Slide 21-Orthogonality III

MAT2040 Linear Algebra

Inner Product Spaces

In previous lectures, we have discussed the orthogonality of vectors and subspaces in Euclidean vector space $\mathbb{R}^n$. In this lecture, we will discuss the orthogonality of vectors and subspaces in general vector space. In fact, we will need to discuss these concepts in the general inner product space.

Inner Product Spaces

Definition 21.1 (Inner Product Space over Real Number Field)
Let V be a vector space, an **inner product** is an operation on V which assigns a real number $\langle \mathbf{x}, \mathbf{y} \rangle$ for each pair of vectors $\mathbf{x}, \mathbf{y} \in V$. The operation $\langle \cdot, \cdot \rangle$ satisfies:

(1) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ with equality if and only if $\mathbf{x} = \mathbf{0}$.

(2) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle, \ \forall \ \mathbf{x}, \ \mathbf{y} \in V$.

(3) $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle, \ \forall \ \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \text{ and } \alpha, \ \beta \in \mathbb{R}$.

If the vector space V has an inner product operation on V, then V is called the inner product space.

Examples

1. The standard inner product defined on the vector space $\mathbb{R}^n$
The standard inner product for $\mathbb{R}^n$ is the scalar product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T \mathbf{x}$.
(Without further notice, in this course, when discussing the Euclidean vector space $\mathbb{R}^n$, it is always associated with this standard inner product.)

2. Inner product defined on $\mathbb{R}^{m\times n}$ (**Frobenius inner product**)
Given $A, B \in \mathbb{R}^{m\times n}$, we can define an inner product as

$$\langle A, B\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}$$

Three conditions:
(1) $\langle A, A\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 \geq 0$, and the equality valid only when $a_{ij} = 0, i = 1, \cdots m, j = 1, \cdots, n$.
(2)

$$\langle A, B\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}a_{ij} = \langle B, A\rangle$$

(3)

$$
\begin{aligned}
\langle \alpha A + \beta B, C \rangle =& \sum_{i=1}^{m} \sum_{j=1}^{n} (\alpha a_{ij} + \beta b_{ij}) c_{ij} \\
=& \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} c_{ij} + \beta \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} c_{ij} = \alpha \langle A, C \rangle + \beta \langle B, C \rangle
\end{aligned}
$$

3. The vector space $C[a, b]$. For $f, g \in C[a, b]$, the inner product on $C[a, b]$ is defined by

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

Three conditions:

(1) $\langle f, f \rangle = \int_a^b f^2(x)dx \geq 0$. If $\langle f, f \rangle = \int_a^b f^2(x)dx = 0$, then we can show that $f(x) \equiv 0$. Otherwise if there exists a point x_0 s.t. $f(x_0) \neq 0$, say $f(x_0) > 0$, then there exists a interval $(x_0 - \delta, x_0 + \delta)$ containing the point x_0, s.t. $f(x) > 0$ when $x \in (x_0 - \delta, x_0 + \delta)$. Thus $0 < \int_{x_0-\delta}^{x_0+\delta} f^2(x)dx < \int_a^b f^2(x)dx = 0$. This is a contradiction.

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

(3) $\langle \alpha f + \beta g, h \rangle = \int_a^b (\alpha f(x) + \beta g(x))h(x)dx =$
$\alpha \int_a^b f(x)h(x)dx + \beta \int_a^b g(x)h(x)dx = \alpha \langle f, h \rangle + \beta \langle g, h \rangle$

Definition 21.2 (Length of the vector in inner product space) Let V be an inner product space, the **length** of $\mathbf{v}$ is defined as $\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Example:

For $\forall\ \mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$, which is the Euclidean length.

For $\forall\ f(x) \in C[a, b]$, $\|f\| = (\int_a^b f^2(x)dx)^{\frac{1}{2}}$.

For $\forall\ A \in \mathbb{R}^{m\times n}$, $\|A\| = (\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2)^{\frac{1}{2}}$.

Definition 21.3 (**Orthogonal in the Inner Product Space**) Two vectors $\mathbf{x}$ and $\mathbf{y}$ in the inner product space V is said to be orthogonal if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. (Generalization of orthogonality in $\mathbb{R}^n$).

Example: For $C[-1,1]$, $f(x) = 1, g(x) = x$, then $< f, g >= \int_{-1}^{1} f(x)g(x)dx = \int_{-1}^{1} xdx = 0$, f and g are orthogonal.

Example: For $\mathbb{R}^{2\times 2}$, $A = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 \\ -3 & -2 \end{bmatrix}$

$$< A, B >= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} = 4 * 1 + 3 * 3 + 3 * (-3) + 2 * (-2) = 0.$$

A and B are orthogonal in the sense of Frobenius inner product.

Theorem 21.4 **(Pythagorean's Law for inner product space)** If $\mathbf{u}, \mathbf{v}$ are two orthogonal vectors in the inner product space V, then

$$\| \mathbf{u} + \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2, \quad \| \mathbf{u} - \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2$$

Proof. $\| \mathbf{u} \pm \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 \pm 2\langle \mathbf{u}, \mathbf{v} \rangle$ and $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ will give the result.

Theorem 21.5 (**Cauchy-Schwartz Inequality**) If $\mathbf{u}$ and $\mathbf{v}$ are any two vectors in the inner product space V, then

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

Proof. See the appendix.

Normed linear vector spaces

Remark: in fact, the length of vector in inner product defines a norm. And the word norm in mathematical has its own meaning, independent of inner product space. The following is the definition for normed linear space.

Definition 21.6 (**Normed Vector Space**) A vector space V is said to be a normed linear space if, each vector $\mathbf{v} \in V$ is associated with a real number $\| \mathbf{v} \| \in \mathbb{R}$, called the **norm** of $\mathbf{v}$, satisfying:
(I) $\| \mathbf{v} \| \geq 0$ with equality if and only if $\mathbf{v} = 0$.
(II) $\| \alpha \mathbf{v} \| = |\alpha| \| \mathbf{v} \|$ for any scalar α.
(III) $\| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \|$ for any $\mathbf{u}, \mathbf{v} \in V$ (triangle inequality).

Theorem 21.7 (**Norm on the Inner Product Space**) For the inner product space V, for any $\mathbf{v} \in V$, the length $\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ defines a norm on V.

Proof. The condition (I) and (II) can be readily seen. For condition (III), by using the Cauchy-Schwartz inequality, one has

$$\begin{aligned} \| \mathbf{u} + \mathbf{v} \|^2 &= \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle \\ &\leq \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 + 2 \| \mathbf{u} \| \| \mathbf{v} \| \\ &\leq (\| \mathbf{u} \| + \| \mathbf{v} \|)^2 \end{aligned}$$

Definition 21.8 (Orthogonal Set) Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be nonzero vectors in an inner product space V. If $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ when $i \neq j$, then $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is said to be an **orthogonal set** of vectors.

Example: For $C[-1,1]$, $f(x) = 1, g(x) = x$, then $< f, g >= \int_{-1}^{1} f(x)g(x)dx = \int_{-1}^{1} xdx = 0$, f, g are orthogonal set in the inner product space $C[-1,1]$.

Definition 21.9 (**Orthonormal Set**) An **orthonormal set** of vectors is an orthogonal set of **unit** vectors, where the **unit** vector means the norm of the vector is 1.

Example: For $C[-1,1]$, $f(x) = \frac{1}{\sqrt{2}}, g(x) = \frac{x}{\sqrt{\frac{2}{3}}}$, then

$< f, g >= \int_{-1}^{1} f(x)g(x)dx = \int_{-1}^{1} \frac{x}{\frac{2}{\sqrt{3}}} dx = 0,$

$\|f\| = (\int_{-1}^{1} \frac{1}{2} dx)^{\frac{1}{2}} = 1, \quad \|g\| = (\int_{-1}^{1} \frac{x^2}{\frac{2}{3}} dx)^{\frac{1}{2}} = 1.$ Thus,

$f(x) = \frac{1}{\sqrt{2}}, g(x) = \frac{x}{\sqrt{\frac{2}{3}}}$ are orthonormal set in the inner product space $C[-1,1]$.

Theorem 21.10 (**Orthogonal set are linearly independent**) Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be the set of orthogonal vectors in an inner product space V, then $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is linearly independent.

Proof. Suppose that the following linear combination is zero:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}, \quad c_1, c_2, \cdots, c_n \text{ are scalars.}$$

For $1 \leq i \leq n$, taking the inner product with $\mathbf{v}_i$ on both sides of the equation yields

$$c_i \parallel \mathbf{v}_i \parallel^2 = 0$$

Then $c_i = 0$ $(0 \leq i \leq n)$ since $\parallel \mathbf{v}_i \parallel > 0$ and $\mathbf{v}_i \neq \mathbf{0}$.

Example: It will be an excise to check that $1, x, x^2 - \frac{1}{3}$ are orthogonal set in the inner product space $C[-1, 1]$. Thus, $1, x, x^2 - \frac{1}{3}$ are linearly independent.

Remark 1: Orthogonal set is linearly independent, but linearly independent set may not the orthogonal set. Eg. $\{[1,0]^T, [1,1]^T\}$ is linearly independent but not orthogonal.

Remark 2: The set $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is orthonormal if and only if

$$\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = \delta_{ij}$$

where

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Remark 3: Given the orthogonal set $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$, we can use **the method of normalization** to form the orthonormal set as

$$\left\{ \frac{\mathbf{u}_1}{\| \mathbf{u}_1 \|}, \frac{\mathbf{u}_2}{\| \mathbf{u}_2 \|}, \ldots, \frac{\mathbf{u}_n}{\| \mathbf{u}_n \|} \right\}$$

Example 21.11 Let

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix}$$

then $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, the orthonormal set is

$$\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$$

Definition 21.12 **(Orthonormal Basis)** $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ is the orthonormal basis for the inner product vector space V, if the following conditions are satisfied:

1. $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ is the orthonormal set.
2. $V = \mathrm{Span}\,(\mathbf{u}_1, \cdots, \mathbf{u}_m)$.

Example For $\mathbb{R}^3$, the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ will be the orthonormal basis.

Moreover, from above example 21.10,

$$\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$$

will also be an orthonormal basis for $\mathbb{R}^3$.

Theorem 21.13 **(Coordinate w.r.t orthonormal basis)** Let $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ be the orthonormal basis for the inner product vector space V, and for any $\mathbf{v} \in V$, $\mathbf{v}$ can be decomposed as

$$\mathbf{v} = \sum_{i=1}^{m} < \mathbf{v}, \mathbf{u_i} > \mathbf{u}_i$$

Proof. For any $\mathbf{v} \in V$, it can be written as a linear combination of the orthonormal basis as follows:

$$\mathbf{v} = c_1\mathbf{u}_1 + \cdots + c_m\mathbf{u}_m$$

Taking the inner product with $\mathbf{u}_j$ on both sides of the above equation, one has:

$$< \mathbf{v}, \mathbf{u}_i >= c_i \parallel \mathbf{u}_i \parallel^2 = c_i$$

since $\mathbf{u}_i(\ i = 1, \cdots, m)$ are the unit vectors.

Example 21.14 For $\mathbb{R}^3$,

$$\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$$

will also be an orthonormal basis.
For any $\mathbf{x} = [x, y, z]^T \in \mathbb{R}^3$, one has

$$\begin{aligned} \mathbf{x} &= < \mathbf{x}, \mathbf{v}_1 > \mathbf{v}_1 + < \mathbf{x}, \mathbf{v}_2 > \mathbf{v}_2 + < \mathbf{x}, \mathbf{v}_3 > \mathbf{v}_3 \\ &= \frac{x + y + z}{\sqrt{3}} \mathbf{v}_1 + \frac{2x + y - 3z}{\sqrt{14}} \mathbf{v}_2 + \frac{4x - 5y + z}{\sqrt{42}} \mathbf{v}_3 \end{aligned}$$

Remark: If $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ is an orthogonal set and is the basis for the inner product vector space V, then for any $\mathbf{v} \in V$, one has:

$$\mathbf{v} = \sum_{i=1}^{m} \frac{< \mathbf{v}, \mathbf{u}_i >}{\| \mathbf{u}_i \|^2} \mathbf{u}_i$$

Orthogonal Matrix

One important matrix is $n \times n$ matrix whose columns are an orthonormal set in $\mathbb{R}^n$.

Definition 21.15 (**Orthogonal Matrix**)
Let $Q \in \mathbb{R}^{n\times n}$, Q is said to be the orthogonal matrix if the column vectors of Q is an orthonormal set in $\mathbb{R}^n$.

Remark: In fact, the column vectors of orthogonal matrix Q are orthonormal basis for $\mathbb{R}^n$ since the column vectors of Q are linearly independent set in $\mathbb{R}^n$ and the number of columns is n.

Orthogonal Matrix

Theorem 21.16 (**Equivalent Condition for Orthogonal Matrix**) An $n \times n$ matrix Q is orthogonal matrix if and only if $Q^{-1} = Q^T$.

Recall: For square matrices $A, B \in \mathbb{R}^{n\times n}$, $AB = I_n$ implies $BA = I_n$.

Proof. Since Q is a square matrix, $Q^{-1} = Q^T$ is equivalent to $Q^T Q = I_n$. Thus, only need to show that Q is orthogonal matrix if and only if $Q^T Q = I_n$. Let $Q = [\mathbf{q}_1, \cdots, \mathbf{q}_n]$, then $Q^T = \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix}$.

$$
\begin{aligned}
& Q = [\mathbf{q}_1, \cdots, \mathbf{q}_n] \text{ is an orthogonal matrix} \\
\Leftrightarrow\ & \{\mathbf{q}_1, \cdots, \mathbf{q}_n\} \text{ is an orthonormal set in } \mathbb{R}^n \\
\Leftrightarrow\ & \{\mathbf{q}_1, \cdots, \mathbf{q}_n\} \text{ is an orthonormal basis for } \mathbb{R}^n \\
\Leftrightarrow\ & (\mathbf{q}_i^T \mathbf{q}_j)_{n\times n} = Q^T Q = (\delta_{ij})_{n\times n} = I_n \ .
\end{aligned}
$$

Orthogonal Matrix

Example 21.17 For any fixed θ, the matrix $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, and $Q^{-1} = Q^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

Orthogonal Matrix

Property 21.18 (Properties for orthogonal matrix) If Q is an $n \times n$ orthogonal matrix, then

(a) the column vectors of Q form an orthonormal basis for $\mathbb{R}^n$.

(b) $Q^{-1} = Q^T$

(c) $Q^T Q = I_n$

(d) $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$,

$$\langle Q\mathbf{x}, Q\mathbf{y} \rangle = (Q\mathbf{y})^T Q\mathbf{x} = \mathbf{y}^T Q^T Q\mathbf{x} = \mathbf{y}^T \mathbf{x} = \langle \mathbf{x}, \mathbf{y} \rangle$$

(e) $\| Q\mathbf{x} \| = \| \mathbf{x} \|$

Appendix: The proof of Cauchy-Schwartz Inequality for inner product space

Theorem 21.5 (**Cauchy-Schwartz Inequality**) If $\mathbf{u}$ and $\mathbf{v}$ are any two vectors in the inner product space V, then

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

Proof. If $\mathbf{v} = \mathbf{0}$, the inequality becomes equality. If $\mathbf{v} \neq \mathbf{0}$, then
$\langle \mathbf{u} - k\mathbf{v}, \mathbf{u} - k\mathbf{v} \rangle \geq 0$ for any $k \in \mathbb{R}$.
$\langle \mathbf{u} - k\mathbf{v}, \mathbf{u} - k\mathbf{v} \rangle = \| \mathbf{u} \|^2 - 2k\langle \mathbf{u}, \mathbf{v} \rangle + k^2 \| \mathbf{v} \|^2 \geq 0$ for any $k \in \mathbb{R}$. Thus

$$\triangle = 4|\langle \mathbf{u}, \mathbf{v} \rangle|^2 - 4 \| \mathbf{u} \|^2 \| \mathbf{v} \|^2 \leq 0$$

This gives the result.