate generalization of property (2) when using the standard inner product in $\mathbb{C}^n$. Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ be vectors in $\mathbb{C}^n$. Then, from Definition 4.11.8,

$$\langle \mathbf{v}, \mathbf{u} \rangle = v_1\overline{u}_1 + v_2\overline{u}_2 + \cdots + v_n\overline{u}_n = \overline{u_1\overline{v}_1 + u_2\overline{v}_2 + \cdots + u_n\overline{v}_n} = \overline{\langle \mathbf{u}, \mathbf{v} \rangle}.$$

Thus,

$$\langle \mathbf{v}, \mathbf{u} \rangle = \overline{\langle \mathbf{u}, \mathbf{v} \rangle}.$$

We now use the properties satisfied by the standard inner product in $\mathbb{C}^n$ to define an inner product in an arbitrary (that is, real or complex) vector space.

[10] Recall that if $z = a + ib$, then $\overline{z} = a - ib$ and $|z|^2 = z\overline{z} = (a+ib)(a-ib) = a^2 + b^2$.

DEFINITION 4.11.10

Let V be a (real or complex) vector space. A mapping that associates with each pair of vectors **u**, **v** in V a scalar, denoted $\langle \mathbf{u}, \mathbf{v} \rangle$, is called an **inner product** in V, provided it satisfies the following properties. For all **u**, **v** and **w** in V and all (real or complex) scalars k,

1. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$. Furthermore, $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.
2. $\langle \mathbf{v}, \mathbf{u} \rangle = \overline{\langle \mathbf{u}, \mathbf{v} \rangle}$.
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$.
4. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$.

The **norm** of **u** is defined in terms of the inner product by

$$||\mathbf{u}|| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

Remark Notice that the properties in the preceding definition reduce to those in Definition 4.11.3 in the case that V is a *real* vector space, since in such a case the complex conjugates are unnecessary. Thus, this definition is a consistent extension of Definition 4.11.3.

Example 4.11.11 Use properties 2 and 3 of Definition 4.11.10 to prove that in an inner product space

$$\langle \mathbf{u}, k\mathbf{v} \rangle = \overline{k}\langle \mathbf{u}, \mathbf{v} \rangle$$

for all vectors **u**, **v** and all scalars k.

Solution: From properties 2 and 3, we have

$$\langle \mathbf{u}, k\mathbf{v} \rangle = \overline{\langle k\mathbf{v}, \mathbf{u} \rangle} = \overline{k\langle \mathbf{v}, \mathbf{u} \rangle} = \overline{k}\;\overline{\langle \mathbf{v}, \mathbf{u} \rangle} = \overline{k}\;\langle \mathbf{u}, \mathbf{v} \rangle.$$

Notice that in the particular case of a real vector space, the foregoing result reduces to

$$\langle \mathbf{u}, k\mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle,$$

since in such a case the scalars are real numbers. $\square$

Exercises for 4.11

Key Terms

Inner product, Axioms of an inner product, Real (complex) inner product space, Norm, Angle, Cauchy-Schwarz inequality.

Skills

- Know the four inner product space axioms.
- Be able to check whether or not a proposed inner product on a vector space V satisfies the inner product space axioms.
- Be able to compute the inner product of two vectors in an inner product space.
- Be able to find the norm of a vector in an inner product space.
- Be able to find the angle between two vectors in an inner product space.

True-False Review

For Questions 1–7, decide if the given statement is **true** or **false**, and give a brief justification for your answer. If true, you can quote a relevant definition or theorem from the text. If false, provide an example, illustration, or brief explanation of why the statement is false.

1. If $\mathbf{v}$ and $\mathbf{w}$ are linearly independent vectors in an inner product space V, then $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

2. In any inner product space V, we have
$$\langle k\mathbf{v}, k\mathbf{w} \rangle = k\langle \mathbf{v}, \mathbf{w} \rangle.$$

3. If $\langle \mathbf{v}_1, \mathbf{w} \rangle = \langle \mathbf{v}_2, \mathbf{w} \rangle = 0$ in an inner product space V, then
$$\langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2, \mathbf{w} \rangle = 0.$$

4. In any inner product space V, $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle < 0$ if and only if $||\mathbf{x}|| < ||\mathbf{y}||$.

5. In any vector space V, there is at most one valid inner product $\langle \, , \rangle$ that can be defined on V.

6. The angle between the vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V is the same as the angle between the vectors $-2\mathbf{v}$ and $-2\mathbf{w}$.

7. If $p(x) = a_0+a_1x+a_2x^2$ and $q(x) = b_0+b_1x+b_2x^2$, then we can define an inner product on P_2 via $\langle p, q \rangle = a_0b_0$.

Problems

1. Use the standard inner product in $\mathbb{R}^4$ to determine the angle between the vectors $\mathbf{v} = (1, 3, -1, 4)$ and $\mathbf{w} = (-1, 1, -2, 1)$.

2. If $f(x) = \sin x$ and $g(x) = x$ on $[0, \pi]$, use the function inner product defined in the text to determine the angle between f and g.

3. If $\mathbf{v} = (2+i, 3-2i, 4+i)$ and $\mathbf{w} = (-1+i, 1-3i, 3-i)$, use the standard inner product in $\mathbb{C}^3$ to determine, $\langle \mathbf{v}, \mathbf{w} \rangle$, $||\mathbf{v}||$, and $||\mathbf{w}||$.

4. Let
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$
be vectors in $M_2(\mathbb{R})$. Show that the mapping
$$\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$$
defines an inner product in $M_2(\mathbb{R})$.

5. Referring to A and B in the previous problem, show that the mapping
$$\langle A, B \rangle = a_{11}b_{22} + a_{12}b_{21} + a_{21}b_{12} + a_{22}b_{11}$$
does *not* define a valid inner product on $M_2(\mathbb{R})$.

For Problems 6–7, use the inner product given in Problem 4 to determine $\langle A, B \rangle$, $|||A|||$, and $|||B|||$.

6. $A = \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$.

7. $A = \begin{bmatrix} 3 & 2 \\ -2 & 4 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$.

8. Let $p_1(x) = a + bx$ and $p_2(x) = c + dx$ be vectors in P_1. Determine a mapping $\langle p_1, p_2 \rangle$ that defines an inner product on P_1.

9. Verify that for all $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ in $\mathbb{R}^2$,
$$\langle \mathbf{v}, \mathbf{w} \rangle = 2v_1w_1 + v_1w_2 + v_2w_1 + 2v_2w_2$$
defines an inner product on $\mathbb{R}^2$.

For Problems 10–12, determine the inner product of the given vectors using **(a)** the inner product given in Problem 9, **(b)** the standard inner product in $\mathbb{R}^2$.

10. $\mathbf{v} = (1, 0), \mathbf{w} = (-1, 2)$.

11. $\mathbf{v} = (2, -1), \mathbf{w} = (3, 6)$.

12. $\mathbf{v} = (1, -2), \mathbf{w} = (2, 1)$.

13. Consider the vector space $\mathbb{R}^2$. Define the mapping $\langle \, , \rangle$ by
$$\langle \mathbf{v}, \mathbf{w} \rangle = v_1w_1 - v_2w_2,$$
for all vectors $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$. Verify that all of the properties in Definition 4.11.3 except (1) are satisfied by $\langle \, , \rangle$.

The mapping in Problem 13 is called a **pseudo-inner product** in $\mathbb{R}^2$ and, when generalized to $\mathbb{R}^4$, is of fundamental importance in Einstein's special relativity theory.

14. Using Equation in Problem 13, determine all nonzero vectors satisfying $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. Such vectors are called **null** vectors.

15. Using Equation in Problem 13, determine all vectors satisfying $\langle \mathbf{v}, \mathbf{v} \rangle < 0$. Such vectors are called **timelike** vectors.

16. Using Equation (4.11.11), determine all vectors satisfying $\langle \mathbf{v}, \mathbf{v} \rangle > 0$. Such vectors are called **spacelike** vectors.

17. Make a sketch of $\mathbb{R}^2$ and indicate the position of the null, timelike, and spacelike vectors.

18. Consider the vector space $\mathbb{R}^n$, and let $\mathbf{v} = (v_1, v_2, \dots, v_n)$ and $\mathbf{w} = (w_1, w_2, \dots, w_n)$ be vectors in $\mathbb{R}^n$. Show that the mapping $\langle , \rangle$ defined by
$$\langle \mathbf{v}, \mathbf{w} \rangle = k_1v_1w_1 + k_2v_2w_2 + \cdots + k_nv_nw_n$$
is a valid inner product on $\mathbb{R}^n$ if and only if the constants $k_1, k_2, \dots, k_n$ are all positive.

19. Prove from the inner product axioms that, in any inner product space V, $\langle \mathbf{v}, \mathbf{0} \rangle = 0$ for all $\mathbf{v}$ in V.

20. Let V be a real inner product space.

(a) Prove that for all $\mathbf{v}, \mathbf{w} \in V$,
$$||\mathbf{v} + \mathbf{w}||^2 = ||\mathbf{v}||^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + ||\mathbf{w}||^2.$$
[**Hint**: $||\mathbf{v} + \mathbf{w}||^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle$.]

(b) Two vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V are called *orthogonal* if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Use (a) to prove the general **Pythagorean theorem:** If $\mathbf{v}$ and $\mathbf{w}$ are orthogonal in an inner product space V, then
$$||\mathbf{v} + \mathbf{w}||^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2.$$

(c) Prove that for all $\mathbf{v}, \mathbf{w}$ in V,

(i) $||\mathbf{v} + \mathbf{w}||^2 - ||\mathbf{v} - \mathbf{w}||^2 = 4\langle \mathbf{v}, \mathbf{w} \rangle$.

(ii) $||\mathbf{v} + \mathbf{w}||^2 + ||\mathbf{v} - \mathbf{w}||^2 = 2(||\mathbf{v}||^2 + ||\mathbf{w}||^2)$.

21. Let V be a complex inner product space. Prove that for all $\mathbf{v}, \mathbf{w}$ in V,
$$||\mathbf{v} + \mathbf{w}||^2 = ||\mathbf{v}||^2 + 2\mathrm{Re}(\langle \mathbf{v}, \mathbf{w} \rangle) + ||\mathbf{v}||^2,$$
where Re denotes the real part of a complex number.

4.12 Orthogonal Sets of Vectors and the Gram-Schmidt Process

The discussion in the previous section has shown how an inner product can be used to define the angle between two nonzero vectors. In particular, if the inner product of two nonzero vectors is zero, then the angle between those two vectors is $\pi/2$ radians, and therefore it is natural to call such vectors orthogonal (perpendicular). The following definition extends the idea of orthogonality into an arbitrary inner product space.

DEFINITION 4.12.1

Let V be an inner product space.

1. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in V are said to be **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
2. A set of nonzero vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in V is called an **orthogonal set** of vectors if
$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0, \qquad \text{whenever } i \neq j.$$
(That is, every vector is orthogonal to every other vector in the set.)
3. A vector $\mathbf{v}$ in V is called a **unit vector** if $||\mathbf{v}|| = 1$.
4. An *orthogonal* set of *unit* vectors is called an **orthonormal set** of vectors. Thus, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in V is an orthonormal set if and only if

   (a) $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ whenever $i \neq j$.

   (b) $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ for all $i = 1, 2, \dots, k$.

Remarks

1. The conditions in (4a) and (4b) can be written compactly in terms of the Kronecker delta symbol as
$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}, \qquad i, j = 1, 2, \dots, k.$$
2. Note that the inner products occurring in Definition 4.12.1 will depend upon which inner product space we are working in.
3. If $\mathbf{v}$ is any nonzero vector, then $\frac{1}{||\mathbf{v}||}\mathbf{v}$ is a unit vector, since the properties of an inner product imply that
$$\left\langle \frac{1}{||\mathbf{v}||}\mathbf{v}, \frac{1}{||\mathbf{v}||}\mathbf{v} \right\rangle = \frac{1}{||\mathbf{v}||^2}\langle \mathbf{v}, \mathbf{v} \rangle = \frac{1}{||\mathbf{v}||^2}||\mathbf{v}||^2 = 1.$$

Using Remark 3 above, we can take an orthogonal set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ and create a new set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$, where $\mathbf{u}_i = \frac{1}{||\mathbf{v}_i||}\mathbf{v}_i$ is a unit vector for each i. Using the properties of an inner product, it is easy to see that the new set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is an orthonormal set (see Problem 31). The process of replacing the $\mathbf{v}_i$ by the $\mathbf{u}_i$ is called **normalization**.

Example 4.12.2 Verify that $\{(-2, 1, 3, 0), (0, -3, 1, -6), (-2, -4, 0, 2)\}$ is an orthogonal set of vectors in $\mathbb{R}^4$, and use it to construct an orthonormal set of vectors in $\mathbb{R}^4$.

Solution: Let $\mathbf{v}_1 = (-2, 1, 3, 0)$, $\mathbf{v}_2 = (0, -3, 1, -6)$, and $\mathbf{v}_3 = (-2, -4, 0, 2)$. Then
$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0, \qquad \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = 0, \qquad \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0,$$
so that the given set of vectors is an orthogonal set. Dividing each vector in the set by its norm yields the following orthonormal set:
$$\left\{ \frac{1}{\sqrt{14}}\mathbf{v}_1, \frac{1}{\sqrt{46}}\mathbf{v}_2, \frac{1}{2\sqrt{6}}\mathbf{v}_3 \right\}. \qquad \square$$

Example 4.12.3 Verify that the functions $f_1(x) = 1$, $f_2(x) = \sin x$, and $f_3(x) = \cos x$ are orthogonal in $C^0[-\pi, \pi]$, and use them to construct an orthonormal set of functions in $C^0[-\pi, \pi]$.

Solution: In this case, we have
$$\langle f_1, f_2 \rangle = \int_{-\pi}^{\pi} \sin x \, dx = 0, \qquad \langle f_1, f_3 \rangle = \int_{-\pi}^{\pi} \cos x \, dx = 0,$$
$$\langle f_2, f_3 \rangle = \int_{-\pi}^{\pi} \sin x \cos x \, dx = \left[\frac{1}{2}\sin^2 x\right]_{-\pi}^{\pi} = 0,$$
so that the functions are indeed orthogonal on $[-\pi, \pi]$. Taking the norm of each function, we obtain
$$||f_1|| = \sqrt{\int_{-\pi}^{\pi} 1 \, dx} = \sqrt{2\pi},$$
$$||f_2|| = \sqrt{\int_{-\pi}^{\pi} \sin^2 x \, dx} = \sqrt{\int_{-\pi}^{\pi} \frac{1}{2}(1 - \cos 2x) \, dx} = \sqrt{\pi},$$
$$||f_3|| = \sqrt{\int_{-\pi}^{\pi} \cos^2 x \, dx} = \sqrt{\int_{-\pi}^{\pi} \frac{1}{2}(1 + \cos 2x) \, dx} = \sqrt{\pi}.$$

Thus an orthonormal set of functions on $[-\pi, \pi]$ is
$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\sin x, \frac{1}{\sqrt{\pi}}\cos x \right\}. \qquad \square$$

Orthogonal and Orthonormal Bases

In the analysis of geometric vectors in elementary calculus courses, it is usual to use the standard basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Notice that this set of vectors is in fact an orthonormal set. The introduction of an inner product in a vector space opens up the possibility of using similar bases in a general finite-dimensional vector space. The next definition introduces the appropriate terminology.

DEFINITION 4.12.4

A basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ for a (finite-dimensional) inner product space is called an **orthogonal** basis if
$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \qquad \text{whenever } i \neq j,$$
and it is called an **orthonormal** basis if
$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}, \qquad i, j = 1, 2, \dots, n.$$

There are two natural questions at this point: (1) How can we obtain an orthogonal or orthonormal basis for an inner product space V? (2) Why is it beneficial to work with an orthogonal or orthonormal basis of vectors? We address the second question first.

In light of our work in previous sections of this chapter, the importance of our next theorem should be self-evident.

Theorem 4.12.5 If $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is an *orthogonal* set of *nonzero* vectors in an inner product space V, then $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is linearly independent.

Proof Assume that
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}. \qquad (4.12.1)$$
We will show that $c_1 = c_2 = \cdots = c_k = 0$. Taking the inner product of each side of (4.12.1) with $\mathbf{v}_i$, we find that
$$\langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k, \mathbf{v}_i \rangle = \langle \mathbf{0}, \mathbf{v}_i \rangle = 0.$$
Using the inner product properties on the left side, we have
$$c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_i \rangle + \cdots + c_k\langle \mathbf{v}_k, \mathbf{v}_i \rangle = 0.$$
Finally, using the fact that for all $j \neq i$, we have $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$, we conclude that
$$c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 0.$$
Since $\mathbf{v}_i \neq \mathbf{0}$, it follows that $c_i = 0$, and this holds for each i with $1 \le i \le k$. ■

Example 4.12.6 Let $V = M_2(\mathbb{R})$, let W be the subspace of all 2×2 *symmetric* matrices, and let
$$S = \left\{ \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & -3 \end{bmatrix} \right\}.$$

Define an inner product on V via[11]

$$\left\langle \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}.$$

Show that S is an orthogonal basis for W.

Solution: According to Example 4.6.18, we already know that $\dim[W] = 3$. Using the given inner product, it can be directly shown that S is an orthogonal set, and hence, Theorem 4.12.5 implies that S is linearly independent. Therefore, by Theorem 4.6.10, S is a basis for W. □

Let V be a (finite-dimensional) inner product space, and suppose that we have an *orthogonal* basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ for V. As we saw in Section 4.7, any vector $\mathbf{v}$ in V can be written *uniquely* in the form

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n, \tag{4.12.2}$$

where the unique n-tuple $(c_1, c_2, \dots, c_n)$ consists of the components of $\mathbf{v}$ relative to the given basis. It is easier to determine the components c_i in the case of an orthogonal basis than it is for other bases, because we can simply form the inner product of both sides of (4.12.2) with $\mathbf{v}_i$ as follows:

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v}_i \rangle &= \langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n, \mathbf{v}_i \rangle \\ &= c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_i \rangle + \cdots + c_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle \\ &= c_i ||\mathbf{v}_i||^2, \end{aligned}$$

where the last step follows from the orthogonality properties of the basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Therefore, we have proved the following theorem.

Theorem 4.12.7 Let V be a (finite-dimensional) inner product space with orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Then any vector $\mathbf{v} \in V$ may be expressed in terms of the basis as

$$\mathbf{v} = \left(\frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2}\right)\mathbf{v}_1 + \left(\frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{||\mathbf{v}_2||^2}\right)\mathbf{v}_2 + \cdots + \left(\frac{\langle \mathbf{v}, \mathbf{v}_n \rangle}{||\mathbf{v}_n||^2}\right)\mathbf{v}_n.$$

Theorem 4.12.7 gives a simple formula for writing an arbitrary vector in an inner product space V as a linear combination of vectors in an orthogonal basis for V. Let us illustrate with an example.

Example 4.12.8 Let V, W, and S be as in Example 4.12.6. Find the components of the vector

$$\mathbf{v} = \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}$$

relative to S.

Solution: From the formula given in Theorem 4.12.7, we have

$$\mathbf{v} = \frac{2}{6}\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix} + \frac{2}{7}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} - \frac{10}{21}\begin{bmatrix} 2 & 2 \\ 2 & -3 \end{bmatrix},$$

[11] This defines a valid inner product on V by Problem 4 in Section 4.11.

so the components of $\mathbf{v}$ relative to S are

$$\left(\frac{1}{3}, \frac{2}{7}, -\frac{10}{21}\right).$$ □

If the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ for V is in fact orthonormal, then since $||\mathbf{v}_i|| = 1$ for each i, we immediately deduce the following corollary of Theorem 4.12.7.

Corollary 4.12.9 Let V be a (finite-dimensional) inner product space with an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Then any vector $\mathbf{v} \in V$ may be expressed in terms of the basis as

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{v}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n.$$

Remark Corollary 4.12.9 tells us that the components of a given vector $\mathbf{v}$ relative to the orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ are precisely the numbers $\langle \mathbf{v}, \mathbf{v}_i \rangle$, for $1 \le i \le n$. Thus, by working with an orthonormal basis for a vector space, we have a simple method for getting the components of any vector in the vector space.

Example 4.12.10 We can write an arbitrary vector in $\mathbb{R}^n$, $\mathbf{v} = (a_1, a_2, \dots, a_n)$, in terms of the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ by noting that $\langle \mathbf{v}, \mathbf{e}_i \rangle = a_i$. Thus, $\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n$. □

Example 4.12.11 We can equip the vector space P_1 of all polynomials of degree ≤ 1 with inner product

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

thus making P_1 into an inner product space. Verify that the vectors $p_0 = 1/\sqrt{2}$ and $p_1 = \sqrt{1.5}x$ form an orthonormal basis for P_1 and use Corollary 4.12.9 to write the vector $q = 1 + x$ as a linear combination of p_0 and p_1.

Solution: We have

$$\langle p_0, p_1 \rangle = \int_{-1}^{1} \frac{1}{\sqrt{2}} \cdot \sqrt{1.5}x\,dx = 0,$$

$$||p_0|| = \sqrt{\langle p_0, p_0 \rangle} = \sqrt{\int_{-1}^{1} p_0^2\,dx} = \sqrt{\int_{-1}^{1} \frac{1}{2}\,dx} = \sqrt{1} = 1,$$

$$||p_1|| = \sqrt{\langle p_1, p_1 \rangle} = \sqrt{\int_{-1}^{1} p_1^2\,dx} = \sqrt{\int_{-1}^{1} \frac{3}{2}x^2\,dx} = \sqrt{\frac{1}{2}x^3\Big|_{-1}^{1}} = \sqrt{1} = 1.$$

Thus, $\{p_0, p_1\}$ is an orthonormal (and hence linearly independent) set of vectors in P_1. Since $\dim[P_1] = 2$, Theorem 4.6.10 shows that $\{p_0, p_1\}$ is an (orthonormal) basis for P_1.

Finally, we wish to write $q = 1 + x$ as a linear combination of p_0 and p_1, by using Corollary 4.12.9. We leave it to the reader to verify that $\langle q, p_0 \rangle = \sqrt{2}$ and $\langle q, p_1 \rangle = \sqrt{\frac{2}{3}}$. Thus, we have

$$1 + x = \sqrt{2}\,p_0 + \sqrt{\frac{2}{3}}\,p_1 = \sqrt{2} \cdot \frac{1}{\sqrt{2}} + \sqrt{\frac{2}{3}} \cdot \left(\sqrt{\frac{3}{2}}x\right).$$

So the component vector of $1 + x$ relative to $\{p_0, p_1\}$ is $(\sqrt{2}, \sqrt{\frac{2}{3}})^T$. □

The Gram-Schmidt Process

Next, we return to address the first question we raised earlier: How can we obtain an orthogonal or orthonormal basis for an inner product space V? The idea behind the process is to begin with *any* basis for V, say $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, and to successively replace these vectors with vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ that are orthogonal to one another, and to ensure that, throughout the process, the span of the vectors remains unchanged. This is known as the **Gram-Schmidt process**. To describe it, we shall once more appeal to a look at geometric vectors.

If $\mathbf{v}$ and $\mathbf{w}$ are any two linearly independent (noncollinear) geometric vectors, then the **orthogonal projection** of $\mathbf{w}$ on $\mathbf{v}$ is the vector $\mathbf{P}(\mathbf{w}, \mathbf{v})$ shown in Figure 4.12.1. We see from the figure that an orthogonal basis for the subspace (plane) of 3-space spanned by $\mathbf{v}$ and $\mathbf{w}$ is $\{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \mathbf{v} \quad \text{and} \quad \mathbf{v}_2 = \mathbf{w} - \mathbf{P}(\mathbf{w}, \mathbf{v}).$$

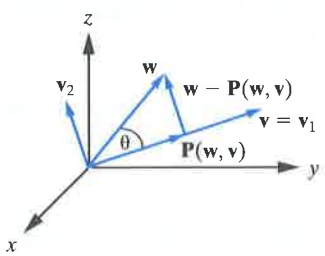

Figure 4.12.1: Obtaining an orthogonal basis for a two-dimensional subspace of $\mathbb{R}^3$.

In order to generalize this result to an arbitrary inner product space, we need to derive an expression for $\mathbf{P}(\mathbf{w}, \mathbf{v})$ in terms of the dot product. We see from Figure 4.12.1 that the norm of $\mathbf{P}(\mathbf{w}, \mathbf{v})$ is

$$||\mathbf{P}(\mathbf{w}, \mathbf{v})|| = ||\mathbf{w}|| \cos\theta,$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$. Thus

$$\mathbf{P}(\mathbf{w}, \mathbf{v}) = ||\mathbf{w}|| \cos\theta \frac{\mathbf{v}}{||\mathbf{v}||},$$

which we can write as

$$\mathbf{P}(\mathbf{w}, \mathbf{v}) = \left(\frac{||\mathbf{w}||\,||\mathbf{v}||}{||\mathbf{v}||^2} \cos\theta\right) \mathbf{v}. \tag{4.12.3}$$

Recalling that the dot product of the vectors $\mathbf{w}$ and $\mathbf{v}$ is defined by

$$\mathbf{w} \cdot \mathbf{v} = ||\mathbf{w}||\,||\mathbf{v}|| \cos\theta,$$

it follows from Equation (4.12.3) that

$$\mathbf{P}(\mathbf{w}, \mathbf{v}) = \frac{(\mathbf{w} \cdot \mathbf{v})}{||\mathbf{v}||^2} \mathbf{v},$$

or equivalently, using the notation for the inner product introduced in the previous section,

$$\mathbf{P}(\mathbf{w}, \mathbf{v}) = \frac{\langle \mathbf{w}, \mathbf{v} \rangle}{||\mathbf{v}||^2} \mathbf{v}.$$

Now let $\mathbf{x}_1$ and $\mathbf{x}_2$ be linearly independent vectors in an arbitrary inner product space V. We show next that the foregoing formula can also be applied in V to obtain an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the subspace of V spanned by $\{\mathbf{x}_1, \mathbf{x}_2\}$. Let

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

and

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{P}(\mathbf{x}_2, \mathbf{v}_1) = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2} \mathbf{v}_1. \tag{4.12.4}$$

Note from (4.12.4) that $\mathbf{v}_2$ can be written as a linear combination of $\{\mathbf{x}_1, \mathbf{x}_2\}$, and hence, $\mathbf{v}_2 \in \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$. Since we also have that $\mathbf{x}_2 \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$, it follows that $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$. Next we claim that $\mathbf{v}_2$ is orthogonal to $\mathbf{v}_1$. We have

$$\begin{aligned}\langle \mathbf{v}_2, \mathbf{v}_1 \rangle &= \left\langle \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2}\mathbf{v}_1, \mathbf{v}_1 \right\rangle = \langle \mathbf{x}_2, \mathbf{v}_1 \rangle - \left\langle \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2}\mathbf{v}_1, \mathbf{v}_1 \right\rangle \\ &= \langle \mathbf{x}_2, \mathbf{v}_1 \rangle - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 0,\end{aligned}$$

which verifies our claim. We have shown that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal set of vectors which spans the same subspace of V as $\mathbf{x}_1$ and $\mathbf{x}_2$.

The calculations just presented can be generalized to prove the following useful result (see Problem 32).

Lemma 4.12.12 Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be an orthogonal set of vectors in an inner product space V. If $\mathbf{x} \in V$, then the vector

$$\mathbf{x} - \mathbf{P}(\mathbf{x}, \mathbf{v}_1) - \mathbf{P}(\mathbf{x}, \mathbf{v}_2) - \dots - \mathbf{P}(\mathbf{x}, \mathbf{v}_k)$$

is orthogonal to $\mathbf{v}_i$ for each i.

Now suppose we are given a linearly independent set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ in an inner product space V. Using Lemma 4.12.12, we can construct an orthogonal basis for the subspace of V spanned by these vectors. We begin with the vector $\mathbf{v}_1 = \mathbf{x}_1$ as above, and we define $\mathbf{v}_i$ by subtracting off appropriate projections of $\mathbf{x}_i$ on $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}$. The resulting procedure is called the **Gram-Schmidt orthogonalization procedure**. The formal statement of the result is as follows.

Theorem 4.12.13 **(Gram-Schmidt Process)**

Let $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ be a linearly independent set of vectors in an inner product space V. Then an *orthogonal basis* for the subspace of V spanned by these vectors is $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$, where

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{x}_1 \\ \mathbf{v}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2}\mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{||\mathbf{v}_1||^2}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{||\mathbf{v}_2||^2}\mathbf{v}_2 \\ &\vdots \\ \mathbf{v}_i &= \mathbf{x}_i - \sum_{k=1}^{i-1} \frac{\langle \mathbf{x}_i, \mathbf{v}_k \rangle}{||\mathbf{v}_k||^2}\mathbf{v}_k \\ &\vdots \\ \mathbf{v}_m &= \mathbf{x}_m - \sum_{k=1}^{m-1} \frac{\langle \mathbf{x}_m, \mathbf{v}_k \rangle}{||\mathbf{v}_k||^2}\mathbf{v}_k.\end{aligned}$$

Proof Lemma 4.12.12 shows that $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is an orthogonal set of vectors. Thus, both $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ and $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ are linearly independent sets, and hence

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\} \quad \text{and} \quad \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$$

are m-dimensional subspaces of V. (Why?) Moreover, from the formulas given in Theorem 4.12.13, we see that each $\mathbf{x}_i \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$, and so $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is a subset of $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$. Thus, by Corollary 4.6.14,

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}.$$

We conclude that $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is a basis for the subspace of V spanned by $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. ■

Example 4.12.14 Obtain an orthogonal basis for the subspace of $\mathbb{R}^4$ spanned by

$$\mathbf{x}_1 = (1, 0, 1, 0), \quad \mathbf{x}_2 = (1, 1, 1, 1), \quad \mathbf{x}_3 = (-1, 2, 0, 1).$$

Solution: Following the Gram-Schmidt process, we set $\mathbf{v}_1 = \mathbf{x}_1 = (1, 0, 1, 0)$. Next, we have

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = (1, 1, 1, 1) - \frac{2}{2}(1, 0, 1, 0) = (0, 1, 0, 1)$$

and

$$\begin{aligned}\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \\ &= (-1, 2, 0, 1) + \frac{1}{2}(1, 0, 1, 0) - \frac{3}{2}(0, 1, 0, 1) \\ &= \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right).\end{aligned}$$

The orthogonal basis so obtained is

$$\left\{(1, 0, 1, 0), (0, 1, 0, 1), \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)\right\}.$$ □

Of course, once an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is obtained for a subspace of V, we can normalize this basis by setting $\mathbf{u}_i = \dfrac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$ to obtain an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$. For instance, an orthonormal basis for the subspace of $\mathbb{R}^4$ in the preceding example is

$$\left\{\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0\right), \left(0, \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)\right\}.$$

Example 4.12.15 Determine an orthogonal basis for the subspace of $C^0[-1, 1]$ spanned by the functions $f_1(x) = x$, $f_2(x) = x^3$, $f_3(x) = x^5$, using the same inner product introduced in the previous section.

Solution: In this case, we let $\{g_1, g_2, g_3\}$ denote the orthogonal basis, and we apply the Gram-Schmidt process. Thus, $g_1(x) = x$, and

$$g_2(x) = f_2(x) - \frac{\langle f_2, g_1 \rangle}{\|g_1\|^2}g_1(x). \quad (4.12.5)$$

We have

$$\langle f_2, g_1 \rangle = \int_{-1}^{1} f_2(x)g_1(x)\,dx = \int_{-1}^{1} x^4\,dx = \tfrac{2}{5} \quad \text{and}$$

$$\|g_1\|^2 = \langle g_1, g_1 \rangle = \int_{-1}^{1} x^2\,dx = \tfrac{2}{3}.$$

Substituting into Equation (4.12.5) yields

$$g_2(x) = x^3 - \tfrac{3}{5}x = \tfrac{1}{5}x(5x^2 - 3).$$

We now compute $g_3(x)$. According to the Gram-Schmidt process,

$$g_3(x) = f_3(x) - \frac{\langle f_3, g_1 \rangle}{\|g_1\|^2}g_1(x) - \frac{\langle f_3, g_2 \rangle}{\|g_2\|^2}g_2(x). \quad (4.12.6)$$

We first evaluate the required inner products:

$$\begin{aligned}\langle f_3, g_1 \rangle &= \int_{-1}^{1} f_3(x)g_1(x)\,dx = \int_{-1}^{1} x^6\,dx = \tfrac{2}{7}, \\ \langle f_3, g_2 \rangle &= \int_{-1}^{1} f_3(x)g_2(x)\,dx = \frac{1}{5}\int_{-1}^{1} x^6(5x^2 - 3)\,dx = \tfrac{1}{5}\left(\tfrac{10}{9} - \tfrac{6}{7}\right) = \tfrac{16}{315}, \\ \|g_2\|^2 &= \int_{-1}^{1} [g_2(x)]^2\,dx = \tfrac{1}{25}\int_{-1}^{1} x^2(5x^2 - 3)^2\,dx \\ &= \tfrac{1}{25}\int_{-1}^{1} (25x^6 - 30x^4 + 9x^2)\,dx = \tfrac{8}{175}.\end{aligned}$$

Substituting into Equation (4.12.6) yields

$$g_3(x) = x^5 - \tfrac{3}{7}x - \tfrac{2}{9}x(5x^2 - 3) = \tfrac{1}{63}(63x^5 - 70x^3 + 15x).$$

Thus, an orthogonal basis for the subspace of $C^0[-1, 1]$ spanned by f_1, f_2, and f_3 is

$$\left\{x, \tfrac{1}{5}x(5x^2 - 3), \tfrac{1}{63}x(63x^4 - 70x^2 + 15)\right\}.$$ □

Exercises for 4.12

Key Terms

Orthogonal vectors, Orthogonal set, Unit vector, Orthonormal vectors, Orthonormal set, Normalization, Orthogonal basis, Orthonormal basis, Gram-Schmidt process, Orthogonal projection.

Skills

- Be able to determine whether a given set of vectors are orthogonal and/or orthonormal.
- Be able to determine whether a given set of vectors forms an orthogonal and/or orthonormal basis for an inner product space.
- Be able to replace an orthogonal set with an orthonormal set via normalization.
- Be able to readily compute the components of a vector $\mathbf{v}$ in an inner product space V relative to an orthogonal (or orthonormal) basis for V.
- Be able to compute the orthogonal projection of one vector $\mathbf{w}$ along another vector $\mathbf{v}$: $\mathbf{P}(\mathbf{w}, \mathbf{v})$.
- Be able to carry out the Gram-Schmidt process to replace a basis for V with an orthogonal (or orthonormal) basis for V.

True-False Review

For Questions 1–7, decide if the given statement is **true** or **false**, and give a brief justification for your answer. If true, you can quote a relevant definition or theorem from the text. If false, provide an example, illustration, or brief explanation of why the statement is false.

1. Every orthonormal basis for an inner product space V is also an orthogonal basis for V.

2. Every linearly independent set of vectors in an inner product space V is orthogonal.

3. With the inner product $\langle f, g\rangle = \int_0^{\pi} f(t)g(t)\,dt$, the functions $f(x) = \cos x$ and $g(x) = \sin x$ are an orthogonal basis for span$\{\cos x, \sin x\}$.

4. The Gram-Schmidt process applied to the vectors $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ yields the same basis as the Gram-Schmidt process applied to the vectors $\{\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1\}$.

5. In expressing the vector $\mathbf{v}$ as a linear combination of the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ for an inner product space V, the coefficient of $\mathbf{v}_i$ is

$$c_i = \frac{\langle \mathbf{v}, \mathbf{v}_i\rangle}{||\mathbf{v}_i||^2}.$$

6. If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors and $\mathbf{w}$ is any vector, then

$$\mathbf{P}(\mathbf{P}(\mathbf{w}, \mathbf{v}), \mathbf{u}) = \mathbf{0}.$$

7. If $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{v}$ are vectors in an inner product space V, then

$$\mathbf{P}(\mathbf{w}_1 + \mathbf{w}_2, \mathbf{v}) = \mathbf{P}(\mathbf{w}_1, \mathbf{v}) + \mathbf{P}(\mathbf{w}_2, \mathbf{v}).$$

Problems

For Problems 1–4, determine whether the given set of vectors is an orthogonal set in $\mathbb{R}^n$. For those that are, determine a corresponding orthonormal set of vectors.

1. $\{(2, -1, 1), (1, 1, -1), (0, 1, 1)\}$.

2. $\{(1, 3, -1, 1), (-1, 1, 1, -1), (1, 0, 2, 1)\}$

3. $\{(1, 2, -1, 0), (1, 0, 1, 2), (-1, 1, 1, 0), (1, -1, -1, 0)\}$.

4. $\{(1, 2, -1, 0, 3), (1, 1, 0, 2, -1), (4, 2, -4, -5, -4)\}$

5. Let $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (1, 1, -1)$. Determine all nonzero vectors $\mathbf{w}$ such that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}\}$ is an orthogonal set. Hence obtain an orthonormal set of vectors in $\mathbb{R}^3$.

For Problems 6–7, show that the given set of vectors is an orthogonal set in $\mathbb{C}^n$, and hence obtain an orthonormal set of vectors in $\mathbb{C}^n$ in each case.

6. $\{(1 - i, 3 + 2i), (2 + 3i, 1 - i)\}$.

7. $\{(1 - i, 1 + i, i), (0, i, 1 - i), (-3 + 3i, 2 + 2i, 2i)\}$.

8. Consider the vectors $\mathbf{v} = (1-i, 1+2i)$, $\mathbf{w} = (2+i, z)$ in $\mathbb{C}^2$. Determine the complex number z such that $\{\mathbf{v}, \mathbf{w}\}$ is an orthogonal set of vectors, and hence obtain an orthonormal set of vectors in $\mathbb{C}^2$.

For Problems 9–10, show that the given functions in $C^0[-1, 1]$ are orthogonal, and use them to construct an orthonormal set of functions in $C^0[-1, 1]$.

9. $f_1(x) = 1$, $f_2(x) = \sin \pi x$, $f_3(x) = \cos \pi x$.

10. $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = \frac{1}{2}(3x^2 - 1)$. These are the **Legendre polynomials** that arise as solutions of the Legendre differential equation

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0,$$

when $n = 0, 1, 2$, respectively.

For Problems 11–12, show that the given functions are orthonormal on $[-1, 1]$.

11. $f_1(x) = \sin \pi x$, $f_2(x) = \sin 2\pi x$, $f_3(x) = \sin 3\pi x$. [**Hint:** The trigonometric identity

$$\sin a \sin b = \tfrac{1}{2}[\cos(a + b) - \cos(a - b)]$$

will be useful.]

12. $f_1(x) = \cos \pi x$, $f_2(x) = \cos 2\pi x$, $f_3(x) = \cos 3\pi x$.

13. Let

$$A_1 = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, \text{ and}$$

$$A_3 = \begin{bmatrix} -1 & -3 \\ 0 & 2 \end{bmatrix}.$$

Use the inner product

$$\langle A, B\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$$

to find all matrices

$$A_4 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

such that $\{A_1, A_2, A_3, A_4\}$ is an orthogonal set of matrices in $M_2(\mathbb{R})$.

For Problems 14–19, use the Gram-Schmidt process to determine an *orthonormal* basis for the subspace of $\mathbb{R}^n$ spanned by the given set of vectors.

14. $\{(1, -1, -1), (2, 1, -1)\}$.

15. $\{(2, 1, -2), (1, 3, -1)\}$.

16. $\{(-1, 1, 1, 1), (1, 2, 1, 2)\}$.

17. $\{(1, 0, -1, 0), (1, 1, -1, 0), (-1, 1, 0, 1)\}$

18. $\{(1, 2, 0, 1), (2, 1, 1, 0), (1, 0, 2, 1)\}$.

19. $\{(1, 1, -1, 0), (-1, 0, 1, 1), (2, -1, 2, 1)\}$.

20. If

$$A = \begin{bmatrix} 3 & 1 & 4 \\ 1 & -2 & 1 \\ 1 & 5 & 2 \end{bmatrix},$$

determine an orthogonal basis for rowspace(A).

For Problems 21–22, determine an *orthonormal* basis for the subspace of $\mathbb{C}^3$ spanned by the given set of vectors. Make sure that you use the appropriate inner product in $\mathbb{C}^3$.

21. $\{(1 - i, 0, i), (1, 1 + i, 0)\}$.

22. $\{(1 + i, i, 2 - i), (1 + 2i, 1 - i, i)\}$.

For Problems 23–25, determine an orthogonal basis for the subspace of $C^0[a, b]$ spanned by the given vectors, for the given interval $[a, b]$.

23. $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$, $a = 0$, $b = 1$.

24. $f_1(x) = 1$, $f_2(x) = x^2$, $f_3(x) = x^4$, $a = -1$, $b = 1$.

25. $f_1(x) = 1$, $f_2(x) = \sin x$, $f_3(x) = \cos x$, $a = -\pi/2$, $b = \pi/2$.

On $M_2(\mathbb{R})$ define the inner product $\langle A, B\rangle$ by

$$\langle A, B\rangle = 5a_{11}b_{11} + 2a_{12}b_{12} + 3a_{21}b_{21} + 5a_{22}b_{22}$$

for all matrices $A = [a_{ij}]$ and $B = [b_{ij}]$. For Problems 26–27, use this inner product in the Gram-Schmidt procedure to determine an orthogonal basis for the subspace of $M_2(\mathbb{R})$ spanned by the given matrices.

26. $A_1 = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 2 & -3 \\ 4 & 1 \end{bmatrix}$.

27. $A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

Also identify the subspace of $M_2(\mathbb{R})$ spanned by $\{A_1, A_2, A_3\}$.

On P_n, define the inner product $\langle p_1, p_2\rangle$ by

$$\langle p_1, p_2\rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$$

for all polynomials

$$p_1(x) = a_0 + a_1x + \cdots + a_nx^n,$$
$$p_2(x) = b_0 + b_1x + \cdots + b_nx^n.$$

For Problems 28–29, use this inner product to determine an orthogonal basis for the subspace of P_n spanned by the given polynomials.

28. $p_1(x) = 1 - 2x + 2x^2$, $p_2(x) = 2 - x - x^2$.

29. $p_1(x) = 1+x^2$, $p_2(x) = 2-x+x^3$, $p_3(x) = 2x^2-x$.

30. Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}\}$ be linearly independent vectors in an inner product space V, and suppose that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal. Define the vector $\mathbf{u}_3$ in V by

$$\mathbf{u}_3 = \mathbf{v} + \lambda\mathbf{u}_1 + \mu\mathbf{u}_2,$$

where λ, μ are scalars. Derive the values of λ and μ such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for the subspace of V spanned by $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}\}$.

31. Prove that if $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is an orthogonal set of vectors in an inner product space V and if $\mathbf{u}_i = \frac{1}{||\mathbf{v}_i||}\mathbf{v}_i$ for each i, then $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ form an orthonormal set of vectors.

32. Prove Lemma 4.12.12.

Let V be an inner product space, and let W be a subspace of V. Set

$$W^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w}\rangle = 0 \text{ for all } \mathbf{w} \in W\}.$$

The set $W^{\perp}$ is called the **orthogonal complement** of W in V. Problems 33–38 explore this concept in some detail. Deeper applications can be found in Project 1 at the end of this chapter.

33. Prove that $W^{\perp}$ is a subspace of V.

34. Let $V = \mathbb{R}^3$ and let

$$W = \text{span}\{(1, 1, -1)\}.$$

Find $W^{\perp}$.

35. Let $V = \mathbb{R}^4$ and let

$$W = \text{span}\{(0, 1, -1, 3), (1, 0, 0, 3)\}.$$

Find $W^{\perp}$.

36. Let $V = M_2(\mathbb{R})$ and let W be the subspace of 2×2 symmetric matrices. Compute $W^{\perp}$.

37. Prove that $W \cap W^{\perp} = \mathbf{0}$. (That is, W and $W^{\perp}$ have no nonzero elements in common.)

38. Prove that if W_1 is a subset of W_2, then $(W_2)^{\perp}$ is a subset of $(W_1)^{\perp}$.

39. The subject of Fourier series is concerned with the representation of a 2π-periodic function f as the following *infinite* linear combination of the set of functions $\{1, \sin nx, \cos nx\}_{n=1}^{\infty}$:

$$f(x) = \tfrac{1}{2}a_0 + \textstyle\sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx). \quad (4.12.7)$$

In this problem, we investigate the possibility of performing such a representation.

(a) Use appropriate trigonometric identities, or some form of technology, to verify that the set of functions

$$\{1, \sin nx, \cos nx\}_{n=1}^{\infty}$$

is orthogonal on the interval $[-\pi, \pi]$.

(b) By multiplying (4.12.7) by $\cos mx$ and integrating over the interval $[-\pi, \pi]$, show that

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx$$

and

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos mx\,dx.$$

[**Hint**: You may assume that interchange of the infinite summation with the integral is permissible.]

(c) Use a similar procedure to show that

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin mx\,dx.$$

It can be shown that if f is in $C^1(-\pi, \pi)$, then Equation (4.12.7) holds for each $x \in (-\pi, \pi)$. The series appearing on the right-hand side of (4.12.7) is called the **Fourier series of** f, and the constants in the summation are called the **Fourier coefficients for** f.

(d) Show that the Fourier coefficients for the function $f(x) = x, -\pi < x \le \pi, f(x + 2\pi) = f(x)$, are

$$a_n = 0, \qquad n = 0, 1, 2, \ldots,$$

$$b_n = -\frac{2}{n}\cos n\pi, \qquad n = 1, 2, \ldots,$$

and thereby determine the Fourier series of f.

(e) ◇ Using some form of technology, sketch the approximations to $f(x) = x$ on the interval $(-\pi, \pi)$ obtained by considering the first three terms, first five terms, and first ten terms in the Fourier series for f. What do you conclude?

4.13 Chapter Review

In this chapter we have derived some basic results in linear algebra regarding vector spaces. These results form the framework for much of linear mathematics. Following are listed some of the chapter highlights.

The Definition of a Vector Space

A vector space consists of four different components:

1. A set of vectors V.
2. A set of scalars F (either the set of real numbers $\mathbb{R}$, or the set of complex numbers $\mathbb{C}$).
3. A rule, $+$, for adding vectors in V.
4. A rule, $\cdot$, for multiplying vectors in V by scalars in F.

Then $(V, +, \cdot)$ is a vector space over F if and only if axioms A1–A10 of Definition 4.2.1 are satisfied. If F is the set of all real numbers, then $(V, +, \cdot)$ is called a *real* vector space, whereas if F is the set of all complex numbers, then $(V, +, \cdot)$ is called a *complex*

vector space. Since it is usually quite clear what the addition and scalar multiplication operations are, we usually specify a vector space by giving only the set of vectors V. The major vector spaces we have dealt with are the following:

$\mathbb{R}^n$	the (real) vector space of all ordered n-tuples of real numbers.
$\mathbb{C}^n$	the (complex) vector space of all ordered n-tuples of complex numbers.
$M_n(\mathbb{R})$	the (real) vector space of all $n \times n$ matrices with real elements.
$C^k(I)$	the vector space of all real-valued functions that are continuous and have (at least) k continuous derivatives on I.
P_n	the vector space of all polynomials of degree $\le n$ with real coefficients.

Subspaces

Usually the vector space V that underlies a given problem is known. It is often one that appears in the list above. However, the solution of a given problem in general involves only a subset of vectors from this vector space. The question that then arises is whether this subset of vectors is itself a vector space under the same operations of addition and scalar multiplication as in V. In order to answer this question, Theorem 4.3.2 tells us that *a nonempty subset of a vector space V is a subspace of V if and only if the subset is closed under addition and closed under scalar multiplication.*

Spanning Sets

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in a vector space V is said to *span* V if *every* vector in V can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$—that is, if for every $\mathbf{v} \in V$, there exist scalars $c_1, c_2, \ldots, c_k$ such that

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

Given a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in a vector space V, we can form the set of *all* vectors that can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. This collection of vectors is a subspace of V called the *subspace spanned by* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, and denoted $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$. Thus,

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \{\mathbf{v} \in V : \mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k\}.$$

Linear Dependence and Linear Independence

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of vectors in a vector space V, and consider the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}. \quad (4.13.1)$$

Clearly this equation will hold if $c_1 = c_2 = \cdots = c_k = 0$. The question of interest is whether there are nonzero values of some or all of the scalars $c_1, c_2, \ldots, c_k$ such that (4.13.1) holds. This leads to the following two ideas:

Linear dependence: There exist scalars $c_1, c_2, \ldots, c_k$, not all zero, such that (4.13.1) holds.

Linear independence: The *only* values of the scalars $c_1, c_2, \ldots, c_k$ such that (4.13.1) holds are $c_1 = c_2 = \cdots = c_k = 0$.

To determine whether a set of vectors is linearly dependent or linearly independent we usually have to use (4.13.1). However, if the vectors are from $\mathbb{R}^n$, then we can use Corollary 4.5.15, whereas for vectors in $C^{k-1}(I)$ the Wronskian can be useful.

34. Let A be an $m \times n$ matrix, let $\mathbf{v} \in \text{colspace}(A)$ and let $\mathbf{w} \in \text{nullspace}(A^T)$. Prove that $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

35. Let W denote the set of all 3×3 skew-symmetric matrices.

(a) Show that W is a subspace of $M_3(\mathbb{R})$.

(b) Find a basis and the dimension of W.

(c) Extend the basis you constructed in part (b) to a basis for $M_3(\mathbb{R})$.

36. Let W denote the set of all 3×3 matrices whose rows and columns add up to zero.

(a) Show that W is a subspace of $M_3(\mathbb{R})$.

(b) Find a basis and the dimension of W.

(c) Extend the basis you constructed in part (b) to a basis for $M_3(\mathbb{R})$.

37. Let $(V, +_V, \cdot_V)$ and $(W, +_W, \cdot_W)$ be vector spaces and define

$$V \oplus W = \{(\mathbf{v}, \mathbf{w}) : \mathbf{v} \in V \text{ and } \mathbf{w} \in W\}.$$

Prove that

(a) $V \oplus W$ is a vector space, under componentwise operations.

(b) Via the identification $\mathbf{v} \mapsto (\mathbf{v}, 0)$, V is a subspace of $V \oplus W$, and likewise for W.

(c) If $\dim[V] = n$ and $\dim[W] = m$, then $\dim[V \oplus W] = m + n$. **[Hint:** Write a basis for $V \oplus W$ in terms of bases for V and W.]

38. Show that a basis for P_3 need not contain a polynomial of each degree 0, 1, 2, 3.

39. Prove that if A is a matrix whose nullspace and column space are the same, then A must have an even number of columns.

40. Let

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix}.$$

Prove that if all entries $b_1, b_2, \dots, b_n$ and $c_1, c_2, \dots, c_n$ are nonzero, then the $n \times n$ matrix $A = BC$ has nullity $n - 1$.

For Problems 41–44, find a basis and the dimension for the row space, column space, and null space of the given matrix A.

41. $A = \begin{bmatrix} -3 & -6 \\ -6 & -12 \end{bmatrix}$.

42. $A = \begin{bmatrix} -1 & 6 & 2 & 0 \\ 3 & 3 & 1 & 5 \\ 7 & 21 & 7 & 15 \end{bmatrix}$.

43. $A = \begin{bmatrix} -4 & 0 & 3 \\ 0 & 10 & 13 \\ 6 & 5 & 2 \\ -2 & 5 & 10 \end{bmatrix}$.

44. $A = \begin{bmatrix} 3 & 5 & 5 & 2 & 0 \\ 1 & 0 & 2 & 2 & 1 \\ 1 & 1 & 1 & -2 & -2 \\ -2 & 0 & -4 & -2 & -2 \end{bmatrix}$.

For Problems 45–46, find an orthonormal basis for the row space, column space, and null space of the given matrix A.

45. $A = \begin{bmatrix} 1 & 2 & 6 \\ 2 & 1 & 6 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{bmatrix}$.

46. $A = \begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$.

For Problems 47–50, find an orthogonal basis for the span of the set S, where S is given in

47. Problem 25.

48. Problem 26.

49. Problem 29, using $p \cdot q = \int_0^1 p(t)q(t)\,dt$.

50. Problem 32, using the inner product defined in Problem 4 of Section 4.11.

For Problems 51–54, determine the angle between the given vectors $\mathbf{u}$ and $\mathbf{v}$ using the standard inner product on $\mathbb{R}^n$.

51. $\mathbf{u} = (2, 3)$ and $\mathbf{v} = (4, -1)$.

52. $\mathbf{u} = (-2, -1, 2, 4)$ and $\mathbf{v} = (-3, 5, 1, 1)$.

53. Repeat Problems 51–52 for the inner product on $\mathbb{R}^n$ given by

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 + u_2v_2 + u_3v_3 + \cdots + u_nv_n.$$

54. Let $t_0, t_1, \dots, t_n$ be real numbers. For p and q in P_n, define

$$p \cdot q = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n).$$

(a) Prove that $p \cdot q$ defines a valid inner product on P_n.

(b) Let $t_0 = -3$, $t_1 = -1$, $t_2 = 1$, and $t_3 = 3$. Let $p_0(t) = 1$, $p_1(t) = t$, and $p_2(t) = t^2$. Find a polynomial q that is orthogonal to p_0 and p_1, such that $\{p_0, p_1, q\}$ is an orthogonal basis for $\text{span}\{p_0, p_1, p_2\}$.

55. Find the distance from the point $(2, 3, 4)$ to the line in $\mathbb{R}^3$ passing through $(0, 0, 0)$ and $(6, -1, -4)$.

56. Let V be an inner product space with basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. If $\mathbf{x}$ and $\mathbf{y}$ are vectors in V such that $\mathbf{x} \cdot \mathbf{v}_i = \mathbf{y} \cdot \mathbf{v}_i$ for each $i = 1, 2, \dots, n$, prove that $\mathbf{x} = \mathbf{y}$.

57. State as many conditions as you can on an $n \times n$ matrix A that are equivalent to its invertibility.

Project I: Orthogonal Complement

Let V be an inner product space and let W be a subspace of V.

Part 1 Definition Let

$$W^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}.$$

Show that $W^\perp$ is a subspace of V and that $W^\perp$ and W share only the zero vector: $W^\perp \cap W = \{\mathbf{0}\}$.

Part 2 Examples

(a) Let $V = M_2(\mathbb{R})$ with inner product

$$\left\langle \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}.$$

Find the orthogonal complement of the set W of 2×2 *symmetric matrices*.

(b) Let A be an $m \times n$ matrix. Show that

$$(\text{rowspace}(A))^\perp = \text{nullspace}(A)$$

and

$$(\text{colspace}(A))^\perp = \text{nullspace}(A^T).$$

Use this to find the orthogonal complement of the row space and column space of the matrices below:

(i) $A = \begin{bmatrix} 3 & 1 & -1 \\ 6 & 0 & -4 \end{bmatrix}$.

(ii) $A = \begin{bmatrix} -1 & 0 & 6 & 2 \\ 3 & -1 & 0 & 4 \\ 1 & 1 & 1 & -1 \end{bmatrix}$.

(c) Find the orthogonal complement of

(i) the line in $\mathbb{R}^3$ containing the points $(0, 0, 0)$ and $(2, -1, 3)$.

(ii) the plane $2x + 3y - 4z = 0$ in $\mathbb{R}^3$.

Part 3 Some Theoretical Results Let W be a subspace of a finite-dimensional inner product space V.

(a) Show that every vector in V can be written *uniquely* in the form $\mathbf{w} + \mathbf{w}^{\perp}$, where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. [**Hint:** By Gram-Schmidt, $\mathbf{v}$ can be projected onto the subspace W as, say, $\text{proj}_W(\mathbf{v})$, and so $\mathbf{v} = \text{proj}_W(\mathbf{v}) + \mathbf{w}^{\perp}$, where $w^{\perp} \in W^{\perp}$. For the uniqueness, use the fact that $W \cap W^{\perp} = \{\mathbf{0}\}$.]

(b) Use part (a) to show that

$$\dim[V] = \dim[W] + \dim[W^{\perp}].$$

(c) Show that

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

Project II: Line-Fitting Data Points

Suppose data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ in the xy-plane have been collected. Unless these data points are collinear, there will be no line that contains all of them. We wish to find a line, commonly known as a **least-squares line**, that approximates the data points as closely as possible.

How do we go about finding such a line? The approach we take[12] is to write the line as $y = mx + b$, where m and b are unknown constants.

Part 1 Derivation of the Least-Squares Line

(a) By substituting the data points (x_i, y_i) for x and y in the equation $y = mx + b$, show that the matrix equation $A\mathbf{x} = \mathbf{y}$ is obtained, where

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} m \\ b \end{bmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Unless the data points are collinear, the system $A\mathbf{x} = \mathbf{y}$ obtained in part (a) has no solution for $\mathbf{x}$. In other words, the vector $\mathbf{y}$ does not lie in the column space of A. The goal then becomes to find $\mathbf{x}_0$ such that the distance $||\mathbf{y} - A\mathbf{x}_0||$ is as small as possible. This will happen precisely when $\mathbf{y} - A\mathbf{x}_0$ is perpendicular to the column space of A. In other words, for all $\mathbf{x} \in \mathbb{R}^2$, we must have

$$(A\mathbf{x}) \cdot (\mathbf{y} - A\mathbf{x}_0) = 0.$$

(b) Using the fact that the dot product of vectors $\mathbf{u}$ and $\mathbf{v}$ can be written as a matrix multiplication,

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

show that

$$(A\mathbf{x}) \cdot (\mathbf{y} - A\mathbf{x}_0) = \mathbf{x} \cdot (A^T\mathbf{y} - A^T A\mathbf{x}_0).$$

(c) Conclude that

$$A^T\mathbf{y} = A^T A\mathbf{x}_0.$$

Provided that A has linearly independent columns, the matrix $A^T A$ is invertible (see Problem 34, in Section 4.13).

[12] We can also obtain the least-squares line by using optimization techniques from multivariable calculus, but the goal here is to illustrate the use of linear systems and projections.

(d) Show that the least-squares solution is

$$\mathbf{x}_0 = (A^T A)^{-1} A^T \mathbf{y}$$

and therefore,

$$A\mathbf{x}_0 = A(A^T A)^{-1} A^T \mathbf{y}$$

is the point in the column space of A that is closest to $\mathbf{y}$. Therefore, it is the **projection** of $\mathbf{y}$ onto the column space of A, and we write

$$A\mathbf{x}_0 = A(A^T A)^{-1} A^T \mathbf{y} = P\mathbf{y},$$

where

$$P = A(A^T A)^{-1} A^T \quad (4.13.2)$$

is called a **projection matrix.** If A is $m \times n$, what are the dimensions of P?

(e) Referring to the projection matrix P in (4.13.2), show that $PA = A$ and $P^2 = P$. Geometrically, why are these facts to be expected? Also show that P is a symmetric matrix.

Part 2 Some Applications In parts (a)–(d) below, find the equation of the least-squares line to the given data points.

(a) $(0, -2), (1, -1), (2, 1), (3, 2), (4, 2)$.

(b) $(-1, 5), (1, 1), (2, 1), (3, -3)$.

(c) $(-4, -1), (-3, 1), (-2, 3), (0, 7)$.

(d) $(-3, 1), (-2, 0), (-1, 1), (0, -1), (2, -1)$.

In parts (e)–(f), by using the ideas in this project, find the distance from the point P to the given plane.

(e) $P(0, 0, 0)$; $2x - y + 3z = 6$.

(f) $P(-1, 3, 5)$; $-x + 3y + 3z = 8$.

Part 3 A Further Generalization Instead of fitting data points to a least-squares line, one could also attempt to do a parabolic approximation of the form $ax^2 + bx + c$. By following the outline in Part 1 above, try to determine a procedure for finding the best parabolic approximation to a set of data points. Then try out your procedure on the data points given in Part 2, (a)–(d).

Problems

1. $\text{nullspace}(A) = \{(x, y, z, w) : x - 6z - w = 0\}$; $\text{nullity}(A) = 3$; $\text{rank}(A) = 1$.
3. $\text{nullspace}(A) = \{\mathbf{0}\}$; $\text{nullity}(A) = 0$; $\text{rank}(A) = 3$.
5. 1.
7. 1.
9. $\mathbf{x} = c\mathbf{x}_1 + \mathbf{x}_p$, where $\mathbf{x}_1 = (34, -11, 1)$ and $\mathbf{x}_p = (-5, 3, 0)$.
11. $\mathbf{x} = \mathbf{x}_p = (2, -3, 1)$.
13. No.

Section 4.10

True-False Review

1. True
3. False
5. True
7. False
9. False

Section 4.11

True-False Review

1. False
3. True
5. False
7. False

Problems

1. $\theta = 0.95$ rad.
3. $\langle \mathbf{v}, \mathbf{w} \rangle = 19 + 11i$, $||\mathbf{v}|| = \sqrt{35}$, $||\mathbf{v}|| = \sqrt{22}$.
7. $\langle A, B \rangle = 13$, $||A|| = \sqrt{13}$, $||B|| = \sqrt{7}$.
11. (a) 9.
    (b) 0.

Section 4.12

True-False Review

1. True
3. True
5. True
7. True

Problems

1. Orthonormal set: $\left\{ \frac{1}{\sqrt{6}}(2, -1, 1), \frac{1}{\sqrt{3}}(1, 1, -1), \frac{1}{\sqrt{2}}(0, 1, 1) \right\}$.
3. Not orthogonal.
5. Orthonormal set: $\left\{ \frac{1}{\sqrt{14}}(1, 2, 3), \frac{1}{\sqrt{3}}(1, 1, -1), \frac{1}{\sqrt{42}}(5, -4, 1) \right\}$.
7. Orthonormal set: $\left\{ \frac{1}{\sqrt{5}}(1 - i, 1 + i, i), \frac{1}{\sqrt{3}}(0, i, 1 - i), \frac{1}{\sqrt{30}}(-3 + 3i, 2 + 2i, 2i) \right\}$.
9. Orthonormal set: $\left\{ \frac{1}{\sqrt{2}}, \sin \pi x, \cos \pi x \right\}$.
13. $A_4 = k \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$.
15. Orthonormal basis: $\left\{ \frac{1}{3}(2, 1, -2), \frac{1}{3\sqrt{2}}(-1, 4, 1) \right\}$.
17. Orthonormal basis: $\{ \frac{1}{\sqrt{2}}(1, 0, -1, 0), (0, 1, 0, 0), \frac{1}{\sqrt{6}}(-1, 0, -1, 2) \}$.
19. Orthonormal basis: $\left\{ \frac{1}{\sqrt{3}}(1, 1, -1, 0), \frac{1}{\sqrt{15}}(-1, 2, 1, 3), \frac{1}{\sqrt{15}}(3, -1, 2, 1) \right\}$.
21. Orthonormal basis: $\left\{ \frac{1}{\sqrt{3}}(1 - i, 0, i), \frac{1}{\sqrt{21}}(1, 3 + 3i, 1 - i) \right\}$.
23. Orthogonal basis: $\left\{ 1, \frac{1}{2}(2x - 1), \frac{1}{6}(6x^2 - 6x + 1) \right\}$.
25. Orthogonal basis: $\left\{ 1, \sin x, \frac{1}{\pi}(\pi \cos x - 2) \right\}$.
27. Orthogonal basis: $\left\{ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$, the subspace of all symmetric matrices in $M_2(\mathbb{R})$.
29. Orthogonal basis: $\left\{ 1 + x^2, 1 - x - x^2 + x^3, -3 - 5x + 3x^2 + x^3 \right\}$.
35. $W^\perp = \text{span}\{(0, 1, 1, 0), (-3, -3, 0, 1)\}$.

Section 4.13

Additional Problems

3. No.
5. No.
7. No.
9. Yes.
11. No.
13. No.
19. No.
21. No.
23. Yes.
25. (b) only.
27. (b) only.
29. (a) and (b).
31. (b) only.
35. (b) One possible basis:
$$\left\{ \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \right\}.$$
    (c) Add E_{11}, E_{22}, E_{33}, E_{12}, E_{13}, and E_{23}.
41. Basis for rowspace(A): $\{(-3, -6)\}$. Basis for colspace(A): $\{(-3, -6)\}$. Basis for nullspace(A): $\{(-2, 1)\}$.
43. Basis for rowspace(A): $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. Basis for colspace($A$): $\{(-4, 0, 6, -2), (0, 10, 5, 5), (3, 13, 2, 10)\}$. The part about the basis for nullspace(A) should be $\emptyset$.