

Week 6

Inner Product Space

Christopher Ting

<http://cting.x10host.com/>

Hiroshima University

✉: cting@hiroshima-u.ac.jp

☎: +81 082-424-6451

🏠: A1棟 131-1

Table of Contents

- 1 Introduction
- 2 Inner Product
- 3 Orthonormal Basis
- 4 Compliment Space
- 5 Orthogonal Matrices
- 6 Takeaways

Quotable Quote

We share a philosophy about linear algebra: we think basis-free, we write basis-free, but when the chips are down we close the office door and compute with matrices like fury.

— Irving Kaplansky

Learning Outcomes

- 📌 Discuss the difference, if any, between the inner product vis-à-vis the standard inner product.
- 📌 Create a mind map to interpret the Cauchy–Schwarz inequality and the triangular inequality and develop a scheme to show how these two inequalities are related.
- 📌 Examine the notions of norm, length, and the angle formed by two vectors by clarifying their relationships.
- 📌 Develop an intuition for a pair of vectors being orthogonal to each other and connect the concept of orthogonality to linear independence.
- 📌 Apply the Gram–Schmidt process to obtain an orthonormal basis and discuss QR decomposition.
- 📌 Define orthogonal matrix and orthogonal transformation and discuss their relationships with norm and orthonormal basis.



Inner Product and Inner Product Space

Definition 2.1 (Inner Product).

With respect to two vectors \mathbf{a} and \mathbf{b} of a linear space V , when a real number (scalar) satisfies the following four conditions, then we can define the **inner product** of \mathbf{a} and \mathbf{b} and denote it as $\langle \mathbf{a}, \mathbf{b} \rangle$:

- (1) $\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle$
- (2) $\langle \mathbf{a} + \mathbf{b}, \mathbf{c} \rangle = \langle \mathbf{a}, \mathbf{c} \rangle + \langle \mathbf{b}, \mathbf{c} \rangle$
- (3) $\langle s\mathbf{a}, \mathbf{b} \rangle = s\langle \mathbf{a}, \mathbf{b} \rangle$
- (4) If $\mathbf{a} \neq \mathbf{0}$, then $\langle \mathbf{a}, \mathbf{a} \rangle \geq 0$.

Here, $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$, and $s \in \mathfrak{R}$.

-  A linear space that is defined by the inner product is called the **inner product space**.
-  For a vector \mathbf{a} that belongs to an inner product space, the length or **norm** of \mathbf{a} is defined as $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$.

Standard Inner Product

Two vectors in \mathfrak{R}^n are defined as $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

Define the inner product $\mathbf{a}'\mathbf{b}$ as

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}'\mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

This inner product is called the **standard inner product** of \mathfrak{R}^n because it satisfies

- (1) $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} = \langle \mathbf{b}, \mathbf{a} \rangle$
- (2) $\langle \mathbf{a} + \mathbf{b}, \mathbf{c} \rangle = (\mathbf{a} + \mathbf{b})'\mathbf{c} = \mathbf{a}'\mathbf{c} + \mathbf{b}'\mathbf{c} = \langle \mathbf{a}, \mathbf{c} \rangle + \langle \mathbf{b}, \mathbf{c} \rangle$
- (3) $\langle s\mathbf{a}, \mathbf{b} \rangle = (s\mathbf{a})'\mathbf{b} = s(\mathbf{a}'\mathbf{b}) = s\langle \mathbf{a}, \mathbf{b} \rangle$
- (4) If $\mathbf{a} \neq \mathbf{0}$, then $\langle \mathbf{a}, \mathbf{a} \rangle = a_1^2 + a_2^2 + \cdots + a_n^2 \geq 0$.

Square Matrix and Standard Inner Product

🐟 Suppose A is a square matrix. With respect to the standard inner product,
 $\langle Aa, b \rangle = \langle a, A'b \rangle$.

🐟 In fact,

$$\langle Aa, b \rangle = (Aa)'b = (a'A')b = a'(A'b) = \langle a, A'b \rangle.$$

Inner Product of Functions

🐟 Let $C(I)$ be the linear space of all real-valued functions that are continuous over the interval $I = [a, b]$.

🐟 For any arbitrary $f, g \in C(I)$, we define

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

🐟 This is an inner product of $C(I)$.

Properties of Norm

Theorem 2.2.


For any vectors \mathbf{a} and \mathbf{b} of an inner product space, and for any real-value scalar c , the following statements hold:

(1) $\|c\mathbf{a}\| = |c| \|\mathbf{a}\|$

(2) $|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\| \|\mathbf{b}\|$ (The Cauchy–Schwarz inequality)

(3) $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$ (The triangular inequality)

Proof of (1).

 Since $\|c\mathbf{a}\|^2 = \langle c\mathbf{a}, c\mathbf{a} \rangle = c^2 \langle \mathbf{a}, \mathbf{a} \rangle = c^2 \|\mathbf{a}\|^2$, after taking the square root on both sides, we have $\|c\mathbf{a}\| = |c| \|\mathbf{a}\|$.



Properties of Norm (Cont'd)

Proof of (2) The Cauchy–Schwarz inequality.

- 🐟 The case of $\mathbf{a} = \mathbf{0}$ gives rise to 0 on both sides and thus they are equal.
- 🐟 Now, suppose $\mathbf{a} \neq \mathbf{0}$ and let $f(t) = \|t\mathbf{a} + \mathbf{b}\|^2$.
- 🐟 Hence, $f(t) = \langle t\mathbf{a} + \mathbf{b}, t\mathbf{a} + \mathbf{b} \rangle = \|\mathbf{a}\|^2 t^2 + 2\langle \mathbf{a}, \mathbf{b} \rangle t + \|\mathbf{b}\|^2$, which is quadratic in t .
- 🐟 For $f(t)$ to be positive, we need

$$(2\langle \mathbf{a}, \mathbf{b} \rangle)^2 - 4\|\mathbf{a}\|^2\|\mathbf{b}\|^2 < 0.$$

- 🐟 It follows that $4\langle \mathbf{a}, \mathbf{b} \rangle^2 < 4\|\mathbf{a}\|^2\|\mathbf{b}\|^2$.
- 🐟 Taking the square root, we obtain $|\langle \mathbf{a}, \mathbf{b} \rangle| < \|\mathbf{a}\|\|\mathbf{b}\|$.



Properties of Norm (Cont'd)

Proof of (3) The triangular inequality.

🐟 Consider $\|\mathbf{a} + \mathbf{b}\|^2$, which is $\langle \mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b} \rangle$.

🐟 Upon expansion, we obtain $\|\mathbf{a}\|^2 + 2\langle \mathbf{a}, \mathbf{b} \rangle + \|\mathbf{b}\|^2$.

🐟 It follows that

$$\begin{aligned}\|\mathbf{a} + \mathbf{b}\|^2 &= \|\mathbf{a}\|^2 + 2|\langle \mathbf{a}, \mathbf{b} \rangle| + \|\mathbf{b}\|^2 \\ &\leq \|\mathbf{a}\|^2 + 2\|\mathbf{a}\|\|\mathbf{b}\| + \|\mathbf{b}\|^2 \quad (\text{by the Cauchy-Schwarz inequality}) \\ &\leq (\|\mathbf{a}\| + \|\mathbf{b}\|)^2.\end{aligned}$$

🐟 The upshot is that $\|\mathbf{a} + \mathbf{b}\|^2 \leq (\|\mathbf{a}\| + \|\mathbf{b}\|)^2$.

🐟 Taking the square root on both sides, we obtain the triangular inequality.



Angle Formed by Two Vectors

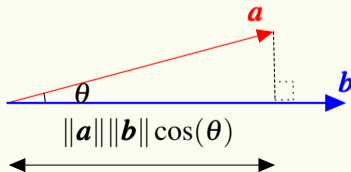
🐟 By the Cauchy-Schwarz inequality,

$$-\|a\|\|b\| \leq \langle a, b \rangle \leq \|a\|\|b\|.$$

🐟 Dividing by $\|a\|\|b\|$, we obtain

$$-1 \leq \frac{\langle a, b \rangle}{\|a\|\|b\|} \leq 1.$$

🐟 It follows that $-1 \leq \cos(\theta) \leq 1$, with $0 \leq \theta \leq \pi$, as θ is the angle formed by a and b .



Orthogonal System

Definition 3.1 (Orthogonal).

When $\langle \mathbf{a}, \mathbf{b} \rangle = 0$ for two vectors of an inner product space, they are said to be **orthogonal**, and we denote this special case as $\mathbf{a} \perp \mathbf{b}$.

Definition 3.2 (Orthogonal System).

Consider r non-null vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ of a linear space V , which are mutually orthogonal. That is

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0, \quad \text{for } i \neq j, 1 \leq i, j \leq r.$$

Then $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$ is said to form an **orthogonal system**.

Examples

🐟 The 3 vectors in \mathfrak{R}^3 , i.e., $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, constitute an orthogonal system.

Example 3.3.

Consider the linear space of all functions that are continuous on the interval $[-\pi, \pi]$, with the inner product defined by

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) dx.$$

Show that, for non-negative integer n , $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(nx), \sin(nx)$ form an orthogonal system.

Solution

➤ First consider, for $k = 1, 2, \dots$,

$$\langle 1, \cos(kx) \rangle = \int_{-\pi}^{\pi} \cos(kx) dx = 0$$

$$\langle 1, \sin(kx) \rangle = \int_{-\pi}^{\pi} \sin(kx) dx = 0$$

➤ From trigonometry, we have, for $k, \ell = 1, 2, \dots$,

$$\sin(kx) \sin(\ell x) = \frac{1}{2} (\cos(kx - \ell x) - \cos(kx + \ell x)),$$

$$\cos(kx) \cos(\ell x) = \frac{1}{2} (\cos(kx - \ell x) + \cos(kx + \ell x)),$$

$$\sin(kx) \cos(\ell x) = \frac{1}{2} (\sin(kx - \ell x) + \sin(kx + \ell x)).$$

Solution (Cont'd)

⇒ Accordingly,

$$\int_{-\pi}^{\pi} \sin(kx) \sin(\ell x) dx = \int_{-\pi}^{\pi} \sin(kx) \sin(\ell x) dx = 0 \quad (k \neq \ell),$$
$$\int_{-\pi}^{\pi} \cos(kx) \sin(\ell x) dx = 0.$$

⇒ In other words,

$$\langle \cos(kx), \sin(\ell x) \rangle = 0,$$

and for $k \neq \ell$

$$\langle \cos(kx), \cos(\ell x) \rangle = \langle \sin(kx), \sin(\ell x) \rangle = 0.$$

⇒ Therefore, we can conclude that $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(nx), \sin(nx)$ form an orthogonal system.

Orthogonality and Independence

Theorem 3.4 (Orthogonality Implies Linear Independence).

If $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ of an inner product space V form an orthogonal system, then $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ are linearly independent.

Proof.

✦ By assumption, $\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0$ when $i \neq j$. To examine independence, we consider, with $c_1, c_2, \dots, c_n \in \mathfrak{R}$,

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = \mathbf{0}.$$

✦ Now, take the inner product of both sides with \mathbf{a}_i . Since $\langle \mathbf{a}_i, \mathbf{0} \rangle = 0$, it must be that

$$\langle \mathbf{a}_i, c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n \rangle = \sum_{j=1}^r c_j \langle \mathbf{a}_i, \mathbf{a}_j \rangle = c_i \langle \mathbf{a}_i, \mathbf{a}_i \rangle = c_i \|\mathbf{a}_i\|^2.$$

✦ Since it is any i , for $c_i \|\mathbf{a}_i\|^2 = 0$, as $\|\mathbf{a}_i\|^2 \neq 0$, it can only be that $c_i = 0, i = 1, 2, \dots, r$.



Orthonormal System and Basis

Definition 3.5 (Orthonormal System and Basis).

Consider r non-null vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ of an inner product space V , which are mutually orthogonal and of unit length. That is

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

Then $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ is said to form an **orthonormal system**. Moreover, if the basis of V is an orthonormal system, then it is called the **orthonormal basis**.

Theorem 3.6.

From a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ of an n -dimensional inner product space, it is possible to construct an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$.

Example of Orthonormal Basis

Example 3.7.

- For the trigonometric basis of Example 3.3 we compute the square norm of each element of the basis:

$$\int_{-\pi}^{\pi} 1^2 dx = 2\pi, \quad \int_{-\pi}^{\pi} \sin^2(kx) dx = \pi, \quad \int_{-\pi}^{\pi} \cos^2(kx) dx = \pi,$$

where k is a natural number.

- Dividing 1 by $\sqrt{2\pi}$, $\sin(kx)$ and $\cos(kx)$ by $\sqrt{\pi}$ for $k = 1, 2, \dots, n$, we obtain the orthonormal basis as

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\cos(2x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}}, \dots, \frac{\cos(nx)}{\sqrt{\pi}}, \frac{\sin(nx)}{\sqrt{\pi}}.$$

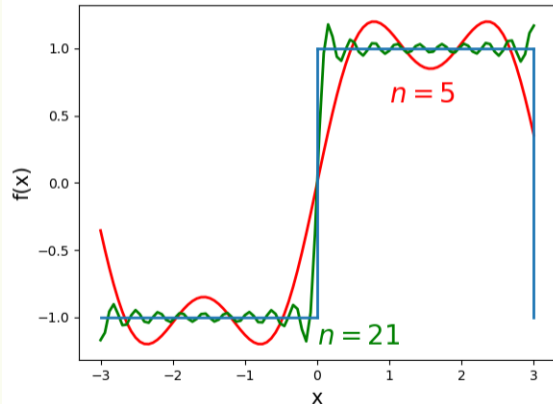
- In **Fourier's series**, these functions are employed to approximate a function.

Application of Orthonormal Basis 1

Example 3.8.

- Let n be an odd number.
- Define a function by using the vectors of Example 3.7:

$$f(x) = \frac{4}{\pi} \sin(x) + \frac{4}{3\pi} \sin(3x) + \dots$$
$$\dots + \frac{4}{n\pi} \sin(nx)$$

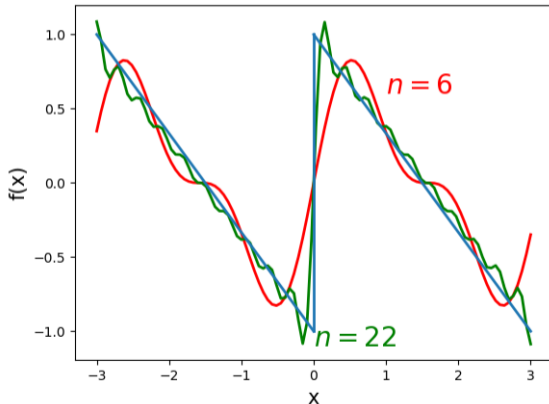


Application of Orthonormal Basis 2

Example 3.9.

- Let n be an even number.
- Define a function by using the vectors of Example 3.7:

$$f(x) = \frac{4}{2\pi} \sin(2x) + \frac{4}{4\pi} \sin(4x) + \dots \\ \dots + \frac{4}{n\pi} \sin(nx)$$



Orthogonal Projection and Remainder

➤ Let V be an inner product space. Let $\mathbf{x}, \mathbf{v} \in V$ and $\mathbf{v} \neq \mathbf{0}$.

➤ The **orthogonal projection** of the vector \mathbf{x} onto the vector \mathbf{v} is given by

$$\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$$

➤ It is described as “orthogonal” because the **remainder** defined as

$$\mathbf{q} = \mathbf{x} - \mathbf{p}$$

is orthogonal to \mathbf{v} .

➤ Proof:
$$\langle \mathbf{v}, \mathbf{q} \rangle = \langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{v}, \mathbf{p} \rangle = \langle \mathbf{v}, \mathbf{x} \rangle - \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle = 0.$$

Orthogonal Projection and Remainder (Cont'd)

➤ In general, if $\{\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ is an orthogonal set of vectors, i.e.,

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \quad \text{when } i \neq j,$$

then the orthogonal projection of the vector \mathbf{x} onto the subspace spanned by $\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$ is

$$\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 + \frac{\langle \mathbf{x}, \mathbf{v}_3 \rangle}{\langle \mathbf{v}_3, \mathbf{v}_3 \rangle} \mathbf{v}_3 + \dots + \frac{\langle \mathbf{x}, \mathbf{v}_n \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle} \mathbf{v}_n.$$

➤ The orthogonality comes from the fact that the remainder $\mathbf{q} = \mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$.

➤ If we write \mathbf{x} as \mathbf{v}_1 , we have **the Gram–Schmidt orthogonalization**.

The Gram–Schmidt Process

Theorem 3.10 (The Gram–Schmidt Process).

It is possible to construct an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ from an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ of n -dimensional inner product space.

Proof of The Gram–Schmidt Process

Proof by construction.

➤ First, set $\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$, which yields $\|\mathbf{u}_1\| = 1$.

➤ Next, define $\tilde{\mathbf{v}}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{u}_1 \rangle \mathbf{u}_1$, and $\mathbf{u}_2 = \frac{\tilde{\mathbf{v}}_2}{\|\tilde{\mathbf{v}}_2\|}$.

➤ It can be readily shown that $\|\mathbf{u}_2\| = 1$ and $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$.

➤ In general, in the same fashion, for $1 \leq h < n$,

$$\tilde{\mathbf{v}}_{h+1} = \mathbf{v}_{h+1} - \sum_{i=1}^h \langle \mathbf{v}_{h+1}, \mathbf{u}_i \rangle \mathbf{u}_i, \quad \mathbf{u}_{h+1} = \frac{\tilde{\mathbf{v}}_{h+1}}{\|\tilde{\mathbf{v}}_{h+1}\|}, \text{ which has the required properties of}$$
$$\|\mathbf{u}_{h+1}\| = 1, \text{ and } \langle \mathbf{u}_{h+1}, \mathbf{u}_i \rangle = 0 \quad (1 \leq h < n).$$

➤ Repeat the same procedure until $h = n - 1$.



Example of the Gram–Schmidt Process

➤ Suppose the basis of \mathfrak{R}^3 is

$$\left\{ \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} \right\}$$

➤ First, let $\mathbf{u}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.

Example of the Gram–Schmidt Process (Cont'd)

Next, define $\tilde{\mathbf{a}}_2 = \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} - \frac{4}{\sqrt{6}} \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \frac{5}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, so

$$\mathbf{u}_2 = \frac{\tilde{\mathbf{a}}_2}{\|\tilde{\mathbf{a}}_2\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

Example of the Gram–Schmidt Process (Cont'd)

➤ Finally, define $\tilde{\mathbf{a}}_3 = \mathbf{a}_3 - \langle \mathbf{a}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{a}_3, \mathbf{u}_2 \rangle \mathbf{u}_2$. Specifically,

$$\tilde{\mathbf{a}}_3 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} - \frac{5}{\sqrt{6}} \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - \frac{-5}{\sqrt{3}} \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \frac{3}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

➤ It follows that $\mathbf{u}_3 = \frac{\tilde{\mathbf{a}}_3}{\|\tilde{\mathbf{a}}_3\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

➤ In this way, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is obtained as an orthonormal basis of \mathfrak{R}^3 .

QR Decomposition

- Suppose $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n]$. Using the Gram–Schmidt process, an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ is obtained.
- With suitable constants r_{ij} where $1 \leq i, j \leq n$, the vector \mathbf{a}_i can be written as a linear combination of the orthonormal basis.

$$\mathbf{a}_1 = r_{11}\mathbf{q}_1$$

$$\mathbf{a}_2 = r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2$$

$$\vdots$$

$$\mathbf{a}_n = r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n$$

QR Decomposition (Cont'd)

➤ In matrix form,

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} =: \mathbf{QR}$$

➤ We call $\mathbf{A} = \mathbf{QR}$ the **QR decomposition** of matrix \mathbf{A} .

Remark 1

- With respect to the basis $\{1, x, x^2, \dots\}$, the polynomial obtained by applying the Gram–Schmidt process is

$$f_n(x) = \sqrt{\frac{2n+1}{2}} L_n(x),$$

where **the Legendre polynomial** is defined as

$$L_n(x) := \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n,$$

for $n = 0, 1, 2, \dots$

Remark 2

➤ Suppose the inner product is defined instead as

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2} dx.$$

➤ By applying the Gram–Schmidt process, we obtain an orthogonal basis $\{1, 2x, 4x^2 - 2, \dots\}$, which is called **the Hermite polynomial**.

➤ Interestingly, the wavefunction of a harmonic oscillator in quantum mechanics can be expressed by **the Hermite polynomial**.

Orthogonal Complement Space

✎ For a subspace W of an inner product space V , consider the set of all vectors in V that are orthogonal to every vector in W . That is

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

Then, W^\perp is said to be the **orthogonal complement space** of W .

✎ W^\perp is also a subspace of V .

Proof:

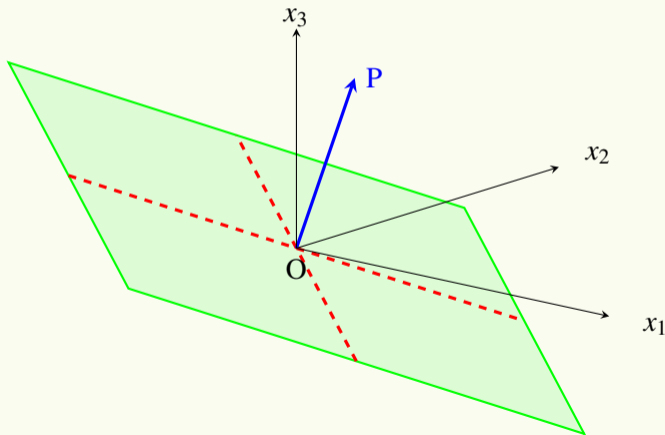
- Let $\mathbf{v}_1, \mathbf{v}_2$ be any two vectors in W^\perp and c a scalar. For any $\mathbf{w} \in W$,


$$\langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{w} \rangle = \langle \mathbf{v}_1, \mathbf{w} \rangle + \langle \mathbf{v}_2, \mathbf{w} \rangle = 0 + 0 = 0.$$

$$\langle c\mathbf{v}_1, \mathbf{w} \rangle = c\langle \mathbf{v}_1, \mathbf{w} \rangle = 0.$$

- Hence, $\mathbf{v}_1 + \mathbf{v}_2 \in W^\perp$ and $c\mathbf{v}_1 \in W^\perp$.
- We can conclude that W^\perp is a subspace of V .

Illustration of Orthogonal Complement Space



 The complement space of the vector \overrightarrow{OP} is the plane.

Example

Example 4.1.

🚀 Suppose the subspace of \mathcal{R}^4 is W and it is generated by two vectors:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 3 \end{bmatrix} \quad \text{and} \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 2 \end{bmatrix}$$

🚀 Find the orthogonal complement space W^\perp .

🚀 Let the vector of W^\perp be $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

Example (Cont'd)


✎ Since $\langle \mathbf{x}, \mathbf{a}_1 \rangle = \langle \mathbf{x}, \mathbf{a}_2 \rangle = 0$, we obtain a system of linear equations:

$$\begin{cases} x_1 - 2x_2 + 3x_4 = 0 \\ x_1 - x_2 + x_3 + 2x_4 = 0. \end{cases}$$

✎ The coefficient matrix is

$$\begin{bmatrix} 1 & -2 & 0 & 3 \\ 1 & -1 & 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & 1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

Example (Cont'd)

 For any real constants, the general solution is

$$\begin{cases} x_1 = -2x_3 - x_4 \\ x_2 = -x_3 + x_4 \end{cases} \implies \mathbf{x} = c_1 \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Orthogonal Matrix

⏏ For a real-valued n -dimensional square matrix \mathbf{A} , if

$$\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I},$$

then \mathbf{A} is said to be an **orthogonal matrix**.

Theorem 5.1.

Let the n -dimensional square matrix \mathbf{A} be expressed in terms of its column vectors $[\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n]$. Then the following 4 statements are equivalent.

- 1 \mathbf{A} is an orthogonal matrix.
- 2 $\|\mathbf{A}\mathbf{a}\| = \|\mathbf{a}\|$, for $\mathbf{a} \in \mathfrak{R}^n$.
- 3 $\langle \mathbf{A}\mathbf{a}, \mathbf{A}\mathbf{b} \rangle = \langle \mathbf{a}, \mathbf{b} \rangle$, for $\mathbf{a}, \mathbf{b} \in \mathfrak{R}^n$.
- 4 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ is the orthonormal basis of \mathfrak{R}^n .

Orthogonal Transformation

⏏ For a transformation T of an inner product space V , with regard to the vectors \mathbf{a} and \mathbf{b} in V , if the following is satisfied:

$$\langle T(\mathbf{a}), T(\mathbf{b}) \rangle = \langle \mathbf{a}, \mathbf{b} \rangle,$$

then T is said to be an **orthogonal transformation**.

Theorem 5.2.

Let $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ be an orthonormal basis of an inner product space V . Then

T is orthogonal. $\iff \{T(\mathbf{a}_1), T(\mathbf{a}_2), \dots, T(\mathbf{a}_n)\}$ is an orthonormal basis of V .

Orthogonal matrix \equiv Orthogonal transformation

Theorem 5.3.

For an n -dimensional real-valued square matrix \mathbf{A} , if the linear transformation on \mathfrak{R}^n is defined as $T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{x} \in \mathfrak{R}^n$, then

\mathbf{A} is an orthogonal matrix. $\iff T_{\mathbf{A}}$ is an orthogonal transformation.

Takeaways

- ✂ Inner product is bilinear, and it maps two (abstract) vectors into a scalar.
- ✂ Norm is a generalization of length.
- ✂ Standard inner product applies to the usual vectors.
- ✂ Inner product is essentially a projection of one vector onto another vector, and thus the inner product cannot be more than the product of two norms.
- ✂ The triangular inequality is essentially saying that the length of the “direct flight” to the final destination is shorter than the total length of two “indirect flights”.
- ✂ A pair of orthogonal vectors form an “angle” of 90° , which is also the intuitive meaning of linear independence.
- ✂ The Gram–Schmidt process is very useful in constructing an orthonormal basis.

Takeaways (Cont'd)

- ✂ QR decomposition facilitates the computation of the determinant.
- ✂ Polynomial is a linear combination of the basis “vectors”, which are powers of x .
- ✂ A linear space can be decomposed by a subspace and its orthogonal complement space.
- ✂ The inverse matrix of an orthogonal matrix is its transpose.
- ✂ Orthogonal transformation is defined as a transformation that does not change the inner product of two vectors.
- ✂ Orthogonal matrix is equivalent to orthogonal transformation, and its column vectors form the orthonormal basis of the Euclidean space.

Keywords

Fourier's series, 19
QR Decomposition, 29, 30
The Cauchy–Schwarz inequality, 9
The Gram–Schmidt Process, 24, 25
The triangular inequality, 9
inner product, 5
inner product space, 5
norm, 5
orthogonal, 13
orthogonal complement space, 33
orthogonal matrix, 38

orthogonal projection, 22
orthogonal system, 13
orthogonal transformation, 39
orthonormal basis, 18
orthonormal system, 18
remainder, 22
standard inner product, 6
the Gram–Schmidt orthogonalization, 23
the Hermite polynomial, 32
the Legendre polynomial, 31