Chapter 5: Inner Product Spaces

Cheng Chang, Ph.D.
Department of Math & CS
Mercy College

Contents

1. Length and Dot Product in $\mathbb{R}^n$

Definition 1.1. *The **length**, or **norm**, of a vector $\vec{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ is*

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

*The length of a vector is also called its **magnitude**. If $\|\vec{v}\| = 1$, then the vector $\vec{v}$ is a **unit vector**.*

Example 1.1. *In $\mathbb{R}^5$, find the length of the vector $\vec{v} = (0, -2, 1, 4, -2)$. In $\mathbb{R}^3$, find the length of $\vec{v} = (2/\sqrt{17}, -2\sqrt{17}, 3\sqrt{17})$.*

Two vectors are parallel if they are constant multiples of each other. Moreover, if $\vec{u} = c\vec{v}$ where $c > 0$, we say $\vec{u}$ and $\vec{v}$ have the same direction. If $c < 0$ then they have opposite direction.

Theorem 1.1. *Let $\vec{v}$ be a vector in $\mathbb{R}^n$ and let c be a scalar. Then*

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

where $|c|$ is the absolute value of c.

Sometimes we need to find the unit vector of a given vector in the same direction, the following theorem deals with that.

Theorem 1.2. *If $\vec{v}$ is a nonzero vector in $\mathbb{R}^n$, then the vector*

$$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$$

*has length 1 and has the same direction as $\vec{v}$. This vector $\vec{u}$ is the **unit vector in the direction of** $\vec{v}$.*

Example 1.2. *Find the unit vector in the direction of $\vec{v} = (3, -1, 2)$.*

In $\mathbb{R}^2$ (or the plane), we know the distance formula between two points (x_1, y_1) and (x_2, y_2) is

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Now we generalize this idea, and define the distance formula in $\mathbb{R}^n$.

Definition 1.2. *The distance between two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ is*

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

In $\mathbb{R}^2$, suppose there are two vectors $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$, the angle between these two vectors, denoted θ, is given by

$$\cos\theta = \frac{u_1 v_1 + u_2 v_2}{\|\vec{u}\|\|\vec{v}\|}$$

Inspired by this angle formula, we define a new operation in $\mathbb{R}^n$.

Definition 1.3 (Dot Product in $\mathbb{R}^n$). *The **dot product** of $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ is the scalar quantity*

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Example 1.3. *Given $\vec{u} = (1, 2, 0, -3)$ and $\vec{v} = (3, -2, 4, 2)$, find $\vec{u} \cdot \vec{v}$.*

Theorem 1.3. *If $\vec{u}, \vec{v}$ and $\vec{w}$ are vectors in $\mathbb{R}^n$ and c is a scalar, then the properties listed below are true.*

(1) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
(2) $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$
(3) $c(\vec{u} \cdot \vec{v}) = (c\vec{u}) \cdot \vec{v} = \vec{u} \cdot (c\vec{v})$
(4) $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$
(5) $\vec{v} \cdot \vec{v} \geq 0$, *and* $\vec{v} \cdot \vec{v} = 0$ *if and only if* $\vec{v} = \vec{0}$.

Example 1.4. *Let* $\vec{u} = (2, -2), \vec{v} = (5, 8)$, *and* $\vec{w} = (-4, 3)$. *Find each quantity.*
(1) $\vec{u} \cdot \vec{v}$
(2) $(\vec{u} \cdot \vec{v})\vec{w}$
(3) $\|\vec{w}\|^2$
(4) $\vec{u} \cdot (\vec{v} - 2\vec{w})$

Example 1.5. *Consider two vectors* $\vec{u}$ *and* $\vec{v}$ *in* $\mathbb{R}^n$ *such that* $\vec{u} \cdot \vec{u} = 39$, $\vec{u} \cdot \vec{v} = -3$, *and* $\vec{v} \cdot \vec{v} = 79$. *Evaluate* $(\vec{u} + 2\vec{v}) \cdot (3\vec{u} + \vec{v})$.

Next, we try to get a generalized angle formula in $\mathbb{R}^n$. The idea is to use the following formula

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|}$$

However, there is a minor detail needed to be taken care of, saying that the absolute value of $\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|}$ should be bounded by 1. The next theorem will address that.

Theorem 1.4 (The Cauchy-Schwarz Inequality)**.** *If* $\vec{u}$ *and* $\vec{v}$ *are vectors in* $\mathbb{R}^n$, *then*

$$|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\|\|\vec{v}\|$$

where $|\vec{u} \cdot \vec{v}|$ *denote the absolute value of* $\vec{u} \cdot \vec{v}$.

A direct application of the Cauchy-Schwarz Inequality shows that

$$\left| \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|} \right| \leq 1$$

So we can officially define the angle formula next.

Definition 1.4. *The* ***angle*** θ *between two nonzero vectors in* $\mathbb{R}^n$ *can be found using*

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|}, \quad 0 \leq \theta \leq \pi.$$

Example 1.6. *Find the angle between* $\vec{u} = (-4, 0, 2, -2)$ *and* $\vec{v} = (2, 0, -1, 1)$.

Definition 1.5. *Two vectors* $\vec{u}$ *and* $\vec{v}$ *in* $\mathbb{R}^n$ *are* ***orthogonal*** *when*

$$\vec{u} \cdot \vec{v} = 0.$$

Example 1.7. *Show that the vectors are orthogonal to each other.*

$$\vec{u} = (3, 2, -1, 4), \quad \vec{v} = (1, -1, 1, 0)$$

Example 1.8. *Determine all vectors in* $\mathbb{R}^2$ *that are orthogonal to* $\vec{u} = (4, 2)$.

In $\mathbb{R}^2$, we have a well known property called the Triangle Inequality, which says that the sum of the two adjacent sides is greater than the opposite side. Now in $\mathbb{R}^n$, we want to generalize this, by using the Cauchy-Schwarz Inequality.

Theorem 1.5 (The Triangle Inequality)**.** *If* $\vec{u}$ *and* $\vec{v}$ *are vectors in* $\mathbb{R}^n$, *then*

$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|.$$

From the proof of the Triangle Inequality, we have

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

So when $\vec{u}$ and $\vec{v}$ are orthogonal, we have the following generalized Pythagorean Theorem.

Theorem 1.6. *If $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^n$, then $\vec{u}$ and $\vec{v}$ are orthogonal if and only if*

$$\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2.$$

Finally, we look at the relationship between two operations, one is the dot product and the other is the matrix multiplication.

It is often useful to represent vectors in $\mathbb{R}^n$ as column vectors. For example,

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

In this notation, the dot product of two vectors $\vec{u}\cdot\vec{v}$ can be represented as the matrix product of the transpose of $\vec{u}$ multiplied by $\vec{v}$.

$$\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v}.$$

2. Inner Product Spaces

In last section, the concepts of length, distance, and angle were extended from $\mathbb{R}^2$ to $\mathbb{R}^n$. This section extends these concepts one step further—to general vector spaces —by using the idea of an inner product of two vectors.

We've seen an example of an inner product: the dot product in $\mathbb{R}^n$. The dot product, called the **Euclidean inner product**, is only one of several inner products that can be defined on $\mathbb{R}^n$. To distinguish between the standard inner product and other possible inner products, use the notation below.

$$\vec{u}\cdot\vec{v} = \text{ dot product (Euclidean inner product for } \mathbb{R}^n)$$

$$\langle\vec{u},\vec{v}\rangle = \text{ general inner product for a vector space } V$$

A general inner product is defined in much the same way that a general vector space is defined —that is, in order for a function to qualify as an inner product, it must satisfy a set of axioms.

Definition 2.1. *Let $\vec{u},\vec{v}$, and $\vec{w}$ be vectors in a vector space V, and let c be any scalar. An **inner product** on V is a function that associates a real number $\langle\vec{u},\vec{v}\rangle$ with each pair of vectors $\vec{u}$ and $\vec{v}$ and satisfies the axioms listed below.*

(1) $\langle\vec{u},\vec{v}\rangle = \langle\vec{v},\vec{u}\rangle$
(2) $\langle\vec{u},\vec{v}+\vec{w}\rangle = \langle\vec{u},\vec{v}\rangle + \langle\vec{u},\vec{w}\rangle$
(3) $c\langle\vec{u},\vec{v}\rangle = \langle c\vec{u},\vec{v}\rangle$
(4) $\langle\vec{v},\vec{v}\rangle \geq 0$*, and* $\langle\vec{v},\vec{v}\rangle = 0$ *if and only if* $\vec{v}=\vec{0}$.

A vector space V, with an inner product on V, is called an inner product space. So inner product spaces are "special" vector spaces.

Example 2.1. *Show that the function below defines an inner product on $\mathbb{R}^2$, where $\vec{u}=(u_1,u_2)$ and $\vec{v}=(v_1,v_2)$.*

$$\langle\vec{u},\vec{v}\rangle = u_1v_1 + 2u_2v_2$$

Example 2.2. *Show that the function below is not an inner product on $\mathbb{R}^3$, where $\vec{u}=(u_1,u_2,u_3)$ and $\vec{v}=(v_1,v_2,v_3)$.*

$$\langle\vec{u},\vec{v}\rangle = u_1v_1 - 2u_2v_2 + u_3v_3$$

Example 2.3. *Let $A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{12}\end{bmatrix}$ and $B=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22}\end{bmatrix}$ be matrices in the vector space $M_{2,2}$. The function*

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

is an inner product on $M_{2,2}$.

Example 2.4. *Let f and g be real-valued continuous functions in the vector space $C[a,b]$. Show that*

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

defines an inner product on $C[a,b]$.

The next theorem lists some properties of inner products.

Theorem 2.1. *Let $\vec{u},\vec{v}$, and $\vec{w}$ be vectors in an inner product space V, and let c be any real number.*

(1) $\langle \vec{0}, \vec{v} \rangle = \langle \vec{v}, \vec{0} \rangle = 0$
(2) $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$
(3) $\langle \vec{u}, c\vec{v} \rangle = c\langle \vec{u}, \vec{v} \rangle$

The definition of length, distance, and angle for general inner product spaces closely parallel those for Euclidean $\mathbb{R}^n$-space.

Definition 2.2. *Let $\vec{u}$ and $\vec{v}$ be vectors in an inner product space V.*
(1) The ***length*** *(or norm) of $\vec{u}$ is $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}$.*
(2) The ***distance*** *between $\vec{u}$ and $\vec{v}$ is $d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$.*
(3) The ***angle*** *between two nonzero vectors $\vec{u}$ and $\vec{v}$ can be found using*

$$\cos\theta = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|\|\vec{v}\|}, \quad 0 \le \theta \le \pi.$$

(4) $\vec{u}$ and $\vec{v}$ are ***orthogonal*** *when $\langle \vec{u}, \vec{v} \rangle = 0$.*

If $\|\vec{u}\| = 1$, then $\vec{u}$ is a **unit vector**. Moreover, if $\vec{v}$ is any nonzero vector in an inner product space V, then the vector $\vec{u} = \vec{v}/\|\vec{v}\|$ is the unit vector in the direction of $\vec{v}$.

Example 2.5. *For polynomials $p = a_0 + a_1x + \cdots + a_nx^n$ and $q = b_0 + b_1x + \cdots + b_nx^n$ in the vector space P_n, the function $\langle p, q \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$ is an inner product. Let $p(x) = 1 - 2x^2, q(x) = 4 - 2x + x^2$, and $r(x) = x + 2x^2$ be polynomials in P_2, and find each quantity.*

(1) $\langle p, q \rangle$ *(2)* $\langle q, r \rangle$ *(3)* $\|q\|$ *(4)* $d(p, q)$

The next theorem lists the general inner product space versions of the Cauchy-Schwarz Inequality, the triangle inequality, and the Pythagorean Theorem.

Theorem 2.2. *Let $\vec{u}$ and $\vec{v}$ be vectors in an inner product space V.*
(1) Cauchy-Schwarz Inequality: $|\langle \vec{u}, \vec{v} \rangle| \le \|\vec{u}\|\|\vec{v}\|$
(2) Triangle inequality: $\|\vec{u} + \vec{v}\| \le \|\vec{u}\| + \|\vec{v}\|$
(3) Pythagorean Theorem: $\vec{u}$ and $\vec{v}$ are orthogonal if and only if

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2.$$

Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^2$. If $\vec{v}$ is nonzero, then $\vec{u}$ can be orthogonally projected onto $\vec{v}$. This projection is denoted by $\operatorname{proj}_{\vec{v}} \vec{u}$ and is a scalar multiple of $\vec{v}$.

Definition 2.3. *Let $\vec{u}$ and $\vec{v}$ be vectors in an inner product space V, such that $\vec{v} \ne \vec{0}$. Then the* ***orthogonal projection*** *of $\vec{u}$ onto $\vec{v}$ is*

$$\operatorname{proj}_{\vec{v}} \vec{u} = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}.$$

Example 2.6. *Use the Euclidean inner product in $\mathbb{R}^3$ to find the orthogonal projection of $\vec{u} = (6, 2, 4)$ onto $\vec{v} = (1, 2, 0)$.*

Example 2.7. *Let $f(x) = 1$ and $g(x) = x$ be functions in $C[0, 1]$. Use the inner product on $C[a, b]$ defined as*

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

to find the orthogonal projection of f onto g.

Theorem 2.3. *Let $\vec{u}$ and $\vec{v}$ be two vectors in an inner product space V, such that $\vec{v} \neq \vec{0}$. Then*

$$d(\vec{u}, \operatorname{proj}_{\vec{v}} \vec{u}) < d(\vec{u}, c\vec{v}), \quad c \neq \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}.$$

3. Orthonormal Bases: Gram-Schmidt Process

Recall the standard basis for $\mathbb{R}^3$, $\{(1,0,0),(0,1,0),(0,0,1)\}$. It has important characteristics that are particularly useful. One important characteristic is that the three vectors are mutually orthogonal. A second important characteristic is that each vector in the basis is a unit vector.

This section identifies some advantages of using bases consisting of mutually orthogonal unit vectors and develops a procedure for constructing such bases, know as the Gram-Schmidt orthonormalization process.

Definition 3.1. *A set S of vectors in an inner product space V is* ***orthogonal*** *when every pair of vectors in S is orthogonal. If, in addition, each vector in the set is a unit vector, then S is* ***orthonormal****.*

Example 3.1. *Show that the set is an orthonormal basis for $\mathbb{R}^3$.*

$$S = \left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right), \left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3} \right), \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3} \right) \right\}$$

Example 3.2. *Show that in P_3, with the inner product*

$$\langle p, q \rangle = a_0b_0 + a_1b_1 + a_2b_2 + a_3b_3$$

the standard basis $B = \{1, x, x^2, x^3\}$ is orthonormal.

Example 3.3. *In $C[0, 2\pi]$, with the inner product*

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)\,dx$$

show that the set $S = \{1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots, \sin nx, \cos nx\}$ is orthogonal.

Each set in the previous examples is linearly independent. This is a characteristic of any orthogonal set of nonzero vectors, as stated in the next theorem.

Theorem 3.1. *If $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.*

Theorem 3.2. *If V is an inner product space of dimension n, then any orthogonal set of n nonzero vectors is a basis for V.*

Example 3.4. *Show that the set S below is a basis for $\mathbb{R}^4$.*

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

In the last chapter, we discussed a procedure for finding a coordinate representation relative to a nonstandard basis. When the basis is orthonormal, this procedure can be streamlined.

Theorem 3.3. *If $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for an inner product space V, then the coordinate representation of a vector $\vec{w}$ relative to B is*

$$\vec{w} = \langle \vec{w}, \vec{v}_1 \rangle \vec{v}_1 + \cdots + \langle \vec{w}, \vec{v}_n \rangle \vec{v}_n.$$

Example 3.5. *Find the coordinate matrix of $\vec{w} = (5, -5, 2)$ relative to the orthonormal basis B for $\mathbb{R}^3$ below.*

$$B = \left\{ \left(\frac{3}{5}, \frac{4}{5}, 0 \right), \left(-\frac{4}{5}, \frac{3}{5}, 0 \right), (0,0,1) \right\}$$

Having seen one of the advantages of orthonormal bases, we will now look at a procedure for finding such a basis.

Theorem 3.4 (**Gram-Schmidt Orthonormalization Process**).

(1) Let $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$ *be a basis for an inner product space* V*.*

(2) Let $B' = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n\}$*, where*

$$\begin{aligned}
\vec{w}_1 &= \vec{v}_1 \\
\vec{w}_2 &= \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{w}_1 \rangle}{\langle \vec{w}_1, \vec{w}_1 \rangle}\vec{w}_1 \\
\vec{w}_3 &= \vec{v}_3 - \frac{\langle \vec{v}_3, \vec{w}_1 \rangle}{\langle \vec{w}_1, \vec{w}_1 \rangle}\vec{w}_1 - \frac{\langle \vec{v}_3, \vec{w}_2 \rangle}{\langle \vec{w}_2, \vec{w}_2 \rangle}\vec{w}_2 \\
&\vdots \\
\vec{w}_n &= \vec{v}_n - \frac{\langle \vec{v}_n, \vec{w}_1 \rangle}{\langle \vec{w}_1, \vec{w}_1 \rangle}\vec{w}_1 - \cdots - \frac{\langle \vec{v}_n, \vec{w}_{n-1} \rangle}{\langle \vec{w}_{n-1}, \vec{w}_{n-1} \rangle}\vec{w}_{n-1}.
\end{aligned}$$

Then B' *is an orthogonal basis for* V*.*

(3) Let $\vec{u}_i = \frac{\vec{w}_i}{\|\vec{w}_i\|}$*. Then* $B'' = \{\vec{u}_1, \ldots, \vec{u}_n\}$ *is an orthonormal basis for* V*. Also,* $\mathrm{Span}\{\vec{v}_1, \ldots, \vec{v}_n\} = \mathrm{Span}\{\vec{u}_1, \ldots, \vec{u}_n\}$ *for* $k = 1, 2, \ldots, n$*.*

Example 3.6. *Apply the Gram-Schmidt orthonormalization process to the basis* B *for* $\mathbb{R}^2$ *below.*

$$B = \{(1,1), (0,1)\}$$

Example 3.7. *The vectors*

$$\vec{v}_1 = (0,1,0) \text{ and } \vec{v}_2 = (1,1,1)$$

span a plane in $\mathbb{R}^3$*. Find an orthonormal basis for this subspace.*

Example 3.8. *Apply the Gram-Schmidt orthonormalization process to the basis* $B = \{1, x, x^2\}$ *in* P_2*, using the inner product*

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