# Lesson 11
# Inner Product Space

**Christopher Ting**

: cting@hiroshima-u.ac.jp

: +81 082-424-7268

: http://cting.x10host.com/

Christopher Ting

June 1, 2020 Version 2.3

# Table of Contents

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

(1) $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{b}, \boldsymbol{a} \rangle$

(2) $\langle \boldsymbol{a} + \boldsymbol{b}, \boldsymbol{c} \rangle = \langle \boldsymbol{a}, \boldsymbol{c} \rangle + \langle \boldsymbol{b}, \boldsymbol{c} \rangle$

(3) $\langle s\boldsymbol{a}, \boldsymbol{b} \rangle = s\langle \boldsymbol{a}, \boldsymbol{b} \rangle$

(4) If $\boldsymbol{a} \neq \boldsymbol{0}$, then $\langle \boldsymbol{a}, \boldsymbol{a} \rangle \geqq 0$.

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

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

# Standard Inner Product

- Two vectors in $\mathfrak{R}^n$ are defined as $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.
- Define the inner product $\boldsymbol{a}'\boldsymbol{b}$ as

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a}'\boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

- This inner product is called the **standard inner product** of $\mathfrak{R}^n$ because it satisfies
  1. $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a}'\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{a} = \langle \boldsymbol{b}, \boldsymbol{a} \rangle$
  2. $\langle \boldsymbol{a} + \boldsymbol{b}, \boldsymbol{c} \rangle = (\boldsymbol{a} + \boldsymbol{b})'\boldsymbol{c} = \boldsymbol{a}'\boldsymbol{c} + \boldsymbol{b}'\boldsymbol{c} = \langle \boldsymbol{a}, \boldsymbol{c} \rangle + \langle \boldsymbol{b}, \boldsymbol{c} \rangle$
  3. $\langle s\boldsymbol{a}, \boldsymbol{b} \rangle = (s\boldsymbol{a})'\boldsymbol{b} = s(\boldsymbol{a}'\boldsymbol{b}) = s\langle \boldsymbol{a}, \boldsymbol{b} \rangle$
  4. If $\boldsymbol{a} \neq \boldsymbol{0}$, then $< \boldsymbol{a}, \boldsymbol{a} >= a_1^2 + a_2^2 + \cdots + a_n^2 \geqq 0$.

# Square Matrix and Standard Inner Product

- Suppose $\boldsymbol{A}$ is a square matrix. With respect to the standard inner product, $\langle \boldsymbol{A}\boldsymbol{a}, \boldsymbol{b}\rangle = \langle \boldsymbol{a}, \boldsymbol{A}'\boldsymbol{b}\rangle$.

- In fact,

$$\langle \boldsymbol{A}\boldsymbol{a}, \boldsymbol{b}\rangle = (\boldsymbol{A}\boldsymbol{a})'\boldsymbol{b} = \left(\boldsymbol{a}'\boldsymbol{A}'\right)\boldsymbol{b} = \boldsymbol{a}'\left(\boldsymbol{A}'\boldsymbol{b}\right) = \langle \boldsymbol{a}, \boldsymbol{A}'\boldsymbol{b}\rangle.$$

# 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 $\boldsymbol{a}$ and $\boldsymbol{b}$ of an inner product space, and for any real-value scalar $c$, the following statements hold:*

(1) $\|c\boldsymbol{a}\| = |c|\,\|\boldsymbol{a}\|$

(2) $|\langle \boldsymbol{a}, \boldsymbol{b} \rangle| \leqq \|\boldsymbol{a}\|\|\boldsymbol{b}\|$ (**The Cauchy–Schwarz inequality**)

(3) $\|\boldsymbol{a}+\boldsymbol{b}\| \leqq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$ (**The triangular inequality**)

**Proof of (1).**

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

□

# Properties of Norm (Cont'd)

## Proof of (2) The Cauchy–Schwarz inequality.

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

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

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

□

# Properties of Norm (Cont'd)

## Proof of (3) The triangular inequality.

- Consider $\|\boldsymbol{a}+\boldsymbol{b}\|^2$, which is $\langle \boldsymbol{a}+\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}\rangle$.
- Upon expansion, we obtain $\|\boldsymbol{a}\|^2 + 2\langle \boldsymbol{a}, \boldsymbol{b}\rangle + \|\boldsymbol{b}\|^2$.
- It follows that

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

- The upshot is that $\|\boldsymbol{a}+\boldsymbol{b}\|^2 \leqq \big(\|\boldsymbol{a}\| + \|\boldsymbol{b}\|\big)^2$.
- Taking the square root on both sides, we obtain the triangular inequality.

□

# Angle Formed by Two Vectors

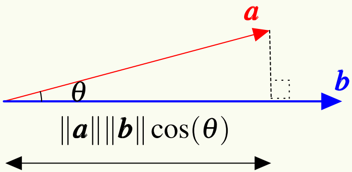
- By the Cauchy-Schwarz inequality,

$$-\|\boldsymbol{a}\|\|\boldsymbol{b}\| \leqq \langle \boldsymbol{a}, \boldsymbol{b} \rangle \leqq \|\boldsymbol{a}\|\|\boldsymbol{b}\|.$$

- Dividing by $\|\boldsymbol{a}\|\|\boldsymbol{b}\|$, we obtain

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

- It follows that $-1 \leqq \cos(\theta) \leqq 1$, with $0 \leqq \theta \leqq \pi$, as $\theta$ is the angle formed by $\boldsymbol{a}$ and $\boldsymbol{b}$.



# Orthogonal System

**Definition 3.1 (Orthogonal).**

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

**Definition 3.2 (Orthogonal System).**

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

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = 0, \qquad \text{for } i \neq j,\ 1 \leqq i, j \leqq r.$$

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

# Examples

- The 3 vectors in $\mathfrak{R}^3$, i.e., $\boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ and $\boldsymbol{c} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, 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 $n$, $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots, \cos(nx), \sin(nx)$ form an orthogonal system.

# Solution

- First consider, for $k = 1, 2, \ldots$,

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

- From trigonometry, we have, for $k, \ell = 1, 2, \ldots$,

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

# Solution (Cont'd)

- Accordingly,

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

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

- 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 $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots, \cos(nx), \sin(nx)$ form an orthogonal system.

# Orthogonality and Independence

**Theorem 3.4 (Orthogonality Implies Linear Independence).**

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

Proof.

- By assumption, $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = 0$ when $i \neq j$. To examine independence, we consider, with $c_1, c_2, \ldots, c_n \in \mathfrak{R}$, $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_n\boldsymbol{a}_n = \mathbf{0}$.
- Now, take the inner product of both sides with $\boldsymbol{a}_i$. Since $\langle \boldsymbol{a}_i, \mathbf{0} \rangle = 0$, it must be that

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

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

□

# Orthonormal System and Basis

**Definition 3.5 (Orthonormal System and Basis).**

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

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

Then $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r$ 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 $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ of an $n$-dimensional inner product space, it is possible to construct an orthonormal basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$.*

# 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 $k = 1, 2, \ldots, n$, we obtain the orthonormal basis as

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

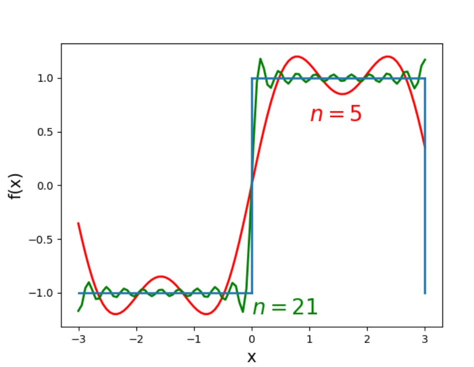
- In **Fourier's series**, these functions are employed to approximate a function.

# 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:

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

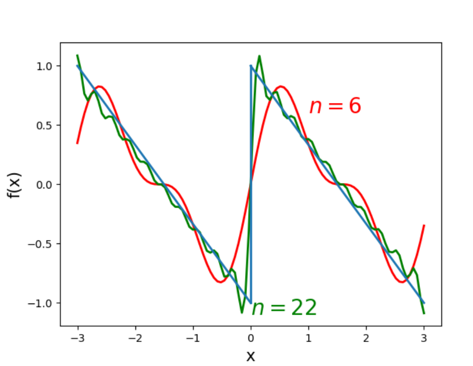


# Application of Orthonormal Basis 2

## Example 3.9.

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

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



# Orthogonal Projection and Remainder

- Let $V$ be an inner product space. Let $\boldsymbol{x}, \boldsymbol{v} \in V$ and $\boldsymbol{v} \neq \mathbf{0}$.
- The **orthogonal projection** of the vector $\boldsymbol{x}$ onto the vector $\boldsymbol{v}$ is given by

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

- It is described as "orthogonal" because the **remainder** defined as

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

  is orthogonal to $\boldsymbol{v}$.
- Proof:

$$\langle \boldsymbol{v}, \boldsymbol{q} \rangle = \langle \boldsymbol{v}, \boldsymbol{x} \rangle - \langle \boldsymbol{v}, \boldsymbol{p} \rangle = \langle \boldsymbol{v}, \boldsymbol{x} \rangle - \frac{\langle \boldsymbol{x}, \boldsymbol{v} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{x} \rangle - \langle \boldsymbol{x}, \boldsymbol{v} \rangle = 0.$$

# Orthogonal Projection and Remainder (Cont'd)

- In general, if $\{\boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set of vectors, i.e.,

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

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

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

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

- If we write $\boldsymbol{x}$ as $\boldsymbol{v}_1$, 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* $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ *from an orthogonal basis* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *of* $n$*-dimensional inner product space.*

**Proof by construction.**

- First, set $\boldsymbol{u}_1 = \dfrac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}$, which yields $\|\boldsymbol{u}_1\| = 1$.
- Next, define $\widetilde{\boldsymbol{v}}_2 := \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1$, and $\boldsymbol{u}_2 = \dfrac{\widetilde{\boldsymbol{v}}_2}{\|\widetilde{\boldsymbol{v}}_2\|}$.
- It can be readily shown that $\|\boldsymbol{u}_2\| = 1$ and $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle = 0$.
- In general, in the same fashion, for $1 \leqq h < n$,

  $$\widetilde{\boldsymbol{v}}_{h+1} = \boldsymbol{v}_{h+1} - \sum_{i=1}^{h} \langle \boldsymbol{v}_{h+1}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i, \quad \boldsymbol{u}_{h+1} = \frac{\widetilde{\boldsymbol{v}}_{h+1}}{\|\widetilde{\boldsymbol{v}}_{h+1}\|},$$

  which has the required properties of $\|u_{h+1}\| = 1$, and $\langle \boldsymbol{u}_{h+1}, \boldsymbol{u}_i \rangle = 0 \ \ (1 \leqq h < n)$.
- Repeat the same procedure until $h = n - 1$. □

# Example of the Gram–Schmidt Process

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

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

- First, let $\boldsymbol{u}_1 = \dfrac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|} = \dfrac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.
- Next, define

$$\widetilde{\boldsymbol{a}}_2 = \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} - \frac{4}{\sqrt{6}} \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \frac{5}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \text{ so}$$

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

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

- Finally, define $\widetilde{\boldsymbol{a}}_3 = \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2$. Specifically,

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

- It follows that $\boldsymbol{u}_3 = \dfrac{\widetilde{\boldsymbol{a}}_3}{\|\widetilde{\boldsymbol{a}}_3\|} = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.
- In this way, $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is obtained as an orthonormal basis of $\mathfrak{R}^3$.

# QR Decomposition

- Suppose $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$. Using the Gram–Schmidt process, an orthonormal basis $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n\}$ is obtained.

- With suitable constants $r_{ij}$ where $1 \leqq i, j \leqq n$, the vector $\boldsymbol{a}_i$ can be written as a linear combination of the orthonormal basis.

$$
\begin{aligned}
\boldsymbol{a}_1 &= r_{11}\boldsymbol{q}_1 \\
\boldsymbol{a}_2 &= r_{12}\boldsymbol{q}_1 + r_{22}\boldsymbol{q}_2 \\
&\vdots \\
\boldsymbol{a}_n &= r_{1n}\boldsymbol{q}_1 + r_{2n}\boldsymbol{q}_2 + \cdots + r_{nn}\boldsymbol{q}_n
\end{aligned}
$$

# QR Decomposition (Cont'd)

- In matrix form,

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

- We call $\boldsymbol{A} = \boldsymbol{QR}$ the **QR decomposition** of matrix $\boldsymbol{A}$.

# Remark 1

- With respect to the basis $\{1, x, x^2, \ldots,\}$, 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 $n = 0, 1, 2, \ldots$.

# 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

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

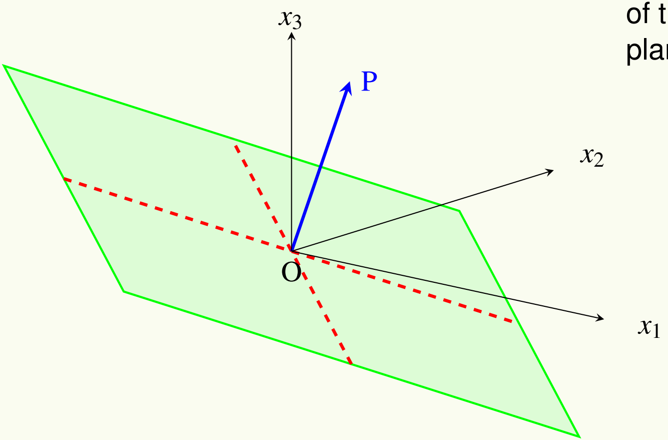
  Then, $W^{\perp}$ is said to be the **orthogonal complement space** of $W$.

- $W^{\perp}$ is also a subspace of $V$.
  Proof:
  - Let $\boldsymbol{v}_1, \boldsymbol{v}_2$ be any two vectors in $W^{\perp}$ and $c$ a scalar. For any $\boldsymbol{w} \in W$,

$$\begin{aligned}\langle \boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w} \rangle &= \langle \boldsymbol{v}_1, \boldsymbol{w} \rangle + \langle \boldsymbol{v}_2, \boldsymbol{w} \rangle = 0 + 0 = 0. \\ \langle c\boldsymbol{v}_1, \boldsymbol{w} \rangle &= c\langle \boldsymbol{v}_1, \boldsymbol{w} \rangle = 0.\end{aligned}$$

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

# Illustration of Orthogonal Complement Space




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

# Example

## Example 3.11.

- Suppose the subspace of $\Re^4$ is $W$ and it is generated by two vectors:

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

- Find the orthogonal complement space $W^\perp$.

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

# Example (Cont'd)

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

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

- The coefficient matrix is

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

- For any real constants, the general solution is

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

# Orthogonal Matrix

For a real-valued $n$-dimensional square matrix $\boldsymbol{A}$, if

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

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

**Theorem 4.1.**

*Let the $n$-dimensional square matrix $\boldsymbol{A}$ be expressed in terms of its column vectors $\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$. Then the following 4 statements are equivalent.*

1. *$\boldsymbol{A}$ is an orthogonal matrix.*
2. *$\|\boldsymbol{A}\boldsymbol{a}\| = \|\boldsymbol{a}\|$, for $\boldsymbol{a} \in \Re^n$.*
3. *$\langle \boldsymbol{A}\boldsymbol{a}, \boldsymbol{A}\boldsymbol{b} \rangle = \langle \boldsymbol{a}, \boldsymbol{b} \rangle$, for $\boldsymbol{a}, \boldsymbol{b} \in \Re^n$.*
4. *$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is the orthonormal basis of $\Re^n$.*

# Orthogonal Transformation

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

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

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

**Theorem 4.2.**

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

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

# Orthogonal matrix $\equiv$ Orthogonal transformation

**Theorem 4.3.**

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

$\boldsymbol{A}$ *is an orthogonal matrix.* $\iff$ $T_{\boldsymbol{A}}$ *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 $90^\circ$, which is also the intuitive meaning of linear independence.

# Takeaways (Cont'd)

- The Gram–Schmidt process is very useful in constructing an orthonormal basis.
- 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