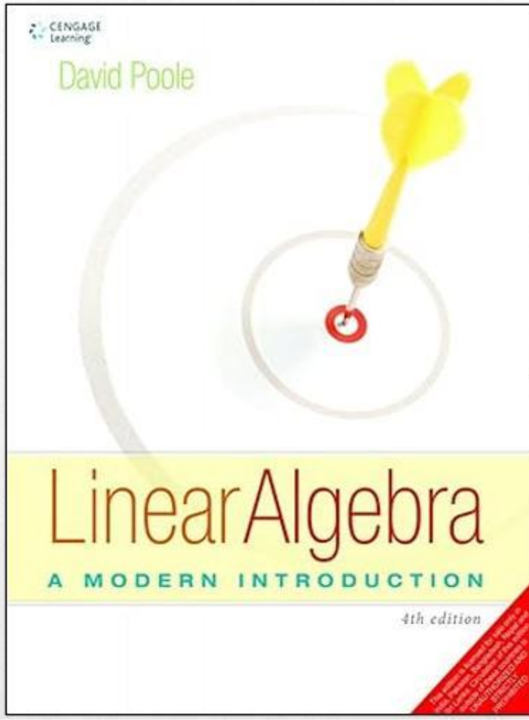

114-2

線性代數 (二)

國立政治大學
統計學系
吳漢銘

5.1 Orthogonality in $\mathbb{R}^n$

https://hmwu.idv.tw

generalize the notion of orthogonality of vectors in $\mathbb{R}^n$ from two vectors to sets of vectors.

Orthogonal and Orthonormal Sets of Vectors

Definition A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in $\mathbb{R}^n$ is called an ***orthogonal set*** if all pairs of distinct vectors in the set are orthogonal—that is, if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{whenever} \quad i \neq j \quad \text{for } i, j = 1, 2, \ldots, k$$

The standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ is an orthogonal set

Example 5.1

Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in $\mathbb{R}^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

perpendicular

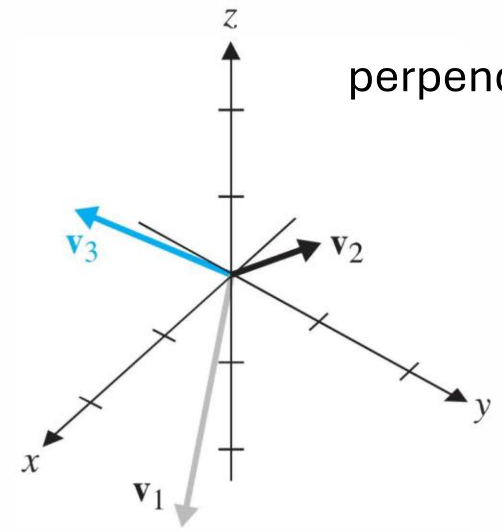

Figure 5.4
An orthogonal set of vectors

Theorem 5.1 If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then these vectors are linearly independent.

Proof If $c_1, \ldots, c_k$ are scalars such that $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, then

$$(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i = 0 \cdot \mathbf{v}_i = 0$$

$$c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) = 0 \qquad (1)$$

Definition An ***orthogonal basis*** for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthogonal set.

Example 5.2 The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

from Example 5.1 are orthogonal and, hence, linearly independent.

Since any three linearly independent vectors in $\mathbb{R}^3$ form a basis for $\mathbb{R}^3$, by the Fundamental Theorem of Invertible Matrices, it follows that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for $\mathbb{R}^3$.

Example 5.3 Find an orthogonal basis for the subspace W of $\mathbb{R}^3$ given by

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x - y + 2z = 0 \right\}$$

Solution Section 5.3 gives a general procedure for problems of this sort.

Theorem 5.2 Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$ and let $\mathbf{w}$ be any vector in W. Then the unique scalars $c_1, \ldots, c_k$ such that

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

are given by

$$c_i = \frac{\mathbf{w} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \quad \text{for } i = 1, \ldots, k$$

Proof Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a basis for W, we know that there are unique scalars $c_1, \ldots, c_k$ such that $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ (from Theorem 3.29).

Example 5.4

Find the coordinates of $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ with respect to the orthogonal basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of Examples 5.1 and 5.2.

With the notation introduced in Section 3.5, we can also write the above equation as

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} \frac{1}{6} \\ \frac{5}{2} \\ \frac{2}{3} \end{bmatrix}$$

Definition A set of vectors in $\mathbb{R}^n$ is an ***orthonormal set*** if it is an orthogonal set of unit vectors. An ***orthonormal basis*** for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthonormal set.

Remark If $S = \{\mathbf{q}_1, \dots, \mathbf{q}_k\}$ is an orthonormal set of vectors, then $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ for $i \neq j$ and $\|\mathbf{q}_i\| = 1$. The fact that each $\mathbf{q}_i$ is a unit vector is equivalent to $\mathbf{q}_i \cdot \mathbf{q}_i = 1$. It follows that we can summarize the statement that S is orthonormal as

$$\mathbf{q}_i \cdot \mathbf{q}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Example 5.5

Show that $S = \{\mathbf{q}_1, \mathbf{q}_2\}$ is an orthonormal set in $\mathbb{R}^3$ if

$$\mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad \text{and} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

Example 5.6

Construct an orthonormal basis for $\mathbb{R}^3$ from the vectors in Example 5.1.

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in $\mathbb{R}^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Theorem 5.3 Let $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}$ be an orthonormal basis for a subspace W of $\mathbb{R}^n$ and let $\mathbf{w}$ be any vector in W. Then

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{w} \cdot \mathbf{q}_2)\mathbf{q}_2 + \cdots + (\mathbf{w} \cdot \mathbf{q}_k)\mathbf{q}_k$$

and this representation is unique.

Proof Apply Theorem 5.2 and use the fact that $\mathbf{q}_i \cdot \mathbf{q}_i = 1$ for $i = 1, \ldots, k$.

Theorem 5.2 Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$ and let $\mathbf{w}$ be any vector in W. Then the unique scalars $c_1, \ldots, c_k$ such that

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

are given by

$$c_i = \frac{\mathbf{w} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \quad \text{for } i = 1, \ldots, k$$

Theorem 5.4 The columns of an $m \times n$ matrix Q form an orthonormal set if and only if $Q^TQ = I_n$.

Definition An $n \times n$ matrix Q whose columns form an orthonormal set is called an ***orthogonal matrix.***

Theorem 5.5 A square matrix Q is orthogonal if and only if $Q^{-1} = Q^T$.

Proof By Theorem 5.4, Q is orthogonal if and only if $Q^TQ = I$. This is true if and only if Q is invertible and $Q^{-1} = Q^T$, by Theorem 3.13.

Example 5.7 Show that the following matrices are orthogonal and find their inverses:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Example 5.7

Show that the following matrices are orthogonal and find their inverses:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Therefore, B is orthogonal, by Theorem 5.5, and

$$B^{-1} = B^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

The word *isometry* literally means "length preserving," since it is derived from the Greek roots *isos* ("equal") and *metron* ("measure").

Matrix B is the matrix of a rotation through the angle θ in $\mathbb{R}^2$.
Any rotation has the property that it is a *length-preserving* transformation (known as an ***isometry*** in geometry).

every orthogonal matrix transformation is an isometry.

Theorem 5.6 Let Q be an $n \times n$ matrix. The following statements are equivalent:

a. Q is orthogonal.
b. $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ for every $\mathbf{x}$ in $\mathbb{R}^n$.
c. $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for every $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

Proof We will prove that (a) $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ (a). To do so, we will need to make use of the fact that if $\mathbf{x}$ and $\mathbf{y}$ are (column) vectors in $\mathbb{R}^n$, then $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$.

(b) => (c)

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

Now if $\mathbf{e}_i$ is the ith standard basis vector, then $\mathbf{q}_i = Q\mathbf{e}_i$. Consequently,

$$\mathbf{q}_i \cdot \mathbf{q}_j = Q\mathbf{e}_i \cdot Q\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Thus, the columns of Q form an orthonormal set, so Q is an orthogonal matrix.

Theorem 5.7 If Q is an orthogonal matrix, then its rows form an orthonormal set.

Proof From Theorem 5.5, we know that $Q^{-1} = Q^T$. Therefore,

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

so Q^T is an orthogonal matrix. Thus, the columns of Q^T—which are just the rows of Q—form an orthonormal set.

Theorem 5.8 Let Q be an orthogonal matrix.

a. Q^{-1} is orthogonal.
b. $\det Q = \pm 1$
c. If λ is an eigenvalue of Q, then $|\lambda| = 1$.
d. If Q_1 and Q_2 are orthogonal $n \times n$ matrices, then so is $Q_1 Q_2$.

Proof

(c) Let λ be an eigenvalue of Q with corresponding eigenvector $\mathbf{v}$. Then $Q\mathbf{v} = \lambda\mathbf{v}$, and, using Theorem 5.6(b), we have

$$\|\mathbf{v}\| = \|Q\mathbf{v}\| = \|\lambda\mathbf{v}\| = |\lambda|\,\|\mathbf{v}\|$$

Since $\|\mathbf{v}\| \neq 0$, this implies that $|\lambda| = 1$.

Exercises 5.1

2, 8, 9, 17, 28