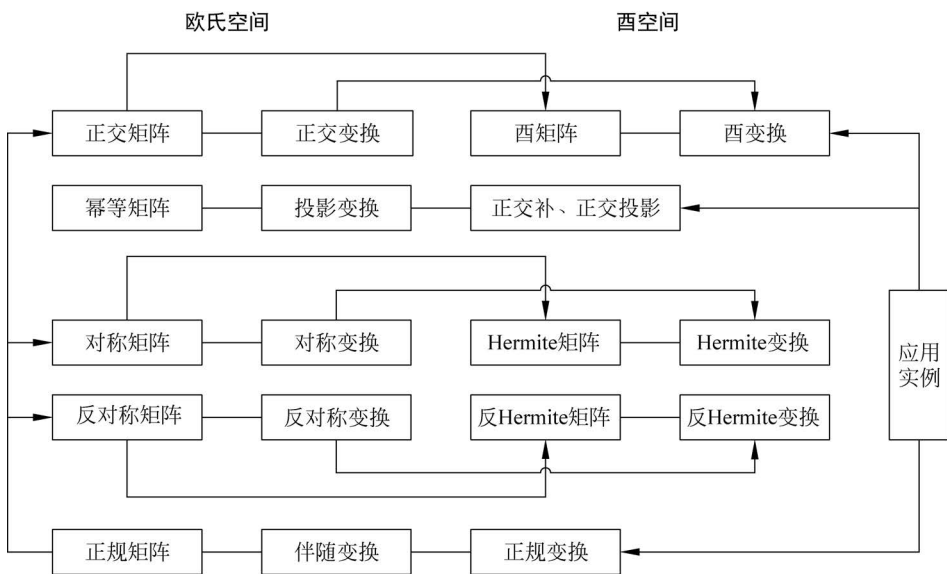


## 第3章

# 典型矩阵与变换

本章的知识网络框图：



在数学形式上,对向量进行线性变换就是用矩阵与之相乘. 本章研究几类典型线性变换及其矩阵的性质,包括正交矩阵与正交变换、酉矩阵与酉变换、幂等矩阵与投影变换、对称矩阵与对称变换、Hermite 矩阵与 Hermite 变换、正规矩阵与正规变换等,在此基础上给出应用实例.

### 3.1 正交矩阵与正交变换、酉矩阵与酉变换

#### 3.1.1 正交矩阵和酉矩阵

**定义 3.1.1** 设  $E$  是单位阵, 若  $n$  阶实矩阵  $A$  满足

$$A^T A = A A^T = E,$$

则称  $A$  是正交矩阵, 记为  $A \in E^{n \times n}$ .

若  $A, B \in E^{n \times n}$ , 则有如下性质:

- (1)  $A^{-1} = A^T \in E^{n \times n}$ .
- (2)  $\det A = \pm 1$ , 即正交矩阵的行列式是  $+1$  或  $-1$ .
- (3)  $AB, BA \in E^{n \times n}$ , 即正交矩阵的乘积仍为正交矩阵.
- (4) 正交矩阵的特征值是  $\pm 1$ , 特征向量互相正交.

**定义 3.1.2** 设复矩阵  $A = (a_{ij})_{m \times n} \in \mathbb{C}^{m \times n}$ ,  $\bar{A}$  是以  $A$  的元素的共轭复数为元素的矩阵, 即  $\bar{A} = (\overline{a_{ij}})_{m \times n}$ , 令  $A^H = (\bar{A})^T$ , 称  $A^H$  为  $A$  的复共轭转置矩阵.

不难证明复共轭转置具有如下性质:

- (1)  $(A+B)^H = A^H + B^H$ ;
- (2)  $(kA)^H = \bar{k}A^H$ ;
- (3)  $(AB)^H = B^H A^H$ ;
- (4)  $(A^H)^H = A$ ;
- (5)  $(A^H)^{-1} = (A^{-1})^H$ , 当  $A$  可逆时.

**定义 3.1.3** 设  $E$  是单位矩阵, 若  $n$  阶复矩阵  $A$  满足

$$A^H A = A A^H = E,$$

则称  $A$  是酉矩阵, 记为  $A \in U^{n \times n}$ .

酉是英文 Unitary 的音译, 酉矩阵也称为幺正矩阵.

若  $A, B \in U^{n \times n}$ , 则有如下性质:

- (1)  $A^{-1} = A^H \in U^{n \times n}$ , 即酉矩阵的逆为其复共轭转置矩阵;
- (2)  $|\det A| = 1$ , 即酉矩阵的行列式的模为 1;
- (3)  $A^T \in U^{n \times n}$ ;
- (4)  $AB, BA \in U^{n \times n}$ , 即酉矩阵的乘积仍为酉矩阵;

(5) 酉矩阵的特征值的模为 1, 特征向量互相正交.

**例 3.1.1** 设  $\alpha \in \mathbb{C}^n$ , 且  $\alpha^H \alpha = 1$ , 豪斯霍尔德变换的矩阵为

$$H = E - 2\alpha\alpha^H \in \mathbb{C}^{n \times n},$$

试证明  $H$  是酉矩阵.

**证:** 由复共轭转置的性质, 有

$$\begin{aligned} H^H H &= (E - 2\alpha\alpha^H)^H (E - 2\alpha\alpha^H) = (E - 2\alpha\alpha^H)(E - 2\alpha\alpha^H) \\ &= E - 4\alpha\alpha^H + 4\alpha\alpha^H\alpha\alpha^H = E, \end{aligned}$$

故  $H$  是酉矩阵. ■

前面介绍了酉矩阵的性质, 下述定理给出了如何构造酉矩阵.

**定理 3.1.1** 设  $A \in \mathbb{C}^{n \times n}$ , 则  $A$  是酉矩阵(正交矩阵)的充要条件是  $A$  的  $n$  个列(行)向量是标准正交向量组.

**证:** (1) 必要性 将  $A \in \mathbb{C}^{n \times n}$  按列分块, 设  $A = [\alpha_1, \alpha_2, \dots, \alpha_n]$ , 则

$$A^H = \begin{bmatrix} \alpha_1^H \\ \alpha_2^H \\ \vdots \\ \alpha_n^H \end{bmatrix},$$

若  $A \in U^{n \times n}$ , 则有  $A^H A = E$ , 于是

$$\begin{bmatrix} \alpha_1^H \\ \alpha_2^H \\ \vdots \\ \alpha_n^H \end{bmatrix} [\alpha_1, \alpha_2, \dots, \alpha_n] = E,$$

展开可得

$$\begin{bmatrix} \alpha_1^H \alpha_1 & \alpha_1^H \alpha_2 & \cdots & \alpha_1^H \alpha_n \\ \alpha_2^H \alpha_1 & \alpha_2^H \alpha_2 & \cdots & \alpha_2^H \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n^H \alpha_1 & \alpha_n^H \alpha_2 & \cdots & \alpha_n^H \alpha_n \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

比较矩阵对应元素可得

$$\alpha_i^H \alpha_j = \delta_{ij} \quad (i, j = 1, 2, \dots, n),$$

故列向量组  $\alpha_1, \alpha_2, \dots, \alpha_n$  是标准正交向量组. 类似可证  $A$  的行向量组也是标准正交向量组.

(2) 充分性 若  $\alpha_1, \alpha_2, \dots, \alpha_n$  是标准正交向量组, 则有

$$\alpha_i^H \alpha_j = \delta_{ij} \quad (i, j = 1, 2, \dots, n),$$

于是

$$\begin{bmatrix} \alpha_1^H \\ \alpha_2^H \\ \vdots \\ \alpha_n^H \end{bmatrix} [\alpha_1, \alpha_2, \dots, \alpha_n] = E,$$

即

$$A^H A = E,$$

类似可证  $AA^H = E$ . 因此,  $A$  是酉矩阵. ■

**定义 3.1.4** 若  $\alpha_1, \alpha_2, \dots, \alpha_r$  为  $n$  维标准正交列向量组 ( $r < n$ ), 则称  $n \times r$  矩阵  $U_1 = (\alpha_1, \alpha_2, \dots, \alpha_r)$  为次酉矩阵, 记为  $U_1 \in U_r^{n \times r}$ .

**定理 3.1.2** 次酉矩阵  $U_1 \in U_r^{n \times r}$  的充要条件为  $U_1^H U_1 = E_r$ .

### 3.1.2 正交变换和酉变换

**定义 3.1.5** 设  $V$  是  $n$  维欧氏空间,  $\sigma$  是  $V$  的线性变换, 若  $\forall \alpha, \beta \in V$ , 都有

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta),$$

则称  $\sigma$  是  $V$  的正交变换.

**定义 3.1.6** 设  $V$  是  $n$  维酉空间,  $\sigma$  是  $V$  的线性变换, 若  $\forall \alpha, \beta \in V$ , 都有

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta),$$

则称  $\sigma$  是  $V$  的酉变换.

**定理 3.1.3** 设  $\sigma$  是欧氏空间(酉空间) $V$  的线性变换, 则下列命题等价:

- (1)  $\sigma$  是正交变换(酉变换);
- (2)  $\|\sigma(\alpha)\| = \|\alpha\|, \forall \alpha \in V$ ;
- (3) 变换前后两向量夹角保持不变;
- (4)  $\sigma$  将  $V$  的标准正交基变换到标准正交基;
- (5)  $\sigma$  在标准正交基下的矩阵是正交矩阵(酉矩阵).

证 (1) $\Rightarrow$ (2)  $\|\sigma(\alpha)\| = \sqrt{(\sigma(\alpha), \sigma(\alpha))} = \sqrt{(\alpha, \alpha)} = \|\alpha\|, \forall \alpha \in V.$

(2) $\Rightarrow$ (1) 由(2), 有

$$(\sigma(\alpha + \beta), \sigma(\alpha + \beta)) = (\alpha + \beta, \alpha + \beta),$$

$$(\sigma(\alpha + i\beta), \sigma(\alpha + i\beta)) = (\alpha + i\beta, \alpha + i\beta),$$

由于  $\sigma$  是线性变换, 因此根据内积性质展开第一个式子, 得

$$\text{左端} = (\sigma(\alpha), \sigma(\alpha)) + (\sigma(\alpha), \sigma(\beta)) + (\sigma(\beta), \sigma(\alpha)) + (\sigma(\beta), \sigma(\beta))$$

$$\text{右端} = (\alpha, \alpha) + (\alpha, \beta) + (\beta, \alpha) + (\beta, \beta)$$

而由(2), 有  $(\sigma(\alpha), \sigma(\alpha)) = (\alpha, \alpha), (\sigma(\beta), \sigma(\beta)) = (\beta, \beta)$ , 因此有

$$(\sigma(\alpha), \sigma(\beta)) + (\sigma(\beta), \sigma(\alpha)) = (\alpha, \beta) + (\beta, \alpha),$$

同理可得

$$(\sigma(\alpha), \sigma(\beta)) - (\sigma(\beta), \sigma(\alpha)) = (\alpha, \beta) - (\beta, \alpha),$$

将以上两式相加, 得

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta),$$

即  $\sigma$  是正交变换(酉变换).

(1)(2) $\Rightarrow$ (3) 两非零向量  $\alpha$  和  $\beta$  变换前的夹角为

$$\theta_1 = \arccos \frac{(\alpha, \beta)}{\|\alpha\| \|\beta\|}$$

变换后的  $\sigma(\alpha)$  和  $\sigma(\beta)$  的夹角为

$$\theta_2 = \arccos \frac{(\sigma(\alpha), \sigma(\beta))}{\|\sigma(\alpha)\| \|\sigma(\beta)\|}$$

由(1), 这两个夹角的分子相同, 由(2), 这两个夹角的分母相同, 故  $\theta_1 = \theta_2$ .

(1) $\Rightarrow$ (4) 设  $\alpha_1, \alpha_2, \dots, \alpha_n$  是  $V$  的标准正交基, 故

$$(\alpha_i, \alpha_j) = \delta_{ij} \quad (i, j = 1, 2, \dots, n),$$

若  $\sigma$  是酉变换, 则有

$$(\sigma(\alpha_i), \sigma(\alpha_j)) = (\alpha_i, \alpha_j) = \delta_{ij},$$

故  $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)$  仍是  $V$  的标准正交基.

(4) $\Rightarrow$ (1) 设  $\alpha_1, \alpha_2, \dots, \alpha_n$  与  $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)$  都是  $V$  的标准正交基,  $\forall \alpha, \beta \in V$  且

$$\alpha = x_1 \alpha_1 + x_2 \alpha_2 + \dots + x_n \alpha_n,$$

$$\beta = y_1 \alpha_1 + y_2 \alpha_2 + \dots + y_n \alpha_n,$$

则

$$\sigma(\alpha) = x_1\sigma(\alpha_1) + x_2\sigma(\alpha_2) + \cdots + x_n\sigma(\alpha_n),$$

$$\sigma(\beta) = y_1\sigma(\alpha_1) + y_2\sigma(\alpha_2) + \cdots + y_n\sigma(\alpha_n),$$

由内积运算性质,有

$$(\sigma(\alpha), \sigma(\beta)) = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n = (\alpha, \beta),$$

即  $\sigma$  是正交变换(酉变换).

(4) $\Rightarrow$ (5) 设  $\alpha_1, \alpha_2, \cdots, \alpha_n$  与  $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$  都是  $V$  的标准正交基,  $\mathbf{A} = (a_{ij})_{n \times n}$  是  $\sigma$  在基  $\alpha_1, \alpha_2, \cdots, \alpha_n$  下的矩阵, 现证  $\mathbf{A}$  为正交矩阵(酉矩阵).

由

$$(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\mathbf{A},$$

且  $\forall i, j (i, j = 1, 2, \cdots, n)$

$$\sigma(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \cdots + a_{ni}\alpha_n,$$

$$\sigma(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n,$$

于是

$$\begin{aligned} \delta_{ij} &= (\sigma(\alpha_i), \sigma(\alpha_j)) = \left( \sum_{k=1}^n a_{ki}\alpha_k, \sum_{h=1}^n a_{hj}\alpha_h \right) \\ &= \sum_{k=1}^n \sum_{h=1}^n a_{ki}\bar{a}_{hj} (\alpha_k, \alpha_h) \\ &= \sum_{k=1}^n \sum_{h=1}^n a_{ki}\bar{a}_{hj} \delta_{kh} \\ &= \sum_{k=1}^n a_{ki}\bar{a}_{kj}. \end{aligned}$$

即  $\mathbf{A}$  的第  $i$  列和第  $j$  列的内积为  $\delta_{ij}$ , 因此  $\mathbf{A}$  的列向量是标准正交向量组,  $\mathbf{A}$  为正交矩阵(酉矩阵).

(5) $\Rightarrow$ (4) 设  $\sigma$  在标准正交基  $\alpha_1, \alpha_2, \cdots, \alpha_n$  下的矩阵  $\mathbf{A} = (a_{ij})_{n \times n}$  是正交矩阵(酉矩阵), 即有

$$\sum_{k=1}^n a_{ki}\bar{a}_{kj} = \delta_{ij} \quad (i, j = 1, 2, \cdots, n)$$

因

$$(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\mathbf{A},$$

从而  $\forall i, j (i, j = 1, 2, \dots, n)$ , 有

$$(\sigma(\alpha_i), \sigma(\alpha_j)) = \left( \sum_{k=1}^n a_{ki} \alpha_k, \sum_{h=1}^n a_{hj} \alpha_h \right) = \sum_{k=1}^n a_{ki} \bar{a}_{kj} = \delta_{ij}$$

即  $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)$  也是标准正交基。

根据命题(2), 酉变换也可称为等距变换。这是因为

$$\begin{aligned} d(\alpha, \beta) &= \|\alpha - \beta\| = \|\sigma(\alpha - \beta)\| \\ &= \|\sigma(\alpha) - \sigma(\beta)\| \\ &= d(\sigma(\alpha), \sigma(\beta)). \end{aligned}$$

即向量  $\alpha, \beta$  之间的距离在线性变换  $\sigma$  下保持不变。

### 3.1.3 正交变换、酉变换实例

#### 1. 正交矩阵(酉矩阵)乘以向量是正交变换(酉变换)

在酉空间  $\mathbb{C}^n$  中, 对任意  $\mathbf{X} \in \mathbb{C}^n$ , 作变换  $T$ :

$$T(\mathbf{X}) = \mathbf{A}\mathbf{X},$$

其中,  $n$  阶方阵  $\mathbf{A}$  为酉矩阵, 则

$$(T(\mathbf{X}_1), T(\mathbf{X}_2)) = (\mathbf{A}\mathbf{X}_1, \mathbf{A}\mathbf{X}_2) = \mathbf{X}_2^H \mathbf{A}^H \mathbf{A} \mathbf{X}_1 = \mathbf{X}_2^H \mathbf{X}_1 = (\mathbf{X}_1, \mathbf{X}_2),$$

所以变换  $T$  是一个酉变换。

显然, 若欧氏空间  $\mathbb{R}^n$  中的方阵  $\mathbf{A}$  为正交矩阵, 令  $T(\mathbf{X}) = \mathbf{A}\mathbf{X}$ , 则  $T$  是一个正交变换。

#### 2. 旋转变换

以  $\mathbb{R}^2$  空间的旋转变换为例。如图 3.1.1 所示, 设向量  $\overrightarrow{OP}$  的长度为  $r$ , 辐角为  $\varphi$ 。将向量  $\overrightarrow{OP}$  逆时针旋转  $\theta$  角, 变为向量  $\overrightarrow{OP'}$ 。

若点  $P$  的坐标是  $(x, y)$ , 点  $P'$  的坐标是  $(x', y')$ , 则有

$$\begin{cases} x = r \cos \varphi \\ y = r \sin \varphi \\ x' = r \cos(\varphi + \theta) = r \cos \varphi \cos \theta - r \sin \varphi \sin \theta = x \cos \theta - y \sin \theta \\ y' = r \sin(\varphi + \theta) = r \cos \varphi \sin \theta + r \sin \varphi \cos \theta = x \sin \theta + y \cos \theta \end{cases},$$

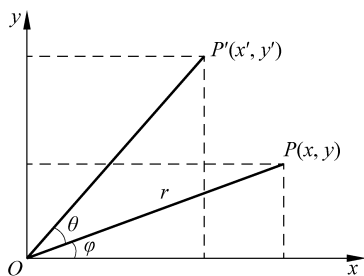


图 3.1.1 向量的旋转

可用矩阵乘法表示为

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

因此该旋转变换的矩阵为 2 阶矩阵

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

容易验证这是一个正交矩阵. 对应的线性变换为将  $xOy$  平面上的向量  $\overrightarrow{OP}$  绕坐标原点旋转  $\theta$  角变为向量  $\overrightarrow{OP'}$ , 因此该变换称为**旋转变换**.

**例 3.1.2** 3 阶矩阵

$$\mathbf{R}_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

是正交矩阵, 它表示三维空间中的向量绕  $x$  轴的旋转变换.

## 3.2 幂等矩阵与投影变换

### 3.2.1 幂等矩阵

**定义 3.2.1** 设  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , 若

$$\mathbf{A}^2 = \mathbf{A},$$

则称  $\mathbf{A}$  是**幂等矩阵**.

例如, 单位矩阵  $\mathbf{E}$  是幂等矩阵. 此外, 容易验证下面几种  $n$  阶方阵都是幂等矩阵.

$$(1) \mathbf{A} = \begin{bmatrix} \mathbf{E}_r & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{n \times n}, \quad \mathbf{A}_1 \in \mathbb{C}^{r \times (n-r)};$$

$$(2) \mathbf{A} = \begin{bmatrix} \mathbf{E}_r & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{n \times n}, \quad \mathbf{A}_2 \in \mathbb{C}^{(n-r) \times r};$$

$$(3) \mathbf{A} = \begin{bmatrix} \mathbf{E}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

用符号  $\mathbf{F}_r^{m \times n}$ 、 $\mathbf{R}_r^{m \times n}$ 、 $\mathbf{C}_r^{m \times n}$  分别表示元素在数域  $\mathbf{F}$ 、实数域  $\mathbf{R}$ 、复数域  $\mathbf{C}$  中的秩

为  $r$  的  $m \times n$  矩阵集合.

**定理 3.2.1** 设  $A \in \mathbb{C}^{n \times n}$  且  $A^2 = A$ , 则

(1)  $A^T, A^H, E-A, E-A^T, E-A^H$  都是幂等矩阵;

(2)  $A(E-A) = (E-A)A = \mathbf{0}$ .

**证** (1)  $(A^H)^2 = A^H A^H = (AA)^H = (A^2)^H = A^H$ .

故  $A^H$  是幂等矩阵, 同理可证  $A^T$  是幂等矩阵.

$$(E-A)^2 = E - 2A + A^2 = E - 2A + A = E - A.$$

故  $E-A$  是幂等矩阵.

$$\begin{aligned} (E-A^H)^2 &= E - 2A^H + (A^H)^2 = E - 2A^H + A^H A^H = E - 2A^H + (A^2)^H \\ &= E - 2A^H + A^H = E - A^H. \end{aligned}$$

故  $E-A^H$  是幂等矩阵, 同理可证  $E-A^T$  是幂等矩阵.

(2)  $A(E-A) = A - A^2 = A - A = \mathbf{0}$ .

$(E-A)A = A - A^2 = A - A = \mathbf{0}$ .

### 3.2.2 正交补与正交投影变换

与几何学中的正交(垂直)概念相似, 酉空间(欧氏空间)也存在子空间正交的概念.

**定义 3.2.2** 设  $V_1$  和  $V_2$  是  $n$  维酉空间  $V$  的子空间, 若对于任意  $x \in V_1, y \in V_2$ , 都有  $(x, y) = 0$ , 则称子空间  $V_1$  与  $V_2$  是正交的, 记为  $V_1 \perp V_2$ .

**例 3.2.1** 如图 3.2.1 所示, 设  $V$  是三维几何空间, 线性子空间  $V_1$  是  $z$  轴, 线性子空间  $V_2$  是  $xOy$  平面, 则可以证明  $V_1 \perp V_2$ .

**例 3.2.2** 设  $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$  是线性空间  $V$  的一组标准正交基, 若  $V_1 = \text{span}\{\alpha_1, \alpha_2\}, V_2 = \text{span}\{\alpha_3, \alpha_4, \alpha_5\}$ , 则可以证明  $V_1 \perp V_2$ .

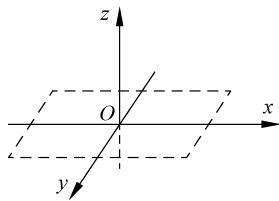


图 3.2.1 例 3.2.1 图示

**定理 3.2.2** 设  $V_1$  和  $V_2$  是  $n$  维酉空间  $V$  的两个正交子空间, 则

(1)  $V_1 \cap V_2 = \{\mathbf{0}\}$ ;

(2)  $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ .

**证:** (1) 设  $x \in V_1 \cap V_2$ , 由  $x \in V_1$ , 由子空间正交的定义可知, 对任意  $y \in V_2$ , 都有  $(x, y) = 0$ , 又因  $x \in V_2$ , 因此取  $y = x$ , 则  $(x, x) = 0$ , 于是  $x = \mathbf{0}$ . 根据  $x$  的任意

性可得  $V_1 \cap V_2 = \{\mathbf{0}\}$ .

(2) 由维数公式和(1)可得.

**定义 3.2.3** 若子空间  $V_1$  和  $V_2$  正交, 则  $V_1 + V_2$  称为  $V_1$  与  $V_2$  的正交和, 记为  $V_1 \oplus V_2$ .

下述定理给出了矩阵的核(值域)与其复共轭转置的值域(核)之间的关系.

**定理 3.2.3** 设  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , 则

$$(1) N(\mathbf{A}) \oplus R(\mathbf{A}^H) = \mathbb{C}^n;$$

$$(2) R(\mathbf{A}) \oplus N(\mathbf{A}^H) = \mathbb{C}^m.$$

**证:** (1) 设  $x \in N(\mathbf{A}), y \in R(\mathbf{A}^H)$ , 则  $\mathbf{A}x = \mathbf{0}, y = \mathbf{A}^H z$ , 其中  $z \in \mathbb{C}^m$ , 则

$$(x, y) = y^H x = z^H \mathbf{A}x = \mathbf{0},$$

故  $N(\mathbf{A}) \perp R(\mathbf{A}^H)$ .

又

$$\dim N(\mathbf{A}) + \dim R(\mathbf{A}^H) = (n - \text{rank} \mathbf{A}) + \text{rank} \mathbf{A} = n,$$

因此

$$N(\mathbf{A}) \oplus R(\mathbf{A}^H) = \mathbb{C}^n.$$

类似可得(2). ■

**定义 3.2.4** 设  $n$  维酉空间  $V$  的子空间  $V_1$  和  $V_2$  满足  $V_1 \oplus V_2 = V$ , 则称  $V_1$  为  $V_2$  的正交补, 记为  $(V_2)_\perp$ , 或

$$V_1 = (V_2)_\perp = \{\alpha \mid (\alpha, \beta) = 0, \forall \beta \in V_2\}.$$

显然, 若  $V_1$  是  $V_2$  的正交补, 则  $V_2$  也是  $V_1$  的正交补.

**定理 3.2.4** 设  $V_1$  是  $n$  维酉空间  $V$  的子空间, 则存在唯一的子空间  $V_2$ , 使得

$$V_1 \oplus V_2 = V.$$

证明略.

**例 3.2.3** 已知  $\alpha_1 = (1, 0, 1, 1)^T, \alpha_2 = (0, 1, 1, 2)^T, W = \text{span}\{\alpha_1, \alpha_2\}$ , 求  $W$  的正交补.

**分析** 取  $\mathbf{A} = (\alpha_1, \alpha_2)$ , 本题实际是求  $R(\mathbf{A})$  的正交补, 由定理 3.2.3(2) 的结论,  $R(\mathbf{A})$  的正交补是  $N(\mathbf{A}^H)$ .

**解:** 取  $\mathbf{A} = (\alpha_1, \alpha_2)$ , 则

$$\mathbf{A}^H = \begin{bmatrix} \alpha_1^H \\ \alpha_2^H \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix},$$



由定理 3.1.2 可知  $P$  是酉矩阵, 有  $P^{-1} = P^H$ , 故

$$P_S = PP_1P^H = P \begin{bmatrix} E_r \\ 0 \end{bmatrix} \begin{bmatrix} E_r \\ 0 \end{bmatrix}^H P^H = \left( P \begin{bmatrix} E_r \\ 0 \end{bmatrix} \right) \left( P \begin{bmatrix} E_r \\ 0 \end{bmatrix} \right)^H.$$

由于

$$\begin{aligned} \alpha_1, \alpha_2, \dots, \alpha_n &= (\gamma_1, \gamma_2, \dots, \gamma_n) P \\ &= (\gamma_1, \gamma_2, \dots, \gamma_n) \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2r} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nr} & \cdots & p_{nm} \end{bmatrix}, \end{aligned}$$

则有

$$P \begin{bmatrix} E_r \\ 0 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nr} \end{bmatrix} = U_1,$$

而  $U_1 \in U_r^{n \times r}$  的  $r$  个列向量是子空间  $S$  的标准正交基  $\alpha_1, \alpha_2, \dots, \alpha_r$  在  $V$  的标准正交基  $\gamma_1, \gamma_2, \dots, \gamma_n$  下的坐标向量. 所以,  $P_S = U_1 U_1^H, U_1 \in U_r^{n \times r}$ . ■

推论  $P_S^H = P_S = P_S^2$ .

### 3.3 对称变换、Hermite 变换及其矩阵

#### 3.3.1 对称变换与对称矩阵

**定义 3.3.1** 设  $\sigma$  是欧氏空间  $V$  的一个线性变换, 如果对任意的  $\alpha, \beta \in V$  都有  $(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta))$ , 则称  $\sigma$  为  $V$  的一个对称变换.

**定义 3.3.2** 设  $\sigma$  是欧氏空间  $V$  的一个线性变换, 如果对任意的  $\alpha, \beta \in V$  都有  $(\sigma(\alpha), \beta) = -(\alpha, \sigma(\beta))$ , 则称  $\sigma$  为  $V$  的一个反对称变换.

**例 3.3.1** 设  $W$  是欧氏空间  $V$  的一个子空间, 试证明  $V$  在  $W$  上的正交投影变换  $P$  是一个对称变换.

证: 任取  $\alpha, \beta \in V$ , 设

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in W, \quad \alpha_2 \in W_\perp,$$

$$\beta = \beta_1 + \beta_2, \quad \beta_1 \in W, \quad \beta_2 \in W_{\perp},$$

由正交投影的定义可知  $P(\alpha) = \alpha_1, P(\beta) = \beta_1$ , 那么

$$(P(\alpha), \beta) = (\alpha_1, \beta_1 + \beta_2) = (\alpha_1, \beta_1) + (\alpha_1, \beta_2) = (\alpha_1, \beta_1),$$

$$(\alpha, P(\beta)) = (\alpha_1 + \alpha_2, \beta_1) = (\alpha_1, \beta_1) + (\alpha_2, \beta_1) = (\alpha_1, \beta_1),$$

于是  $(P(\alpha), \beta) = (\alpha, P(\beta))$ , 故  $P$  是一个对称变换.  $\blacksquare$

**定理 3.3.1** 设  $\sigma$  是欧氏空间  $V$  的一个对称变换, 如果  $W$  是  $\sigma$  的不变子空间, 那么  $W_{\perp}$  也是  $\sigma$  的不变子空间.

**证:** 任取  $\alpha \in W_{\perp}$ , 需证明  $\sigma(\alpha) \in W_{\perp}$ . 对任意的  $\beta \in W$  有  $\sigma(\beta) \in W$ , 那么  $(\sigma(\beta), \alpha) = 0$ , 又  $\sigma$  是  $V$  的一个对称变换, 故

$$(\sigma(\beta), \alpha) = (\beta, \sigma(\alpha)) = 0.$$

这表明  $\sigma(\alpha) \in W_{\perp}$ .  $\blacksquare$

同理可得定理 3.3.2.

**定理 3.3.2** 设  $\sigma$  是欧氏空间  $V$  的一个反对称变换, 如果  $W$  是  $\sigma$  的不变子空间, 那么  $W_{\perp}$  也是  $\sigma$  的不变子空间.

**定理 3.3.3** 欧氏空间  $V$  的线性变换  $\sigma$  是对称变换的充要条件是  $\sigma$  在  $V$  的任意一个标准正交基下的矩阵是对称矩阵.

**证:** ① 必要性 任取  $V$  的一个标准正交基  $\alpha_1, \alpha_2, \dots, \alpha_n$ ,  $\sigma$  在该标准正交基下对应的矩阵为  $A = (a_{ij})_{n \times n}$ , 则由

$$(\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)) = (\alpha_1, \alpha_2, \dots, \alpha_n)A,$$

有

$$\sigma(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \dots + a_{ni}\alpha_n, \quad (\sigma(\alpha_i), \alpha_j) = a_{ji},$$

$$\sigma(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \dots + a_{nj}\alpha_n, \quad (\sigma(\alpha_j), \alpha_i) = a_{ij}.$$

由于  $\sigma$  是对称变换, 所以

$$a_{ji} = (\sigma(\alpha_i), \alpha_j) = (\alpha_i, \sigma(\alpha_j)),$$

$$a_{ij} = (\sigma(\alpha_j), \alpha_i) = (\alpha_i, \sigma(\alpha_j)),$$

故  $a_{ij} = a_{ji}$ , 这表明  $A$  是对称矩阵.

② 充分性 设线性变换  $\sigma$  在  $V$  的一个标准正交基  $\alpha_1, \alpha_2, \dots, \alpha_n$  下的矩阵  $A$  是对称矩阵, 任取  $\alpha, \beta \in V$ , 且在  $\alpha_1, \alpha_2, \dots, \alpha_n$  下的坐标分别为  $X$  和  $Y$ , 即

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)X,$$

$$\beta = (\alpha_1, \alpha_2, \dots, \alpha_n)Y.$$

于是有

$$\begin{aligned}\sigma(\alpha) &= \sigma(\alpha_1, \alpha_2, \dots, \alpha_n)X = (\alpha_1, \alpha_2, \dots, \alpha_n)AX, \\ \sigma(\beta) &= \sigma(\alpha_1, \alpha_2, \dots, \alpha_n)Y = (\alpha_1, \alpha_2, \dots, \alpha_n)AY, \\ (\alpha, \beta) &= ((\alpha_1, \alpha_2, \dots, \alpha_n)X, (\alpha_1, \alpha_2, \dots, \alpha_n)Y) = X^T Y,\end{aligned}$$

所以

$$\begin{aligned}(\sigma(\alpha), \beta) &= ((\alpha_1, \alpha_2, \dots, \alpha_n)AX, (\alpha_1, \alpha_2, \dots, \alpha_n)Y) \\ (AX)^T Y &= X^T A^T Y = X^T AY = (\alpha, \sigma(\beta)),\end{aligned}$$

这表明  $\sigma$  是  $V$  的对称变换. ■

同理可得定理 3.3.4.

**定理 3.3.4** 欧氏空间  $V$  的线性变换  $\sigma$  是反对称变换的充要条件是  $\sigma$  在  $V$  的任意一个标准正交基下的矩阵是反对称矩阵.

在线性代数课程中已经证明: 对  $n$  阶实对称矩阵  $A$ , 一定有  $n$  阶正交矩阵  $Q$ , 使  $Q^{-1}AQ$  为对角阵, 即实对称矩阵一定可以用正交矩阵相似对角化, 因此欧氏空间的对称变换是可对角化的线性变换.

**例 3.3.2** 在  $\mathbb{R}^3$  中, 设  $u$  为过直角坐标系原点的平面  $\pi$  的单位法向量, 变换  $T$  定义为

$$T(\alpha) = \alpha - 2(u, \alpha)u, \quad \alpha \in \mathbb{R}^3,$$

容易验证: 对任意  $\alpha, \beta \in \mathbb{R}^3$ , 任意  $k, l \in \mathbb{R}$ , 都有

$$\begin{aligned}T(k\alpha + l\beta) &= kT(\alpha) + lT(\beta), \\ (T(\alpha), T(\beta)) &= (\alpha, \beta), \\ (T(\alpha), \beta) &= (\alpha, T(\beta)).\end{aligned}$$

因此,  $T$  既是正交变换又是对称变换, 称其为镜面反射.

### 3.3.2 Hermite 矩阵与 Hermite 变换

**定义 3.3.3** 设  $A \in \mathbb{C}^{n \times n}$ , 若  $A^H = A$ , 则称  $A$  为 Hermite 矩阵, 若  $A^H = -A$ , 则称  $A$  为反 Hermite 矩阵.

对  $n$  阶矩阵  $A$ , 有

- (1)  $A^H = A \Leftrightarrow a_{ij} = \overline{a_{ji}} \Leftrightarrow \operatorname{Re}(a_{ij}) = \operatorname{Re}(a_{ji}), \operatorname{Im}(a_{ij}) = -\operatorname{Im}(a_{ji})$ ;
- (2)  $A^H = -A \Leftrightarrow a_{ij} = -\overline{a_{ji}} \Leftrightarrow \operatorname{Re}(a_{ij}) = -\operatorname{Re}(a_{ji}), \operatorname{Im}(a_{ij}) = \operatorname{Im}(a_{ji})$ .

以上两式中,  $i, j = 1, 2, \dots, n$ .

由定义, 一个 Hermite 矩阵的复共轭转置就是它自己, 因此 Hermite 矩阵也称为自共轭矩阵. Hermite 音译为“厄米特”或“埃尔米特”.

例如, 矩阵

$$\mathbf{A} = \begin{bmatrix} 1 & 2+3i \\ 2-3i & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 4 & 1-i & 2+3i \\ 1+i & 6 & 5+i \\ 2-3i & 5-i & 8 \end{bmatrix}$$

是 Hermite 矩阵.

矩阵

$$\mathbf{A} = \begin{bmatrix} 5i & -1-i & 4+3i \\ 1-i & 6i & -2 \\ -4+3i & 2 & 7i \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -3i & -1+i & 2-3i \\ 1+i & 2i & -1-i \\ -2-3i & 1-i & 0 \end{bmatrix}$$

是反 Hermite 矩阵.

由定义可知, Hermite 矩阵的主对角元素一定是实数, 反 Hermite 矩阵的主对角元素一定是 0 或纯虚数.

若 Hermite 矩阵  $\mathbf{A}$  的元素都是实数, 则  $\mathbf{A}$  就是实对称矩阵, 因此 Hermite 矩阵是实对称矩阵的推广. 通常 Hermite 矩阵不对称, 除非所有元素均为实数.

容易证明 Hermite 矩阵具有如下性质:

- (1) 若  $\mathbf{A}$  和  $\mathbf{B}$  是 Hermite 矩阵, 则  $\mathbf{A} + \mathbf{B}$  也是 Hermite 矩阵.
- (2) 若  $\mathbf{A}$  和  $\mathbf{B}$  是 Hermite 矩阵且  $\mathbf{AB} = \mathbf{BA}$ , 则  $\mathbf{AB}$  也是 Hermite 矩阵.
- (3) 若  $\mathbf{A}$  是 Hermite 矩阵且可逆, 则其逆矩阵  $\mathbf{A}^{-1}$  也是 Hermite 矩阵.
- (4) 若  $\mathbf{A}$  是 Hermite 矩阵, 对  $n \in \mathbb{Z}^+$ , 则  $\mathbf{A}^n$  也是 Hermite 矩阵.
- (5) 方阵  $\mathbf{A}$  与其共轭转置的和是 Hermite 矩阵.
- (6) Hermite 矩阵的特征值一定是实数.

齐次式是指合并同类项后每一项关于未知量的次数都相等的多项式. 例如,  $3x+2y, x-2y$  的各项都是 1 次的, 称为一次齐次式,  $3x^2-2xy+y^2, x^2+xy$  的各项都是 2 次的, 称为二次齐次式. 线性代数中的二次型就是二次齐次式.

**定义 3.3.4** 设  $\mathbf{A}$  是  $n$  阶 Hermite 矩阵,  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  是任意一个  $n$  维复向量, 称

$$f(\mathbf{x}) = \mathbf{x}^H \mathbf{A} \mathbf{x}$$

为 **Hermite 二次齐式**, 也称为 **Hermite 二次型**.

由定义可知  $f(\mathbf{x})$  是  $1 \times 1$  维矩阵, 即其值是一个数. 对任意  $n$  维复向量  $\mathbf{x}$ , Hermite 二次型  $f(\mathbf{x})$  的值总是实数, 这是因为

$$\overline{f(\mathbf{x})} = \overline{\mathbf{x}^H \mathbf{A} \mathbf{x}} = \overline{\mathbf{x}^T \mathbf{A} \mathbf{x}} = (\mathbf{x}^T \mathbf{A} \mathbf{x})^T = \mathbf{x}^H \mathbf{A}^H \mathbf{x} = \mathbf{x}^H \mathbf{A} \mathbf{x} = f(\mathbf{x}).$$

**定义 3.3.5** 给定 Hermite 二次型

$$f(\mathbf{x}) = \mathbf{x}^H \mathbf{A} \mathbf{x},$$

其中,  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ , 如果对任意一组不全为 0 的复数  $x_1, x_2, \dots, x_n$ , 都有  $f(\mathbf{x}) > 0$  ( $\geq 0$ ), 则称该 Hermite 二次型是**正定的**(**半正定的**), 并称对应的 Hermite 矩阵  $\mathbf{A}$  是**正定的**(**半正定的**). 如果对任意一组不全为 0 的复数  $x_1, x_2, \dots, x_n$ , 都有  $f(\mathbf{x}) < 0$  ( $\leq 0$ ), 则称该 Hermite 二次型是**负定的**(**半负定的**), 并称对应的 Hermite 矩阵  $\mathbf{A}$  是**负定的**(**半负定的**).

显然, 若 Hermite 二次型  $f(\mathbf{x})$  是正定的(负定的), 则  $-f(\mathbf{x})$  是负定的(正定的).

**定理 3.3.5** 对于 Hermite 二次型  $f(\mathbf{x}) = \mathbf{x}^H \mathbf{A} \mathbf{x}, \mathbf{x} \in \mathbb{C}^n$ , 下列命题等价:

- (1)  $f(\mathbf{x})$  是正定的,  $\mathbf{A}$  是正定的.
- (2) 对于任意  $n$  阶可逆矩阵  $\mathbf{P}$ , 都有  $\mathbf{P}^H \mathbf{A} \mathbf{P}$  为正定矩阵.
- (3)  $\mathbf{A}$  的  $n$  个特征值全大于零.

**定理 3.3.6** 对于 Hermite 二次型  $f(\mathbf{x}) = \mathbf{x}^H \mathbf{A} \mathbf{x}, \mathbf{x} \in \mathbb{C}^n$ , 下列命题等价:

- (1)  $f(\mathbf{x})$  是半正定的,  $\mathbf{A}$  是半正定的.
- (2) 对于任意  $n$  阶可逆矩阵  $\mathbf{P}$ , 都有  $\mathbf{P}^H \mathbf{A} \mathbf{P}$  为半正定矩阵.
- (3)  $\mathbf{A}$  的  $n$  个特征值全是非负实数.

从欧氏空间的对称变换往复数域推广, 可以定义酉空间的 Hermite 变换.

**定义 3.3.6** 设  $V$  是一个酉空间,  $\sigma$  是  $V$  的一个线性变换, 若  $\forall \alpha, \beta \in V$ , 都有

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)),$$

则称  $\sigma$  是  $V$  的一个 **Hermite 变换**.

**定义 3.3.7** 设  $V$  是一个酉空间,  $\sigma$  是  $V$  的一个线性变换, 若  $\forall \alpha, \beta \in V$ , 都有

$$(\sigma(\alpha), \beta) = -(\alpha, \sigma(\beta)),$$

则称  $\sigma$  是  $V$  的一个 **反 Hermite 变换**.

由定义可见, 酉空间中的反 Hermite 变换与欧氏空间中的反对称变换形式相同.

**定理 3.3.7** 酉空间  $V$  的线性变换  $\sigma$  是 Hermite 变换的充要条件是  $\sigma$  在  $V$  的任意一个标准正交基下的矩阵是 Hermite 矩阵.

**证:** ① 必要性 任取  $V$  的一个标准正交基  $\alpha_1, \alpha_2, \dots, \alpha_n$ ,  $\sigma$  在该标准正交基下对应的矩阵为  $\mathbf{A} = (a_{ij})_{n \times n}$ , 则由

$$(\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)) = (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{A}$$

有

$$\sigma(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \dots + a_{ni}\alpha_n, \quad (\sigma(\alpha_i), \alpha_j) = a_{ji},$$

$$\sigma(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \dots + a_{nj}\alpha_n, \quad (\sigma(\alpha_j), \alpha_i) = a_{ij}.$$

由于  $\sigma$  是 Hermite 变换, 所以

$$a_{ji} = (\sigma(\alpha_i), \alpha_j) = (\alpha_i, \sigma(\alpha_j)),$$

$$\overline{a_{ij}} = \overline{(\sigma(\alpha_j), \alpha_i)} = (\alpha_i, \sigma(\alpha_j)),$$

故  $a_{ji} = \overline{a_{ij}}$ , 这表明  $\mathbf{A} = \mathbf{A}^H$ , 即  $\mathbf{A}$  是 Hermite 矩阵.

② 充分性 设线性变换  $\sigma$  在  $V$  的一个标准正交基  $\alpha_1, \alpha_2, \dots, \alpha_n$  下的矩阵  $\mathbf{A}$  是 Hermite 矩阵, 即  $\mathbf{A} = \mathbf{A}^H$ , 任取  $\alpha, \beta \in V$ , 且在  $\alpha_1, \alpha_2, \dots, \alpha_n$  下的坐标分别为  $\mathbf{X}$  和  $\mathbf{Y}$ , 即

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{X},$$

$$\beta = (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{Y}.$$

于是有

$$\sigma(\alpha) = \sigma(\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{X} = (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{A}\mathbf{X},$$

$$\sigma(\beta) = \sigma(\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{Y} = (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{A}\mathbf{Y},$$

$$(\alpha, \beta) = ((\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{X}, (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{Y}) = \mathbf{Y}^H \mathbf{X},$$

所以

$$\begin{aligned} (\sigma(\alpha), \beta) &= ((\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{A}\mathbf{X}, (\alpha_1, \alpha_2, \dots, \alpha_n)\mathbf{Y}) \\ &= \mathbf{Y}^H \mathbf{A}\mathbf{X} = \mathbf{Y}^H \mathbf{A}^H \mathbf{X} = (\mathbf{A}\mathbf{Y})^H \mathbf{X} = (\alpha, \sigma(\beta)), \end{aligned}$$

这表明  $\sigma$  是  $V$  的 Hermite 变换. ■

**定理 3.3.8** 酉空间  $V$  的线性变换  $\sigma$  是反 Hermite 变换的充要条件是  $\sigma$  在  $V$  的任意一个标准正交基下的矩阵是反 Hermite 矩阵.

### 3.4 正规矩阵与正规变换

#### 3.4.1 正规矩阵

**定义 3.4.1** 设  $A \in \mathbb{C}^{n \times n}$ , 若  $AA^H = A^H A$ , 则称  $A$  为正规矩阵.

**定义 3.4.2** 设  $A \in \mathbb{R}^{n \times n}$ , 若  $AA^T = A^T A$ , 则称  $A$  为实正规矩阵.

显然, 对角阵、Hermite 矩阵、反 Hermite 矩阵、酉矩阵是正规矩阵, 实对称矩阵、实反对称矩阵、正交矩阵是实正规矩阵.

**定义 3.4.3** 设  $A, B \in \mathbb{C}^{n \times n}$ , 若存在  $U \in U^{n \times n}$ , 使

$$U^H A U = U^{-1} A U = B,$$

则称  $A$  酉相似于  $B$ .

**定义 3.4.4** 设  $A, B \in \mathbb{R}^{n \times n}$ , 若存在  $U \in E^{n \times n}$ , 使

$$U^T A U = U^{-1} A U = B,$$

则称  $A$  正交相似于  $B$ .

酉相似和正交相似符合矩阵相似的定义, 且相似因子  $U$  是酉矩阵或正交矩阵.

**引理 3.4.1** (Schur 引理) 任何一个  $n$  阶方阵  $A \in \mathbb{C}^{n \times n}$  都酉相似于一个  $n$  阶上三角阵.

**引理 3.4.2** 设  $A$  是正规矩阵, 则与  $A$  酉相似的矩阵都是正规矩阵.

请读者按定义证明.

**引理 3.4.3** 设  $A$  是正规矩阵, 且  $A$  是三角矩阵, 则  $A$  是对角阵.

证: 设

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix}, \quad A^H = \begin{bmatrix} \bar{a}_{11} & & & \\ \bar{a}_{12} & \bar{a}_{22} & & \\ \vdots & & \ddots & \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn} \end{bmatrix},$$

代入  $AA^H = A^H A$ , 比较等式两端矩阵主对角线上的元素, 可得  $n$  个等式:

$$a_{11}\bar{a}_{11} + a_{12}\bar{a}_{12} + \cdots + a_{1n}\bar{a}_{1n} = \bar{a}_{11}a_{11},$$

$$a_{22}\bar{a}_{22} + \cdots + a_{2n}\bar{a}_{2n} = \bar{a}_{12}a_{12} + \bar{a}_{22}a_{22},$$

$\vdots$

$$a_{mn}\bar{a}_{nm} = \bar{a}_{1n}a_{1n} + \bar{a}_{2n}a_{2n} + \cdots + \bar{a}_{nm}a_{nm}.$$

由第一个等式,有  $a_{12}\bar{a}_{12} + a_{13}\bar{a}_{13} + \cdots + a_{1n}\bar{a}_{1n} = 0$ , 又  $a_{ij}\bar{a}_{ij} = |a_{ij}|^2 \geq 0$ , 可得  $a_{12}\bar{a}_{12} = a_{13}\bar{a}_{13} = \cdots = a_{1n}\bar{a}_{1n} = 0$ , 将此结果再代入第二个等式, 以此类推直至第  $n$  个等式, 可得: 当  $i \neq j$  时,  $a_{ij} = 0$ , 即  $\mathbf{A}$  是对角阵. ■

**定理 3.4.1** (正规矩阵的结构定理) 设  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , 则  $\mathbf{A}$  是正规矩阵的充要条件是存在  $\mathbf{U} \in U^{n \times n}$ , 使

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

其中  $\lambda_1, \lambda_2, \cdots, \lambda_n$  是  $\mathbf{A}$  的特征值.

**证:** ① 必要性 根据 Schur 引理知, 存在  $\mathbf{U} \in U^{n \times n}$ , 使得

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{B} \text{ (上三角阵)},$$

根据引理 3.4.2 知,  $\mathbf{B}$  是正规矩阵, 又根据引理 3.4.3 知,  $\mathbf{B}$  是对角阵.

② 充分性 因为对角阵  $\text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$  是正规矩阵, 由引理 3.4.2,  $\mathbf{A}$  是正规矩阵. ■

由定理 3.4.1 可知: 正规矩阵的本质是可以对角化.

**推论 3.4.1** 设  $\mathbf{A}$  是正规矩阵,  $\lambda_i$  是  $\mathbf{A}$  的特征值, 对应的特征向量是  $\mathbf{x}$ , 则  $\bar{\lambda}_i$  是  $\mathbf{A}^H$  的特征值, 其对应的特征向量是  $\mathbf{x}$ .

将定理 3.4.1 中的等式两端取复共轭转置, 可得该推论.

**推论 3.4.2**  $n$  阶正规矩阵  $\mathbf{A}$  有  $n$  个线性无关的特征向量.

**证:** 设  $\mathbf{A}$  是  $n$  阶正规矩阵,  $\lambda_i (i = 1, 2, \cdots, n)$  是  $\mathbf{A}$  的特征值, 则存在  $\mathbf{U} \in U^{n \times n}$ , 使

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

设  $\mathbf{U}$  的列向量组为  $\alpha_1, \alpha_2, \cdots, \alpha_n$ , 则有

$$\mathbf{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n) \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

即  $(A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n)$ , 可得

$$A\alpha_i = \lambda_i\alpha_i \quad (i=1, 2, \dots, n).$$

又由于  $\alpha_1, \alpha_2, \dots, \alpha_n$  是酉矩阵  $U$  的列向量组, 所以  $\alpha_1, \alpha_2, \dots, \alpha_n$  一定是线性无关向量组, 且  $\alpha_i (i=1, 2, \dots, n)$  都是非零的, 由上式可知  $\alpha_i$  是  $A$  的属于特征值  $\lambda_i$  的特征向量  $(i=1, 2, \dots, n)$ , 即  $n$  阶正规矩阵  $A$  有  $n$  个线性无关的特征向量. ■

**定义 3.4.5** 属于某特征值的特征向量和零向量生成的子空间称为属于该特征值的特征子空间.

**推论 3.4.3** 正规矩阵的属于不同特征值的特征子空间是相互正交的.

在线性代数中已经证明:  $n$  阶实对称矩阵  $A$  一定可以用正交矩阵相似对角化, 且该对角阵的主对角线元素恰是  $A$  的全部特征值. 类似地, 现在讨论当  $A$  是  $n$  阶正规矩阵时, 如何求相似因子  $U$ , 使  $U^H A U$  为对角阵.

由定理 3.4.1 和推论 3.4.2 的推导过程可得出相似因子  $U$  的求解方法:

- (1) 求  $A$  的特征值, 即求  $|\lambda E - A| = 0$  的根  $\lambda_1, \lambda_2, \dots, \lambda_n$ ;
- (2) 对每一个相异特征值  $\lambda_i, i=1, 2, \dots, s$ , 求  $\lambda_i$  的特征向量, 即求  $(\lambda_i E - A)X = 0$  的非零解;
- (3) 用 Schmidt 方法求  $\lambda_i (i=1, 2, \dots, s)$  的特征向量生成的空间的标准正交基  $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in_i}$ ;
- (4) 令  $U = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n_1}, \alpha_{21}, \alpha_{22}, \dots, \alpha_{2n_2}, \dots, \alpha_{s1}, \alpha_{s2}, \dots, \alpha_{sn_s})$ , 则得到的酉矩阵  $U$  满足  $U^H A U = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ .

**例 3.4.1** 已知  $A = \begin{bmatrix} -1 & i & 0 \\ -i & 0 & -i \\ 0 & i & -1 \end{bmatrix}$ , 验证  $A$  是正规矩阵, 并求酉矩阵  $U$ , 使

$U^H A U$  为对角阵.

**解:** 由于  $A^H = A$ , 即  $A$  是 Hermite 矩阵, 从而  $A$  是正规矩阵.

$A$  的特征多项式

$$|\lambda E - A| = (\lambda - 1)(\lambda + 1)(\lambda + 2),$$

故  $A$  的特征值是  $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = -2$ .

求得对应的特征向量分别为

$$\alpha_1 = (1, -2i, 1)^T,$$

$$\alpha_2 = (-1, 0, 1)^T,$$

$$\alpha_3 = (1, i, 1)^T.$$

这些特征向量应是两两正交的,单位化得

$$u_1 = \frac{\alpha_1}{\|\alpha_1\|} = \left( \frac{1}{\sqrt{6}}, -\frac{2i}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right)^T,$$

$$u_2 = \frac{\alpha_2}{\|\alpha_2\|} = \left( -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^T,$$

$$u_3 = \frac{\alpha_3}{\|\alpha_3\|} = \left( \frac{1}{\sqrt{3}}, \frac{i}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)^T,$$

于是酉矩阵

$$U = (u_1, u_2, u_3) = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2i}{\sqrt{6}} & 0 & \frac{i}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

满足

$$U^H A U = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -2 \end{bmatrix}.$$

### 3.4.2 伴随变换和正规变换

**定义 3.4.6** 设  $V$  是一个酉空间(欧氏空间),  $\sigma$  是  $V$  的一个线性变换, 如果存在  $V$  的一个线性变换  $\sigma^H$ , 使得

$$(\sigma(\alpha), \beta) = (\alpha, \sigma^H(\beta)) \quad (\forall \alpha, \beta \in V),$$

则称  $\sigma$  有一个伴随变换  $\sigma^H$ .

**例 3.4.2** 设  $\sigma$  是欧氏空间  $V$  的一个对称变换, 则根据对称变换的定义

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)) \quad (\forall \alpha, \beta \in V),$$

则有  $\sigma^H = \sigma$ .

**例 3.4.3** 设  $\sigma$  是酉空间  $V$  的一个 Hermite 变换, 则根据 Hermite 变换的定义

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)) \quad (\forall \alpha, \beta \in V),$$

则有  $\sigma^H = \sigma$ . 因此, Hermite 变换也被称为自伴随变换.

**例 3.4.4** 设  $\sigma$  是欧氏空间  $V$  的一个正交变换, 则根据正交变换的定义

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta) \quad (\forall \alpha, \beta \in V),$$

则有  $\sigma^H = \sigma^{-1}$ . 这是因为

$$(\sigma(\alpha), \beta) = (\sigma(\alpha), \sigma(\sigma^{-1}(\beta))) = (\alpha, \sigma^{-1}(\beta)) \quad (\forall \alpha, \beta \in V).$$

**例 3.4.5** 设  $\sigma$  是酉空间  $V$  的一个酉变换, 则有  $\sigma^H = \sigma^{-1}$ .

**定义 3.4.7** 设  $V$  是酉空间(欧氏空间),  $\sigma$  是  $V$  的一个线性变换, 如果  $\sigma$  满足

$$\sigma^H \sigma = \sigma \sigma^H,$$

则称  $\sigma$  是正规变换.

欧氏空间的正交变换、对称变换、反对称变换, 酉空间的酉变换、Hermite 变换、反 Hermite 变换都是正规变换.

**定理 3.4.2** 设  $V$  是一个  $n$  维酉空间(欧氏空间),  $\sigma$  是  $V$  的一个线性变换,  $\alpha_1, \alpha_2, \dots, \alpha_n$  是  $V$  的一个标准正交基, 且  $\sigma$  在基  $\alpha_1, \alpha_2, \dots, \alpha_n$  下对应的矩阵为  $A = (a_{ij})_{n \times n}$ , 则  $\sigma$  的伴随变换  $\sigma^H$  在基  $\alpha_1, \alpha_2, \dots, \alpha_n$  下的矩阵  $B$  为  $B = A^H$ .

**证:** 由伴随变换定义可知

$$(\sigma(\alpha_i), \alpha_j) = (\alpha_i, \sigma^H(\alpha_j)) \quad (i, j = 1, 2, \dots, n),$$

由

$$(\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)) = (\alpha_1, \alpha_2, \dots, \alpha_n)A$$

有

$$\sigma(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \dots + a_{ni}\alpha_n, \quad (\sigma(\alpha_i), \alpha_j) = a_{ji},$$

设  $\sigma$  的伴随变换  $\sigma^H$  的矩阵  $B = (b_{ij})_{n \times n}$ , 则有

$$(\sigma(\alpha_i), \alpha_j) = (\alpha_i, \sigma^H(\alpha_j)) = \overline{(\sigma^H(\alpha_j), \alpha_i)} = \overline{b_{ij}},$$

从而有  $a_{ji} = \overline{b_{ij}}, b_{ij} = \overline{a_{ji}}$ , 即  $B = A^H$ . ■

**定理 3.4.3** 设  $V$  是一个  $n$  维酉空间(欧氏空间),  $\sigma$  是  $V$  的一个线性变换,  $\sigma$  是正规变换当且仅当  $\sigma$  在  $V$  的任意一个标准正交基下的矩阵是正规矩阵.

**证:** 任取  $V$  的一个标准正交基  $\alpha_1, \alpha_2, \dots, \alpha_n$ , 设  $\sigma$  在该基下的矩阵为  $A$ , 那么其伴随变换  $\sigma^H$  在该基下的矩阵为  $A^H$ , 又因变换与矩阵的一一对应关系, 有

$$\sigma\sigma^H = \sigma^H\sigma \Leftrightarrow AA^H = A^HA,$$

所以,  $\sigma$  是  $V$  的正规变换当且仅当  $A$  为正规矩阵. ■

由于正规变换的矩阵是正规矩阵,且正规矩阵一定酉相似于一个对角阵,因此可得下述定理.

**定理 3.4.4** 设  $\sigma$  是酉空间  $V$  的一个正规变换,则存在  $V$  的一个标准正交基,使得  $\sigma$  在这个基下对应的矩阵为对角阵.

由定理 3.4.4 可知酉空间的正规变换是可对角化的线性变换.

## 3.5 应用实例

### 3.5.1 Householder 镜像变换

**定义 3.5.1** 设  $\alpha \in \mathbb{C}^n$ , 且  $\alpha^H \alpha = 1$ , 则酉矩阵

$$H = E - 2\alpha\alpha^H \in \mathbb{C}^{n \times n}$$

与某一组基下的线性变换对应,该线性变换是酉变换,称为 **Householder**(豪斯霍尔德)镜像变换.

现以  $\mathbb{R}^2$  空间对 Householder 变换进行分析. 如图 3.5.1 所示,  $\mathbb{R}^2$  空间中有两个同样长度的向量  $\xi$  和  $\eta$ , 而  $\alpha$  是与  $\xi - \eta$  同方向的单位向量(向量  $\alpha$  的长度是  $\sqrt{(\alpha, \alpha)} = \sqrt{\alpha^H \alpha} = 1$ ),  $l$  轴(无方向)与向量  $\alpha$  垂直.

向量  $\xi$  在  $\alpha$  上的投影长度是

$$(\xi, \alpha) = \alpha^H \xi,$$

因此

$$\xi - \eta = 2\alpha(\alpha^H \xi),$$

所以

$$\eta = \xi - 2\alpha(\alpha^H \xi) = (E - 2\alpha\alpha^H)\xi = H\xi.$$

又因  $l$  轴与向量  $\alpha$  正交,由几何关系可知向量  $\xi$  和  $\eta$  关于  $l$  轴呈镜像对称关系.

综上,矩阵  $H = E - 2\alpha\alpha^H$  作用在原始向量  $\xi$  上,代表对  $\xi$  进行镜面反射,反射平面(对应图 3.5.1 中的  $l$  轴)正是与向量  $\alpha$  垂直的平面(向量、平面都要过坐标原点),  $\eta = H\xi$  即为变换所得向量,因此,Householder 变换又称为**镜像变换**,酉矩阵  $H$  也称为 Householder 矩阵,镜面的法向量  $\alpha$  也称为 Householder 向量.

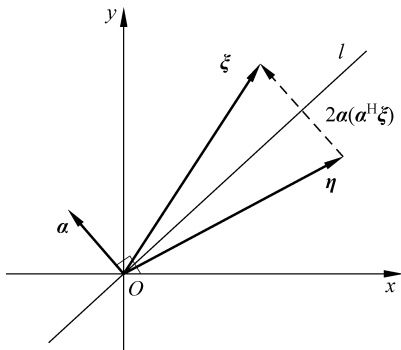


图 3.5.1 Householder 变换示意图

Householder 变换能将向量变换成范数相等的另一向量, 若需变换回原向量, 只需对变换后的向量再进行一次 Householder 变换. 特别地, Householder 变换能将一般向量变换成范数相等且若干个分量为 0 的稀疏向量, 这在参数估计、信息处理等方面有实际应用价值. 此外, Householder 变换还可用于矩阵分解等方面.

### 3.5.2 最小二乘法的数学原理

最小二乘法又称最小平方法, 是一种数学优化方法, 它通过最小化误差的平方之和寻找数据的最佳函数匹配. 本节基于正交补和正交投影介绍最小二乘法的数学原理.

**定理 3.5.1** 设  $W$  是欧氏空间  $V$  的一个线性子空间, 对于任意向量  $\alpha \in V$ ,  $\alpha_1 \in W$  是  $\alpha$  在  $W$  上的正交投影的充要条件是

$$|\alpha - \alpha_1| \leq |\alpha - \beta| \quad (\forall \beta \in W).$$

**证:** ① 必要性 设  $\alpha_1 \in W$  是  $\alpha$  在  $W$  上的正交投影, 则对  $\forall \beta \in W$ , 有

$$(\alpha - \alpha_1) \perp (\alpha_1 - \beta),$$

$|\alpha - \beta|^2 = |(\alpha - \alpha_1) + (\alpha_1 - \beta)|^2 = |\alpha - \alpha_1|^2 + |\alpha_1 - \beta|^2 \geq |\alpha - \alpha_1|^2$ ,  
即  $|\alpha - \alpha_1| \leq |\alpha - \beta|$ .

② 充分性 设  $\delta$  是  $\alpha$  在  $W$  上的正交投影, 则对于  $\alpha_1 \in W$ , 有

$$(\alpha - \delta) \perp (\delta - \alpha_1),$$

$|\alpha - \alpha_1|^2 = |(\alpha - \delta) + (\delta - \alpha_1)|^2 = |\alpha - \delta|^2 + |\delta - \alpha_1|^2 \geq |\alpha - \delta|^2$ ,  
即  $|\alpha - \alpha_1| \geq |\alpha - \delta|$ . 又由于  $|\alpha - \alpha_1| \leq |\alpha - \beta|, \forall \beta \in W$ , 有  $|\alpha - \alpha_1| \leq |\alpha - \delta|$ ,  
所以  $|\alpha - \alpha_1| = |\alpha - \delta|, \forall \alpha \in V$ , 所以  $\alpha_1 = \delta$ . ■

**定义 3.5.2** 设  $W$  是欧氏空间  $V$  的一个线性子空间, 对于任意向量  $\alpha \in V$ , 若存在  $\alpha_1 \in W$ , 使得

$$|\alpha - \alpha_1| \leq |\alpha - \beta| \quad (\forall \beta \in W),$$

则称  $\alpha_1$  是  $\alpha$  在  $W$  上的最佳逼近元.

$\alpha$  在  $W$  上的最佳逼近元也就是  $\alpha$  在  $W$  上的正交投影.

在科学实验和统计分析中, 经常需要分析观测数据的变化规律并用函数近似表示, 从图形上看, 就是求一条能反映数据变化趋势的近似曲线, 当近似曲线为直线时称为直线拟合或线性回归, 为曲线时称为曲线拟合.

(1) 直线拟合.

当观测数据  $y$  随某变量  $x$  的变化关系满足线性关系时, 拟合函数为  $\varphi(x) = a_0 + a_1x$ , 为确定参数  $a_0$  和  $a_1$  的值, 可通过实验测得观测数据, 设总共进行了  $n$  次观测, 数据为  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , 由这组观测数据的散点图可看出, 数据点一般并不在一条直线上, 但却在一条直线附近, 即  $y_i \approx a_0 + a_1x_i, i = 1, 2, \dots, n$ . 由于不存在  $a_0$  和  $a_1$  使得  $y_i = a_0 + a_1x_i$ , 因此方程组

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

在  $n > 2$  时是矛盾方程组, 没有通常意义下的解.

上述方程组的矩阵形式为

$$\mathbf{A}\mathbf{a} = \mathbf{Y},$$

其中,

$$\mathbf{A} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

该问题是求参数向量  $\mathbf{a}$  使  $\mathbf{Y} - \mathbf{A}\mathbf{a}$  长度最小, 若将  $\mathbf{Y}$  看成  $\mathbb{R}^n$  空间中的一个向量, 则问题转化为求  $\mathbf{a}$  使  $\mathbf{A}\mathbf{a}$  是向量  $\mathbf{Y}$  在系数矩阵  $\mathbf{A}$  的列向量  $(1, 1, \dots, 1)^\top$  和  $(x_1, x_2, \dots, x_n)^\top$  生成的子空间  $W$  上的投影.

确定待定参数  $a_0$  和  $a_1$  的原则应是使观测数据的误差  $e_i = y_i - (a_0 + a_1x_i)$  ( $i = 1, 2, \dots, n$ ) 尽可能小, 这是一个多目标优化问题. 将此多目标优化问题转化为单目标优化问题有几种选择:

- ① 使最大误差最小, 即  $\min_{a_0, a_1} \max_{1 \leq i \leq n} |e_i|$ ;
- ② 使误差的绝对值之和最小, 即  $\min_{a_0, a_1} \sum_{i=1}^n |e_i|$ ;
- ③ 使误差的平方之和最小, 即  $\min_{a_0, a_1} \sum_{i=1}^n e_i^2$ .

采用第③种目标函数,即以误差的平方之和最小为目标的方法就是**最小二乘法**.

(2) 多变量拟合.

当观测数据  $y$  与  $n$  个变量  $x_1, x_2, \dots, x_n$  呈线性关系时,拟合函数可表示为

$$\varphi(\mathbf{x}) = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n.$$

设总共进行了  $m$  次观测,为确定待定参数  $a_0, a_1, \dots, a_n$ ,则有方程组

$$\begin{cases} a_0 + a_1x_{11} + a_2x_{21} + \dots + a_nx_{n1} = y_1 \\ a_0 + a_1x_{12} + a_2x_{22} + \dots + a_nx_{n2} = y_2 \\ \vdots \\ a_0 + a_1x_{1m} + a_2x_{2m} + \dots + a_nx_{nm} = y_m \end{cases},$$

其矩阵形式为

$$\mathbf{A}\mathbf{a} = \mathbf{Y},$$

其中

$$\mathbf{A} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{n1} \\ 1 & x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1m} & x_{2m} & \cdots & x_{nm} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

当  $m > n$  时该方程组是矛盾方程组,没有通常意义下的解. 确定待定参数  $a_0, a_1, \dots, a_n$  的原则应是使误差  $e_i = y_i - \varphi(\mathbf{x}_i)$  ( $i=1, 2, \dots, m$ ) 尽可能小,这是一个多目标优化问题,将此多目标优化问题转化为单目标优化问题同样有几种目标函数可供选择,需确定最优目标函数.

(3) 曲线拟合.

当观测数据  $y$  随某变量  $x$  的变化规律为曲线时,需求一条近似曲线来反映数据的变化趋势,通常用多项式表示近似曲线,拟合函数可表示为

$$\varphi(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

其中,  $a_0, a_1, \dots, a_n$  为待定参数. 设总共进行了  $m$  次观测,确定待定参数  $a_0, a_1, \dots, a_n$  的原则应是使误差  $e_i = y_i - \varphi(x_i)$  ( $i=1, 2, \dots, m$ ) 尽可能小,这是一个多目标优化问题,将其转化为单目标优化问题同样需确定最优目标函数.

进行变换

$$\begin{cases} z_1 = x \\ z_2 = x^2 \\ \vdots \\ z_n = x^n \end{cases},$$

则拟合函数可写为

$$\varphi(\mathbf{z}) = a_0 + a_1 z_1 + a_2 z_2 + \cdots + a_n z_n,$$

从而转化为多变量拟合问题进行处理.

对于最小二乘法,有

$$\sum_{i=1}^n e_i^2 = (\mathbf{Y} - \mathbf{A}\mathbf{a})^\top (\mathbf{Y} - \mathbf{A}\mathbf{a}) = \|\mathbf{Y} - \mathbf{A}\mathbf{a}\|^2.$$

令  $\boldsymbol{\delta} = \mathbf{Y} - \mathbf{A}\mathbf{a}$ , 定义损失函数

$$\begin{aligned} J &= \sum_{i=1}^n e_i^2 = (\mathbf{Y} - \mathbf{A}\mathbf{a})^\top (\mathbf{Y} - \mathbf{A}\mathbf{a}) = \boldsymbol{\delta}^\top \boldsymbol{\delta} \\ &= \mathbf{Y}^\top \mathbf{Y} - \mathbf{a}^\top \mathbf{A}^\top \mathbf{Y} - \mathbf{Y}^\top \mathbf{A}\mathbf{a} + \mathbf{a}^\top \mathbf{A}^\top \mathbf{A}\mathbf{a}, \end{aligned}$$

为确定  $\mathbf{a}$ , 求  $J$  对  $\mathbf{a}$  的偏导并令结果等于零, 有

$$\frac{\partial J}{\partial \mathbf{a}} = -\mathbf{A}^\top \mathbf{Y} - \mathbf{A}^\top \mathbf{Y} + (\mathbf{A}^\top \mathbf{A} + \mathbf{A}^\top \mathbf{A})\mathbf{a} = \mathbf{0},$$

$$\mathbf{A}^\top \mathbf{A}\mathbf{a} = \mathbf{A}^\top \mathbf{Y},$$

当  $m \times n$  矩阵  $\mathbf{A}$  列满秩时,  $\mathbf{A}^\top \mathbf{A}$  非奇异, 可得

$$\mathbf{a} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{Y}.$$

此即最小二乘法的参数估计结果.

现在讨论最小二乘法的最优性.

当给定观测值向量  $\mathbf{Y}$  时, 按最小二乘法可得相应的预报值(拟合值)向量  $\hat{\mathbf{Y}}$ , 即

$$\hat{\mathbf{Y}} = \mathbf{A}\mathbf{a} = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{Y} = \mathbf{A}\mathbf{L}^{-1} \mathbf{A}^\top \mathbf{Y} \quad (\mathbf{L} = \mathbf{A}^\top \mathbf{A}),$$

称拟合误差  $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$  为残差向量, 并记  $\mathbf{P} = \mathbf{A}\mathbf{L}^{-1} \mathbf{A}^\top$ , 则有

$$\mathbf{P}^\top = \mathbf{P}, \quad \mathbf{P}^2 = \mathbf{P},$$

$$(\mathbf{E} - \mathbf{P})^\top = \mathbf{E} - \mathbf{P}, \quad (\mathbf{E} - \mathbf{P})^2 = \mathbf{E} - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{E} - \mathbf{P},$$

即  $P$  和  $E - P$  都是幂等矩阵.

$$\hat{Y} = Aa = AL^{-1}A^T Y = PY, \quad (3.5.1)$$

$$e = Y - \hat{Y} = Y - PY = (E - P)Y, \quad (3.5.2)$$

$$PA = AL^{-1}A^T A = A, \quad (E - P)A = 0,$$

$$(\hat{Y}, e) = (PY)^T (E - P)Y = Y^T P(E - P)Y = 0, \quad (3.5.3)$$

即  $\hat{Y} \perp e$ , 由式(3.5.1)~式(3.5.3), 向量之间的关系如图 3.5.2 所示.

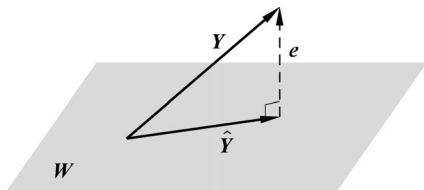


图 3.5.2 向量关系示意图

由图 3.5.2 可见, 预报值  $\hat{Y}$  为观测值  $Y$  在子空间  $W$  上的投影, 子空间  $W$  为系数矩阵  $A$  的列向量生成的子空间, 残差向量  $e = Y - \hat{Y}$  为观测值和预报值之间的误差, 从几何图形上看, 当  $\hat{Y} \perp e$  时, 残差向量  $e$  最小, 而由最小二乘法求得的系数  $a$  恰好使  $\hat{Y} \perp e$  这一条件成立.

综上, 由最小二乘法求得的估计值  $\hat{Y}$  是  $Y$  在子空间  $W$  上的最佳逼近元, 因此目标函数③的方法即最小二乘法是最优的.

## 本章小结

本章介绍了典型矩阵与变换, 包括酉矩阵与酉变换、正交矩阵与正交变换、幂等矩阵与投影变换、次酉矩阵与正交投影、对称矩阵与对称变换、Hermite 矩阵与 Hermite 变换、伴随变换、正规矩阵与正规变换等, 此外给出了应用实例.

学习完本章内容后, 应能达到如下基本要求:

- (1) 掌握各类矩阵和变换的定义与性质、矩阵与变换之间的关系;
- (2) 能进行矩阵、变换的相关计算和证明.

表 3.1 总结了本章内容.

表 3.1 典型矩阵与变换总结

	矩阵	变换	特征值	特征向量	对角化
正规矩阵	酉矩阵 $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H = \mathbf{E}$	酉变换 $(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$	模等于 1	不同特征值对应的特征向量互相正交	可以
	正交矩阵 $\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T = \mathbf{E}$	正交变换 $(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$	$\pm 1$		
	Hermite 矩阵 $\mathbf{A}^H = \mathbf{A}$	Hermite 变换 $(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta))$	实数		
	对称矩阵 $\mathbf{A}^T = \mathbf{A}$	对称变换 $(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta))$	实数	特征值之和不为零的两个特征向量互相正交	
	反 Hermite 矩阵 $\mathbf{A}^H = -\mathbf{A}$	反 Hermite 变换 $(\sigma(\alpha), \beta) = -(\alpha, \sigma(\beta))$	0 或虚数		
	反对称矩阵 $\mathbf{A}^T = -\mathbf{A}$	反对称变换 $(\sigma(\alpha), \beta) = -(\alpha, \sigma(\beta))$	0 或虚数		
幂等矩阵	正规矩阵 $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H$	正规变换 $\sigma^H \sigma = \sigma \sigma^H$	$\mathbf{A}^H$ 和 $\mathbf{A}$ 的特征值互为共轭	$n$ 阶矩阵 $\mathbf{A}$ 有 $n$ 个线性无关的特征向量	
	幂等矩阵 $\mathbf{A}^2 = \mathbf{A}$	投影变换 $\sigma(x+y) = x$	0 或 1	无明显规律	
	$\mathbf{A}^2 = \mathbf{A} = \mathbf{A}^H$	正交投影变换 $\sigma(x+y) = x$ , 其中 $x \perp y$	0 或 1	不同特征值对应的特征向量互相正交	

注：本表的部分内容超出了本书范围，有兴趣的读者可自行查阅相关资料。

### 习题 3

3-1 若  $\mathbf{A} \in U^{n \times n}$ ，试证明：

- (1)  $|\det \mathbf{A}| = 1$ ;
- (2)  $\mathbf{A}^T \in U^{n \times n}$ ;
- (3)  $\mathbf{A}^2 \in U^{n \times n}$ .

**3-2** 试写出:

- (1)  $\mathbb{R}^2$  空间中将向量  $\alpha$  顺时针旋转  $\frac{\pi}{2}$  的旋转变换的矩阵;
- (2)  $\mathbb{R}^2$  空间中将向量  $\alpha$  逆时针旋转  $\frac{\pi}{2}$  的旋转变换的矩阵.

**3-3** 试写出:

- (1) 三维空间中的向量绕  $y$  轴旋转  $\theta$  角的旋转变换的矩阵  $\mathbf{R}_y(\theta)$ ;
- (2) 三维空间中的向量绕  $z$  轴旋转  $\theta$  角的旋转变换的矩阵  $\mathbf{R}_z(\theta)$ .

**3-4** 试写出:

- (1)  $\mathbb{R}^2$  空间中将向量  $\overrightarrow{OP}$  正交投影到  $Ox$  轴上的投影变换的矩阵, 其中  $P$  点的坐标是  $(x, y)$ , 投影点  $P'$  的坐标是  $(x', y')$ ;
- (2)  $\mathbb{R}^2$  空间中将向量  $\overrightarrow{OP}$  正交投影到  $Oy$  轴上的投影变换的矩阵, 其中  $P$  点的坐标是  $(x, y)$ , 投影点  $P'$  的坐标是  $(x', y')$ ;
- (3)  $\mathbb{R}^3$  空间中将向量  $\overrightarrow{OP}$  正交投影到  $Oz$  轴上的投影变换的矩阵, 其中  $P$  点的坐标是  $(x, y, z)$ , 投影点  $P'$  的坐标是  $(x', y', z')$ .

**3-5** 求下列线性变换的矩阵:

- (1)  $\mathbb{R}^2$  空间中将向量  $\alpha$  关于  $y$  轴对称;
- (2)  $\mathbb{R}^2$  空间中将向量  $\alpha$  关于直线  $y=x$  对称;
- (3)  $\mathbb{R}^2$  空间中将向量  $\alpha$  关于直线  $y=-x$  对称.

**3-6** 求正交矩阵  $\mathbf{Q}$ , 使  $\mathbf{Q}^T \mathbf{A} \mathbf{Q}$  为对角阵, 已知

$$(1) \mathbf{A} = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{bmatrix};$$

$$(2) \mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 2 \\ -1 & 0 & 1 & 1 \end{bmatrix}.$$

**3-7** 已知实对称矩阵  $\mathbf{A} = \begin{bmatrix} 3 & -2 & -4 \\ -2 & 6 & -2 \\ -4 & -2 & 3 \end{bmatrix}$ , 求正交矩阵  $\mathbf{Q}$ , 使  $\mathbf{Q}^T \mathbf{A} \mathbf{Q}$  为对

角阵.

**3-8** 填空:

(1) 设  $\sigma$  是欧氏空间  $V$  的一个反对称变换, 则其伴随变换  $\sigma^H =$  \_\_\_\_\_.

(2) 设  $\sigma$  是酉空间  $V$  的一个反 Hermite 变换, 则其伴随变换  $\sigma^H =$  \_\_\_\_\_.

**3-9** 已知  $A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$ , 求矩阵  $Q$ , 使  $Q^T A Q$  为对角阵.

**3-10** 验证  $A = \begin{bmatrix} 0 & -1 & i \\ 1 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}$  是否是正规矩阵, 并求酉矩阵  $U$ , 使  $U^H A U$  为对

角阵.

**3-11** 验证  $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$  是否是正规矩阵, 并求酉矩阵  $U$ , 使  $U^H A U$  为对

角阵.

**3-12** 设  $A, B$  均是 Hermite 矩阵, 试证:  $A$  与  $B$  酉相似的充要条件是  $A$  与  $B$  的特征值相同.

**3-13** 设  $A, B$  均是 Hermite 矩阵, 试证:  $AB$  也是 Hermite 矩阵的充要条件是  $AB = BA$ .

**3-14** 设  $A, B$  均是正规矩阵, 试证:  $A$  与  $B$  酉相似的充要条件是  $A$  与  $B$  的特征值相同.

**3-15** 设  $A, B$  均是正规矩阵, 且  $AB = BA$ , 试证:  $AB$  与  $BA$  均是正规矩阵.

**3-16** 试证: 任一  $n$  阶矩阵都可以表示为一个 Hermite 矩阵与一个反 Hermite 矩阵之和的形式.