Slide 19-Orthogonality I

MAT2040 Linear Algebra

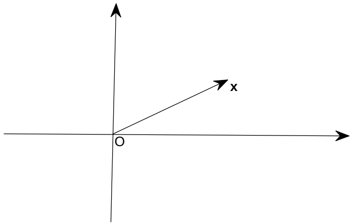
SSE, CUHK(SZ)

Scalar Product and Orthogonality in $\mathbb{R}^n$

Let $\mathbf{x}$ and $\mathbf{y}$ are two vectors in $\mathbb{R}^n$, then the product $\mathbf{x}^\mathbf{T}\mathbf{y}$ is called the **scalar product** since $\mathbf{x}^\mathbf{T}\mathbf{y}$ is a real number. ($\mathbf{x}$ and $\mathbf{y}$ can be regarded as $n \times 1$ matrices, $\mathbf{x}^\mathbf{T}\mathbf{y}$ will be a 1×1 matrix which is a real number). Let $\mathbf{x} = [x_1, \cdots, x_n]^T$, $\mathbf{y} = [y_1, \cdots, y_n]^T$, then

$$\mathbf{x}^\mathbf{T}\mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Given any nonzero vector $\mathbf{x} \in \mathbb{R}^n$, geometrically, we can consider it as a vector with starting point at the origin in n-dimensional space.



Definition 19.1 (Euclidean Length) Let $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T \in \mathbb{R}^n$, the Euclidean length of $\mathbf{x}$ is given by

$$\| \mathbf{x} \| = \sqrt{\mathbf{x}^T \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

It can be regarded as the length of the vector $\mathbf{x}$.

Example 19.2
Let $\mathbf{x} = [3, -2, 1]^T \in \mathbb{R}^3$, the Euclidean length of $\mathbf{x}$ is given by

$$\| \mathbf{x} \| = \sqrt{3^2 + (-2)^2 + 1^2} = \sqrt{14}$$

Definition 19.3 (Distance) Let $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T, \mathbf{y} = [y_1, y_2, \cdots, y_n]^T \in \mathbb{R}^n$, then $\mathbf{x} - \mathbf{y} = [x_1 - y_1, \cdots, x_n - y_n]^T$, the distance between two vectors is given by

$$\| \mathbf{x} - \mathbf{y} \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

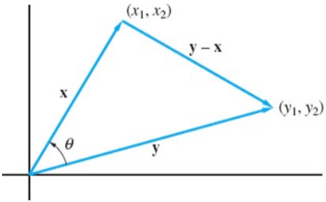


Figure: Illustration for 2D case

Example 19.4 Let $\mathbf{x} = [1, 2, -2, 3]^T, \mathbf{y} = [2, -1, 3, 4]^T \in \mathbb{R}^4$, then

$$\| \mathbf{x} - \mathbf{y} \| = \sqrt{(1-2)^2 + (2-(-1))^2 + (-2-3)^2 + (3-4)^2} = 6$$

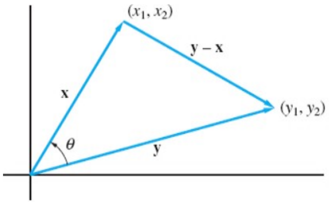


Figure: Illustration for 2D case

Lemma 19.5 (**Cauchy-Schwartz Inequality**)

$$|\mathbf{x}^T\mathbf{y}| \leq \| \mathbf{x} \| \| \mathbf{y} \|$$

The inequality becomes equality only when one vector is zero or $\mathbf{x}$ and $\mathbf{y}$ are in the same direction (one is a multiple of another).
See the appendix for the proof.

Theorem 19.6 (**Scalar Product in terms of Vector Length**) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, suppose θ is the angle between two nonzero vectors, then

$$\mathbf{x}^T\mathbf{y} = \| \mathbf{x} \| \| \mathbf{y} \| \cos\theta, \quad 0 \leq \theta \leq \pi.$$

Proof. By the cosine law, one has

$$\| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 - \| \mathbf{x} - \mathbf{y} \|^2 = 2 \| \mathbf{x} \| \| \mathbf{y} \| \cos\theta.$$

In addition,

$$\begin{aligned} &\| \mathbf{x} - \mathbf{y} \|^2 = (\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y}) \\ =&\mathbf{x}^T\mathbf{x} + \mathbf{y}^T\mathbf{y} - \mathbf{x}^T\mathbf{y} - \mathbf{y}^T\mathbf{x} \\ =& \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 - \mathbf{x}^T\mathbf{y} - \mathbf{y}^T\mathbf{x} \end{aligned}$$

And $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$.
Thus

$$\mathbf{x}^T\mathbf{y} = \| \mathbf{x} \| \| \mathbf{y} \| \cos\theta$$

Since $\mathbf{x}, \mathbf{y}$ are nonzero vectors, one has

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\| \mathbf{x} \| \| \mathbf{y} \|} = \mathbf{u}^T\mathbf{v}$$

where $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ and $\mathbf{v} = \frac{\mathbf{y}}{\|\mathbf{y}\|}$ are the unit vectors in the x, y directions.

Definition 19.7 (**Orthogonal Vectors in $\mathbb{R}^n$**) Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{x}^T\mathbf{y} = 0$. Denote $\mathbf{x} \perp \mathbf{y}$.

Recall:

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\| \mathbf{x} \| \| \mathbf{y} \|} = \mathbf{u}^T\mathbf{v}$$

Thus

$\mathbf{x}$ and $\mathbf{y}$ are orthogonal $\Leftrightarrow \mathbf{x}^T\mathbf{y} = 0 \Leftrightarrow \cos\theta = 0 \Leftrightarrow \theta$ is the right angle.

Example 19.8

(1) Vectors $[3, 2]^T$ and $[-4, 6]^T$ are orthogonal in $\mathbb{R}^2$.

(2) Vectors $[2, -3, 1]^T$ and $[1, 1, 1]^T$ are orthogonal in $\mathbb{R}^3$.

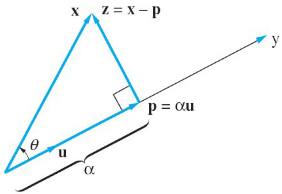
Theorem 19.9 (**Pythagorean's Law**)
Let $\mathbf{x}, \mathbf{y}$ be two vectors in $\mathbb{R}^n$, if they are orthogonal, then

$$\| \mathbf{x} + \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2$$

Since

$$\| \mathbf{x} + \mathbf{y} \|^2 = (\mathbf{x} + \mathbf{y})^T (\mathbf{x} + \mathbf{y}) = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 + \mathbf{x}^T \mathbf{y} + \mathbf{y}^T \mathbf{x}$$

and $\mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} = 0$ because of the orthogonality.



Definition 19.10 (**Scalar and vector projection**) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, find $\mathbf{p}$ in the direction of $\mathbf{y}$ and $\mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{y}$. $\mathbf{u} = \frac{\mathbf{y}}{\|\mathbf{y}\|}$ is the unit vector in $\mathbf{y}$ direction. Suppose $\mathbf{p} = \alpha\mathbf{u}$, then $\mathbf{x} - \alpha\mathbf{u}$ is orthogonal to $\mathbf{u}$, i.e. $(\mathbf{x} - \alpha\mathbf{u})^T\mathbf{u} = 0 \Rightarrow \alpha = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$. Geometrically, $\alpha = \| \mathbf{x} \| \cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$.

(1) α is the **scalar projection** of $\mathbf{x}$ onto $\mathbf{y}$.

(2) $\mathbf{p} = \alpha\mathbf{u} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}$ is the **vector projection** of $\mathbf{x}$ onto $\mathbf{y}$.

Orthogonal Subspaces in $\mathbb{R}^n$

Definition 19.11 (Orthogonal Subspaces in $\mathbb{R}^n$) Two subspaces X and Y of $\mathbb{R}^n$ are said to be orthogonal if

$$\mathbf{x}^T\mathbf{y} = 0, \ \forall\, \mathbf{x} \in X, \ \mathbf{y} \in Y.$$

Denoted by $X \perp Y$.

Corollary If X and Y are orthogonal subspaces of $\mathbb{R}^n$, then $X \cap Y = \{\mathbf{0}\}$

Proof. Suppose that $\mathbf{x} \in X \cap Y$, then $\mathbf{x}^T\mathbf{x} = 0 = \| \mathbf{x} \|^2$, thus $\mathbf{x} = \mathbf{0}$

Example 19.12

(1) Let

$$X = \textbf{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right), \quad Y = \textbf{Span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right)$$

then $X \perp Y$.

(2) Let

$$X = \textbf{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right), \quad Y = \textbf{Span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$$

then X is the *xy* plane while Y is *yz* plane. Geometrically, these two planes are perpendicular with each other but X and Y are not orthogonal since $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in X \cap Y$ ($X \cap Y \neq \{\mathbf{0}\}$).

Definition 19.13 (**Orthogonal Complement**) Let Y be a subspace of $\mathbb{R}^n$, vectors in $\mathbb{R}^n$ that are orthogonal to every vector in Y is said to be the **orthogonal complement** of Y, denoted by $Y^\perp$. Thus

$$Y^\perp = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x}^T \mathbf{y} = 0, \forall\ \mathbf{y} \in Y\}$$

Example 19.14 The subspace $X = \mathbf{Span}(\mathbf{e}_1)$ and $Y = \mathbf{Span}(\mathbf{e}_2)$ of $\mathbb{R}^3$ are orthogonal but they are not orthogonal complements. Indeed,

$$X^\perp = \mathbf{Span}(\mathbf{e}_2, \mathbf{e}_3), \quad Y^\perp = \mathbf{Span}(\mathbf{e}_1, \mathbf{e}_3)$$

Proposition 19.15 (Proposition of Orthogonal Complements)
If Y is a subspace of $\mathbb{R}^n$, then $Y^\perp$ is also a subspace of $\mathbb{R}^n$.

Proof.
Obviously $\mathbf{0} \in Y^\perp$. Now suppose that $\mathbf{y}, \mathbf{z} \in Y^\perp$ and $\alpha_1, \alpha_2 \in \mathbb{R}$, then for any $\mathbf{x} \in Y$, one has

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}^T\mathbf{z} = 0$$

Thus

$$\mathbf{x}^T(\alpha_1\mathbf{y} + \alpha_2\mathbf{z}) = \alpha_1\mathbf{x}^T\mathbf{y} + \alpha_2\mathbf{x}^T\mathbf{z} = 0$$

and

$$\alpha_1\mathbf{y} + \alpha_2\mathbf{z} \in Y^\perp$$

Therefore, $Y^\perp$ is a subspace of $\mathbb{R}^n$.

Example 19.16 Given *yz* plane in $\mathbb{R}^3$

$$Y = \textbf{Span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$$

find $Y^\perp$.

The elements in Y can be written as

$$\begin{bmatrix} 0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$$

For any element

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

in $Y^\perp$, it satisfies:

$$0x + \alpha_1 y + \alpha_2 z = 0, \forall\ \alpha_1, \alpha_2 \in \mathbb{R}$$

Thus, $y = z = 0$ and there is no restriction for x. Thus

$$Y^{\perp} = \textbf{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$$

Theorem 19.17 (**Fundamental Subspaces Theorem**)

Let $A \in \mathbb{R}^{m\times n} = [\mathbf{a}_1, \cdots, \mathbf{a}_n] = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$, $Col(A) = \mathrm{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ is the column space of A, $Row(A) = Col(A^T) = \mathrm{Span}\{\vec{\mathbf{a}}_1^T, \vec{\mathbf{a}}_2^T, \cdots, \vec{\mathbf{a}}_m^T\}$ is the row space of A, then

(1)Null(A)$=Col(A^T)^{\perp}=Row(A)^{\perp}$

(2)Null(A^T)$=Col(A)^{\perp}=Row(A^T)^{\perp}$

Proof. Let $A = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$, then $A^T = [\vec{\mathbf{a}}_1^T, \vec{\mathbf{a}}_2^T, \cdots, \vec{\mathbf{a}}_m^T]$, and

$$
\begin{aligned}
\mathrm{Null}(A) =& \{\mathbf{x} | A\mathbf{x} = \mathbf{0}\} \\
=& \{\mathbf{x} | \vec{\mathbf{a}}_i \mathbf{x} = 0, \forall\ i = 1, \cdots, m\} \\
=& \{\mathbf{x} | \left(\sum_{i=1}^{m} \alpha_i \vec{\mathbf{a}}_i\right) \mathbf{x} = 0, \forall\ \alpha_i \in \mathbb{R},\ \ i = 1, 2, \cdots, m\} \\
=& \{\mathbf{x} | \left(\sum_{i=1}^{m} \alpha_i (\vec{\mathbf{a}}_i)^T\right)^T \mathbf{x} = 0, \forall\ \alpha_i \in \mathbb{R},\ \ i = 1, 2, \cdots, m\} \\
=& \{\mathbf{x} | \mathbf{y}^T \mathbf{x} = 0, \forall \mathbf{y} \in Col(A^T)\} \\
=& Col(A^T)^{\perp} \\
=& Row(A)^{\perp}
\end{aligned}
$$

since $\mathbf{y} = \sum_{i=1}^{m} \alpha_i (\vec{\mathbf{a}}_i)^T$ is the arbitary element in $Col(A^T)$.
In addition,

$$
\mathrm{Null}(A^T) = Col(A)^{\perp} = Row(A^T)^{\perp}
$$

Example 19.18 Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

Then

$$Null(A) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}, \quad Col(A^T) = Row(A) = \mathbb{R}^2$$

$$Null(A^T) = \textbf{Span}\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right), \quad Col(A) = \textbf{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right)$$

Thus

$$Null(A)^{\perp} = Col(A^T) = Row(A)$$

$$Null(A^T)^{\perp} = Col(A) = Row(A^T)$$

Theorem 19.19 If S is a subspace of $\mathbb{R}^n$, then

$$\dim S + \dim S^{\perp} = n.$$

Furthermore, if $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is a basis for S and $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $S^{\perp}$, then $\{\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $\mathbb{R}^n$.

Proof. Skipped. See Steven's book P220 or see the appendix.

Remark. If S is a subspace of $\mathbb{R}^n$, it can be shown that $(S^{\perp})^{\perp} = S$ (the proof is skipped, see Steven's book P221). S and $S^{\perp}$ are mutually orthogonal.

Appendix: The proof of Cauchy-Schwartz Inequality

Theorem 19.5 (**Cauchy-Schwartz Inequality**) If $\mathbf{x}$ and $\mathbf{y}$ are any two vectors in the $\mathbb{R}^n$, then

$$|\mathbf{x}^T\mathbf{y}| \leq \| \mathbf{x} \| \| \mathbf{y} \|$$

Proof. If $\mathbf{y} = \mathbf{0}$, the inequality becomes equality. If $\mathbf{y} \neq \mathbf{0}$, then
$\big(\mathbf{x} - k\mathbf{y}\big)^T(\mathbf{x} - k\mathbf{y}\big) \geq 0$ for any $k \in \mathbb{R}$.
$\big(\mathbf{x} - k\mathbf{y}\big)^T(\mathbf{x} - k\mathbf{y}\big) = \| \mathbf{x} \|^2 - 2k\mathbf{x}^T\mathbf{y} + k^2 \| \mathbf{y} \|^2 \geq 0$ for any $k \in \mathbb{R}$. Thus

$$\triangle = 4|\mathbf{x}^T\mathbf{y}|^2 - 4 \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 \leq 0$$

This gives the result.

Theorem 19.19

If S is a subspace of $\mathbb{R}^n$, then

$$\dim S + \dim S^{\perp} = n.$$

Furthermore, if $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is a basis for S and $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $S^{\perp}$, then $\{\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $\mathbb{R}^n$.

Proof.

(1) If $S = \{\mathbf{0}\}$, then $S^{\perp} = \mathbb{R}^n$, the statement is true.

(2) Assume that $S \neq \{\mathbf{0}\}$, then let $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ be a basis for S, let $A = \left[\mathbf{u}_1, \cdots, \mathbf{u}_r\right]$, then $S = Col(A)$, $rank(A) = rank(A^T) = r$ and

$$S^{\perp} = Col(A)^{\perp} = Null(A^T)$$

By the Rank-Nullity theorem, we have $rank(A^T) + \dim(\mathrm{Null}(A^T)) = n$, thus

$$\dim S + \dim S^{\perp} = n$$

Now suppose that the following linear combination is zero, i.e.,

$$\alpha_1 \mathbf{u}_1 + \cdots + \alpha_r \mathbf{u}_r + \alpha_{r+1} \mathbf{u}_{r+1} + \cdots + \alpha_n \mathbf{u}_n = \mathbf{0}$$

then

$$\alpha_1 \mathbf{u}_1 + \cdots + \alpha_r \mathbf{u}_r = -\alpha_{r+1} \mathbf{u}_{r+1} - \cdots - \alpha_n \mathbf{u}_n$$

The LHS is a vector in S and the RHS is a vector in $S^\perp$, since $S \cap S^\perp = \{\mathbf{0}\}$, then

$$\alpha_1 \mathbf{u}_1 + \cdots + \alpha_r \mathbf{u}_r = 0 = -\alpha_{r+1} \mathbf{u}_{r+1} - \cdots - \alpha_n \mathbf{u}_n$$

Since $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is a basis for S and $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $S^\perp$, thus

$$\alpha_1 = \cdots = \alpha_r = \alpha_{r+1} = \cdots = \alpha_n = 0$$

Thus, $\{\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $\mathbb{R}^n$.