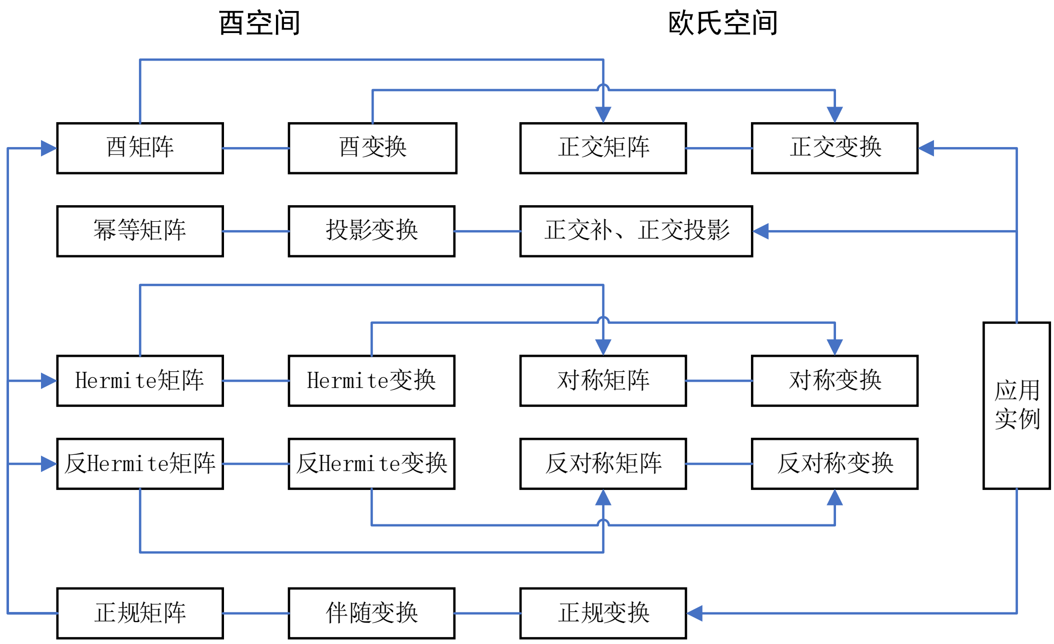
第 3 章　典型矩阵与变换

第 2 章介绍了线性变换及其矩阵表示，线性变换的矩阵表示可简称为线性变换的矩阵，本章开始使用这一简称.

线性空间中向量的线性变换结果等于基的变换结果与该向量的坐标向量的点积，也可以说，线性变换实质上是基的变换. 在数学形式上，对向量进行线性变换就是用矩阵与之相乘. 由于矩阵本身的属性不同，有必要对具体的矩阵与变换进行详细分析，并研究其理论意义和应用价值.

本章研究的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、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.

本章的知识网络框图：



3.1　酉矩阵与酉变换、正交矩阵与正交变换

3.1.1　酉矩阵和正交矩阵

定义 3.1.1　设复矩阵$\boldsymbol{A}=(a_{ij})_{m\times n}\in\mathbb{C}^{m\times n}$，$\overline{\boldsymbol{A}}$是以$\boldsymbol{A}$的元素的共轭复数为元素的矩阵，即$\overline{\boldsymbol{A}}=(\overline{a_{ij}})_{m\times n}$，令$\boldsymbol{A}^{\mathrm{H}}=(\overline{\boldsymbol{A}})^{\mathrm{T}}$，称$\boldsymbol{A}^{\mathrm{H}}$为$\boldsymbol{A}$的**复共轭转置矩阵**.

不难证明复共轭转置具有如下性质:

(1) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}+\boldsymbol{B}^{\mathrm{H}}$;

(2) $(k\boldsymbol{A})^{\mathrm{H}}=\bar{k}\boldsymbol{A}^{\mathrm{H}}$;

(3) $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}$;

(4) $(\boldsymbol{A}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{A}$;

(5) $(\boldsymbol{A}^{\mathrm{H}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{H}}$，当$\boldsymbol{A}$可逆时.

定义 3.1.2　设$\boldsymbol{E}$是单位矩阵，若 n 阶复矩阵$\boldsymbol{A}$满足

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{E},$$

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

酉是英文 Unitary 的音译，酉矩阵也称为**幺正矩阵**.

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

(1) $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{H}}\in U^{n\times n}$;

(2) $|\det\boldsymbol{A}|=1$，即酉矩阵的行列式的模为 1;

(3) $\boldsymbol{A}^{\mathrm{T}}\in U^{n\times n}$;

(4) $\boldsymbol{A}\boldsymbol{B},\boldsymbol{B}\boldsymbol{A}\in U^{n\times n}$;

(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$\boldsymbol{\alpha}\in\mathbb{C}^n$，且$\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\alpha}=1$，若

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

试证明$\boldsymbol{H}$是酉矩阵.

证:　由复共轭转置的性质，有

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

故$\boldsymbol{H}$是酉矩阵. ■

定义 3.1.3 设$\boldsymbol{E}$是单位阵，若n阶实矩阵$\boldsymbol{A}$满足

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{E},$$

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

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

(1) $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}\in E^{n\times n}$.

(2) $\det\boldsymbol{A}=\pm 1$，即正交矩阵的行列式是+1 或−1.

(3) $\boldsymbol{AB},\boldsymbol{BA}\in E^{n\times n}$.

(4) 正交矩阵的特征值是± 1，特征向量互相正交.

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

证: 设$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n]$，则

$$\boldsymbol{A}^{\mathrm{H}}=\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{H}}\\ \boldsymbol{\alpha}_2^{\mathrm{H}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{H}}\end{bmatrix},$$

若$\boldsymbol{A}$是酉矩阵，则$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{E}$，于是

$$\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{H}}\\ \boldsymbol{\alpha}_2^{\mathrm{H}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{H}}\end{bmatrix}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n]=\boldsymbol{E},$$

此即

$$\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{H}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{H}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{H}}\boldsymbol{\alpha}_n\\ \boldsymbol{\alpha}_2^{\mathrm{H}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{H}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^{\mathrm{H}}\boldsymbol{\alpha}_n\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{H}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^{\mathrm{H}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^{\mathrm{H}}\boldsymbol{\alpha}_n\end{bmatrix}=\begin{bmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{bmatrix},$$

比较上式两端，得

$$\boldsymbol{\alpha}_i^{\mathrm{H}}\boldsymbol{\alpha}_j=\delta_{ij}\ (i,j=1,2,\cdots,n),$$

所以列向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是标准正交向量组.

反之，若列向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是标准正交向量组，则有

$$\boldsymbol{\alpha}_i^{\mathrm{H}}\boldsymbol{\alpha}_j=\delta_{ij}\quad(i,j=1,2,\cdots,n),$$

于是

$$\begin{bmatrix} \boldsymbol{\alpha}_1^{\mathrm{H}} \\ \boldsymbol{\alpha}_2^{\mathrm{H}} \\ \vdots \\ \boldsymbol{\alpha}_n^{\mathrm{H}} \end{bmatrix} [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n] = \boldsymbol{E},$$

此即

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{E},$$

因此, $\boldsymbol{A}$是酉矩阵. 用类似方法可证明$\boldsymbol{A}$的行向量组是标准正交向量组. ■

定义 3.1.4 若$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$为 n 维标准正交列向量组($r < n$), 则称$n \times r$矩阵$\boldsymbol{U}_1 = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r)$为**次酉矩阵**, 记为$\boldsymbol{U}_1 \in U_r^{n\times r}$.

定理 3.1.2 次酉矩阵$\boldsymbol{U}_1 \in U_r^{n\times r}$的充要条件为$\boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{U}_1 = \boldsymbol{E}_r$.

3.1.2 酉变换和正交变换

定义 3.1.5 设V是n维酉空间, σ是V的线性变换, 若$\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 都有

$$(\sigma(\boldsymbol{\alpha}), \sigma(\boldsymbol{\beta})) = (\boldsymbol{\alpha}, \boldsymbol{\beta}),$$

则称σ是V的**酉变换**.

定义 3.1.6 设V是n维 Euclid 空间, σ是V的线性变换, 若$\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 都有

$$(\sigma(\boldsymbol{\alpha}), \sigma(\boldsymbol{\beta})) = (\boldsymbol{\alpha}, \boldsymbol{\beta}),$$

则称σ是V的**正交变换**.

定理 3.1.3 设σ是酉空间(Euclid 空间)V的线性变换, 则下列命题等价:

(1) σ是酉变换(正交变换);

(2) $\|\sigma(\boldsymbol{\alpha})\| = \|\boldsymbol{\alpha}\|, \quad \forall \boldsymbol{\alpha} \in V$;

(3) σ将V的标准正交基变换到标准正交基;

(4) 酉变换(正交变换)在标准正交基下的矩阵表示是酉矩阵(正交矩阵).

证: (1) $\Rightarrow$ (2) 显然.

(2) $\Rightarrow$ (1) 由(2), 有

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

$$(\sigma(\boldsymbol{\alpha}+\mathrm{i}\boldsymbol{\beta}), \sigma(\boldsymbol{\alpha}+\mathrm{i}\boldsymbol{\beta})) = (\boldsymbol{\alpha}+\mathrm{i}\boldsymbol{\beta}, \boldsymbol{\alpha}+\mathrm{i}\boldsymbol{\beta}),$$

由于σ是线性变换, 根据内积性质展开以上两式, 得

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

$$(\sigma(\boldsymbol{\alpha}), \sigma(\boldsymbol{\beta})) - (\sigma(\boldsymbol{\beta}), \sigma(\boldsymbol{\alpha})) = (\boldsymbol{\alpha}, \boldsymbol{\beta}) - (\boldsymbol{\beta}, \boldsymbol{\alpha}),$$

将以上两式相加，得

$$(\sigma(\boldsymbol{\alpha}),\sigma(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta}),$$

此即酉变换的定义式.

(1) ⇒ (3)　设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是$V$的标准正交基，故

$$\left(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j\right)=\delta_{ij}\quad(i,j=1,2,\cdots,n),$$

若σ是酉变换，则有

$$\left(\sigma(\boldsymbol{\alpha}_i),\sigma\left(\boldsymbol{\alpha}_j\right)\right)=\left(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j\right)=\delta_{ij},$$

故$\sigma(\boldsymbol{\alpha}_1),\sigma(\boldsymbol{\alpha}_2),\cdots,\sigma(\boldsymbol{\alpha}_n)$仍是$V$的标准正交基.

(3) ⇒ (1)　设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$与$\sigma(\boldsymbol{\alpha}_1),\sigma(\boldsymbol{\alpha}_2),\cdots,\sigma(\boldsymbol{\alpha}_n)$都是$V$的标准正交基，$\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V$且

$$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n,$$

$$\boldsymbol{\beta}=y_1\boldsymbol{\alpha}_1+y_2\boldsymbol{\alpha}_2+\cdots+y_n\boldsymbol{\alpha}_n,$$

则

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

$$\sigma(\boldsymbol{\beta})=y_1\sigma(\boldsymbol{\alpha}_1)+y_2\sigma(\boldsymbol{\alpha}_2)+\cdots+y_n\sigma(\boldsymbol{\alpha}_n),$$

由内积运算性质，有

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

此即酉变换的定义式.

(3) ⇒ (4)　设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$与$\sigma(\boldsymbol{\alpha}_1),\sigma(\boldsymbol{\alpha}_2),\cdots,\sigma(\boldsymbol{\alpha}_n)$是$V$的标准正交基，$n$阶矩阵$\boldsymbol{A}=\left(a_{ij}\right)$是$\sigma$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的矩阵表示，则

$$\left(\sigma(\boldsymbol{\alpha}_1),\sigma(\boldsymbol{\alpha}_2),\cdots,\sigma(\boldsymbol{\alpha}_n)\right)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A},$$

且$\forall i,j\ \ (i,j=1,2,\cdots,n)$

$$\sigma(\boldsymbol{\alpha}_i)=a_{1i}\boldsymbol{\alpha}_1+a_{2i}\boldsymbol{\alpha}_2+\cdots+a_{ni}\boldsymbol{\alpha}_n,$$

$$\sigma\left(\boldsymbol{\alpha}_j\right)=a_{1j}\boldsymbol{\alpha}_1+a_{2j}\boldsymbol{\alpha}_2+\cdots+a_{nj}\boldsymbol{\alpha}_n,$$

于是

$$
\begin{aligned}
\delta_{ij} &= \left(\sigma(\boldsymbol{\alpha}_i), \sigma(\boldsymbol{\alpha}_j)\right) = \left(\sum_{k=1}^{n} a_{ki}\boldsymbol{\alpha}_k, \sum_{h=1}^{n} a_{hj}\boldsymbol{\alpha}_h\right) \\
&= \sum_{k=1}^{n} \sum_{h=1}^{n} a_{ki}\bar{a}_{hj}(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_h) \\
&= \sum_{k=1}^{n} \sum_{h=1}^{n} a_{ki}\bar{a}_{hj}\delta_{kh} \\
&= \sum_{k=1}^{n} a_{ki}\bar{a}_{kj}.
\end{aligned}
$$

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

(4) ⇒ (3)　显然. ■

根据命题(2)，酉变换也可称为**等距变换**．这是因为

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

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

3.1.3　酉变换、正交变换实例

1. 酉矩阵(正交矩阵)乘以向量是酉变换(正交变换)

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

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

其中，n 阶方阵$\boldsymbol{A}$为酉矩阵，则

$$\left(T(\boldsymbol{X}_1), T(\boldsymbol{X}_2)\right) = (\boldsymbol{A}\boldsymbol{X}_1, \boldsymbol{A}\boldsymbol{X}_2) = \boldsymbol{X}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{X}_1 = \boldsymbol{X}_2^{\mathrm{H}}\boldsymbol{X}_1 = (\boldsymbol{X}_1, \boldsymbol{X}_2),$$

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$\boldsymbol{A}$为正交矩阵，令$T(\boldsymbol{X}) = \boldsymbol{A}\boldsymbol{X}$，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φ. 将向量$\overrightarrow{OP}$逆时针旋转θ角，变为向量$\overrightarrow{OP'}$.

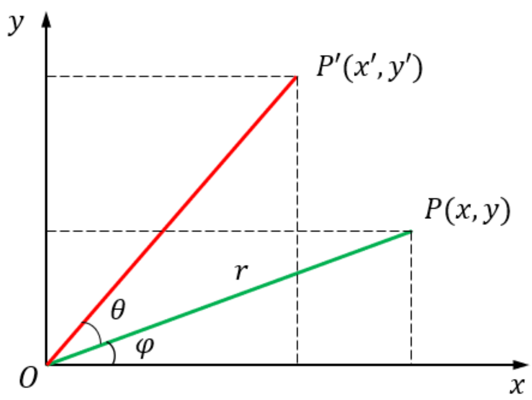


图 3.1.1　向量的旋转

若点P的坐标是(x,y)，点P'的坐标是(x',y')，则有

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi \end{cases},$$

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

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

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

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θ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

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

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$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，若

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

则称$\boldsymbol{A}$是**幂等矩阵**.

例如，矩阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{M}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\in\mathbb{C}^{n\times n},\qquad \boldsymbol{M}\in\mathbb{C}^{r\times(n-r)}$$

是幂等矩阵.

用符号$\mathbb{F}_r^{m\times n}$，$\mathbb{R}_r^{m\times n}$，$\mathbb{C}_r^{m\times n}$分别表示元素在数域$\mathbb{F}$，实数域$\mathbb{R}$，复数域$\mathbb{C}$中的秩为 r 的$m\times n$矩阵集合.

定理 3. 2. 1　矩阵$\boldsymbol{A}\in\mathbb{C}_r^{n\times n}$是幂等矩阵的充要条件是存在$\boldsymbol{P}\in\mathbb{C}_n^{n\times n}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}\boldsymbol{E}_r & \\ & \boldsymbol{0}\end{bmatrix}$$

且$\boldsymbol{A}$的特征值是 0 或 1.

定理 3. 2. 2　设$\boldsymbol{A}\in\mathbb{C}^{n\times n}$且$\boldsymbol{A}^2=\boldsymbol{A}$，则

(1) $\boldsymbol{A}^{\mathrm{T}},\boldsymbol{A}^{\mathrm{H}},\boldsymbol{E}-\boldsymbol{A},\boldsymbol{E}-\boldsymbol{A}^{\mathrm{T}},\boldsymbol{E}-\boldsymbol{A}^{\mathrm{H}}$是幂等矩阵;

(2) $\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})=(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{A}=\boldsymbol{0}$;

(3) $N(\boldsymbol{A})=R(\boldsymbol{E}-\boldsymbol{A})$，$R(\boldsymbol{A})=N(\boldsymbol{E}-\boldsymbol{A})$;

(4) $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\Leftrightarrow\boldsymbol{x}\in R(\boldsymbol{E}-\boldsymbol{A})$，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}\Leftrightarrow\boldsymbol{x}\in R(\boldsymbol{A})$;

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

证：(1)和(2)　由定义$\boldsymbol{A}^2=\boldsymbol{A}$立即可证.

(3)　设$\boldsymbol{x}\in N(\boldsymbol{A})$，则$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，于是

$$(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{x}-\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x},$$

这说明$\boldsymbol{x}\in R(\boldsymbol{E}-\boldsymbol{A})$，故$N(\boldsymbol{A})\subseteq R(\boldsymbol{E}-\boldsymbol{A})$.

另一方面，若$\boldsymbol{x}\in R(\boldsymbol{E}-\boldsymbol{A})$，即存在$\boldsymbol{y}\in\mathbb{C}^n$，使得$(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{y}=\boldsymbol{x}$，故$\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$，由(2)的结果$\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})=\boldsymbol{0}$可得$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，于是$\boldsymbol{x}\in N(\boldsymbol{A})$，故$R(\boldsymbol{E}-\boldsymbol{A})\subseteq N(\boldsymbol{A})$，所以$N(\boldsymbol{A})=R(\boldsymbol{E}-\boldsymbol{A})$.

由于$\boldsymbol{E}-\boldsymbol{A}$也是幂等矩阵，于是由所得等式有

$$N(\boldsymbol{E}-\boldsymbol{A})=R\big(\boldsymbol{E}-(\boldsymbol{E}-\boldsymbol{A})\big)=R(\boldsymbol{A}).$$

(4)　由(3)的第一式知

$$\boldsymbol{x}\in N(\boldsymbol{A})\Leftrightarrow\boldsymbol{x}\in R(\boldsymbol{E}-\boldsymbol{A}),$$

即$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\Leftrightarrow\boldsymbol{x}\in R(\boldsymbol{E}-\boldsymbol{A})$.

由(3)的第二式知

$$\boldsymbol{x}\in N(\boldsymbol{E}-\boldsymbol{A})\Leftrightarrow\boldsymbol{x}\in R(\boldsymbol{A}),$$

即$(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$，也即$\boldsymbol{Ax}=\boldsymbol{x}\Leftrightarrow \boldsymbol{x}\in R(\boldsymbol{A})$.

(5) 先证$\mathbb{C}^n=R(\boldsymbol{A})+N(\boldsymbol{A})$.

对$\mathbb{C}^n$中任一向量$\boldsymbol{x}$都有

$$\boldsymbol{Ax}\in R(\boldsymbol{A})\subseteq \mathbb{C}^n.$$

令$\boldsymbol{x}=\boldsymbol{Ax}+\boldsymbol{\xi}$，其中$\boldsymbol{\xi}\in\mathbb{C}^n$，则$\boldsymbol{Ax}=\boldsymbol{A}^2\boldsymbol{x}+\boldsymbol{A\xi}=\boldsymbol{Ax}+\boldsymbol{A\xi}$，故$\boldsymbol{A\xi}=\boldsymbol{0}$，于是$\boldsymbol{\xi}\in N(\boldsymbol{A})$，所以

$$\mathbb{C}^n=R(\boldsymbol{A})+N(\boldsymbol{A}).$$

取$\boldsymbol{x}\in R(\boldsymbol{A})\cap N(\boldsymbol{A})$，由$\boldsymbol{x}\in R(\boldsymbol{A})$，故由(4)知$\boldsymbol{Ax}=\boldsymbol{x}$，再由$\boldsymbol{x}\in N(\boldsymbol{A})$，故$\boldsymbol{Ax}=\boldsymbol{0}$，综合两式知$\boldsymbol{x}=\boldsymbol{0}$. 于是

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

3.2.2 投影变换

投影分为斜投影和正交投影两类.

在图 3.2.1 所示的斜投影中，若向量$\boldsymbol{\alpha}$沿$\boldsymbol{v}$至$\boldsymbol{u}$的投影为$\boldsymbol{x}$，沿$\boldsymbol{u}$至$\boldsymbol{v}$的投影为$\boldsymbol{y}$，则有$\boldsymbol{\alpha}=\boldsymbol{x}+\boldsymbol{y}$，令投影变换为

$$\sigma：\sigma(\boldsymbol{\alpha})=\sigma(\boldsymbol{x}+\boldsymbol{y})=\boldsymbol{x},$$

显然，这样的投影变换是线性变换，且$\sigma^2=\sigma$.

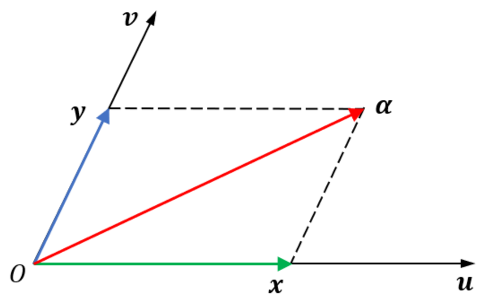


图 3.2.1 斜投影

在工程应用中，常需要提取有用信号，同时抑制干扰和噪声，投影是解决这类问题的重要数学工具，为刻画这类数学方法，将上述几何上的投影变换推广到一般的酉空间或欧氏空间，如图 3.2.2 所示.

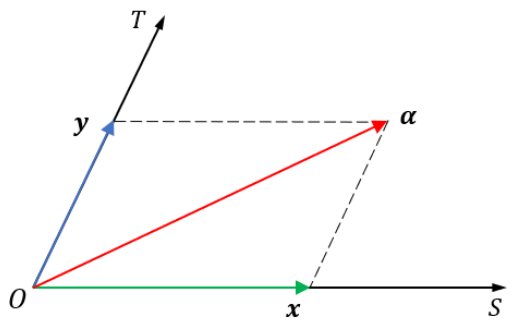


图 3.2.2　酉空间(欧氏空间)的投影

定义 3.2.2　设S,T是n维酉空间V的子空间，且$V=S\oplus T$，若V中任一向量$\boldsymbol{\alpha}$均可唯一地表示为$\boldsymbol{\alpha}=\boldsymbol{x}+\boldsymbol{y}$，$\boldsymbol{x}\in S$，$\boldsymbol{y}\in T$，则称$\boldsymbol{x}$是$\boldsymbol{\alpha}$沿$T$至$S$的投影，$\boldsymbol{y}$是$\boldsymbol{\alpha}$沿$S$至$T$的投影. 由此确定的 n 维酉空间V至子空间S的线性映射$\tau_{S,T}(\boldsymbol{\alpha})=\boldsymbol{x}$称为$n$维酉空间$V$沿$T$至$S$的**投影映射**. 由此确定的 n 维酉空间V的线性变换$\tau_{S,T}(\boldsymbol{\alpha})=\boldsymbol{x}$称为$n$维酉空间$V$沿$T$至$S$的**投影变换**.

显然，$\tau_{S,T}$限制在S上就是恒等变换.

定理 3.2.3　设τ是n维酉空间（欧氏空间）V的线性变换，则下列命题等价:

(1) τ是V上的投影变换;

(2) $\dim\big(R(\tau)\cap N(\tau)\big)=0$;

(3) $V=R(\tau)\oplus N(\tau)$.

证： (1) $\Rightarrow$ (2)　$\forall\boldsymbol{\alpha}\in R(\tau)\cap N(\tau)$，由于$\boldsymbol{\alpha}\in R(\tau)$，故$\tau(\boldsymbol{\alpha})=\boldsymbol{\alpha}$. 又由于$\boldsymbol{\alpha}\in N(\tau)$，故$\tau(\boldsymbol{\alpha})=\boldsymbol{0}$. 综合两式得$\boldsymbol{\alpha}=\boldsymbol{0}$，此即

$$\dim\big(R(\tau)\cap N(\tau)\big)=0.$$

（2）$\Rightarrow$（3）　由维数公式可得

$$\dim\big(R(\tau)+N(\tau)\big)=\dim R(\tau)+\dim N(\tau)-\dim\big(R(\tau)\cap N(\tau)\big)$$

$$=\dim R(\tau)+\dim N(\tau),$$

又由定理 2.1.4 可知$\dim R(\tau)+\dim N(\tau)=n$，即

$$\dim\big(R(\tau)+N(\tau)\big)=n,$$

此即$V=R(\tau)+N(\tau)$，又$R(\tau)\cap N(\tau)=\{\boldsymbol{0}\}$，所以

$$V=R(\tau)\oplus N(\tau).$$

（3）$\Rightarrow$（1）　因为$V=R(\tau)\oplus N(\tau)$，故$\forall\boldsymbol{\alpha}\in V$有$\boldsymbol{\alpha}_2=\boldsymbol{\alpha}-\tau(\boldsymbol{\alpha})$，$\boldsymbol{\alpha}=\boldsymbol{\alpha}_2+$

$\tau(\boldsymbol{\alpha})$，这里$\tau(\boldsymbol{\alpha}) \in R(\tau)$，故$\boldsymbol{\alpha}_2 \in N(\tau)$，即$\tau(\boldsymbol{\alpha}_2) = \mathbf{0}$. 所以，$\tau$是$V$至$R(\tau)$的投影变换. ■

定理 3.2.4 设τ是n维酉空间（欧氏空间）的线性变换，则下列命题等价:

(1) τ是V上的投影变换;

(2) $\tau^2 = \tau$;

(3) τ的矩阵表示$\boldsymbol{A}$满足$\boldsymbol{A}^2 = \boldsymbol{A}$.

证:（1）⇒（2） $\forall \boldsymbol{\alpha} \in V$，有

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2,$$

其中$\boldsymbol{\alpha}_1 \in R(\tau), \boldsymbol{\alpha}_2 \in N(\tau)$，于是$\tau(\boldsymbol{\alpha}_1) = \boldsymbol{\alpha}_1, \tau(\boldsymbol{\alpha}_2) = \mathbf{0}$，因此

$$\tau(\boldsymbol{\alpha}) = \tau(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2) = \tau(\boldsymbol{\alpha}_1) + \tau(\boldsymbol{\alpha}_2) = \tau(\boldsymbol{\alpha}_1) = \boldsymbol{\alpha}_1,$$

$$\tau^2(\boldsymbol{\alpha}) = \tau(\boldsymbol{\alpha}_1) = \boldsymbol{\alpha}_1 = \tau(\boldsymbol{\alpha}).$$

根据$\boldsymbol{\alpha}$的任意性得$\tau^2 = \tau$.

（2）⇒（1） $\forall \boldsymbol{\alpha} \in V$，$\tau(\boldsymbol{\alpha}) \in R(\tau)$.

设$\boldsymbol{\alpha} = \tau(\boldsymbol{\alpha}) + \boldsymbol{\alpha}_2$，则

$$\tau(\boldsymbol{\alpha}) = \tau^2(\boldsymbol{\alpha}) + \tau(\boldsymbol{\alpha}_2),$$

将$\tau = \tau^2$代入上式，得$\tau(\boldsymbol{\alpha}_2) = \mathbf{0}$，于是$\boldsymbol{\alpha}_2 \in N(\tau)$，所以$\tau$是$V$到$R(\tau)$的投影变换.

（2)⇒(3） 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$是$V$的一组基，$\boldsymbol{A}$是$\tau$在该组基下的矩阵表示，于是

$$\tau(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A},$$

$$\tau^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = \tau(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}^2.$$

若$\tau = \tau^2$，有

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}^2,$$

由于$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$线性无关，可得$\boldsymbol{A} = \boldsymbol{A}^2$.

（3）⇒（2） 若

$$\tau(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A},$$

$$\tau^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}^2,$$

当$\boldsymbol{A} = \boldsymbol{A}^2$，由于在同一组基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$下，线性变换与其矩阵一一对应，因此$\tau = \tau^2$. ■

3.2.3 正交补与正交投影

为了把几何学中的正交投影（垂直投影）的概念推广到酉空间（欧氏空间），需要引进子空间的正交与正交补的概念.

定义 3.2.3 设 S, T 是 n 维酉空间 V 的子空间，若对于任意$\boldsymbol{x} \in S, \boldsymbol{y} \in T$，都有$(\boldsymbol{x}, \boldsymbol{y}) = 0$，则称子空间 S 与 T 是**正交**的，记为$S \perp T$.

例 3.2.1 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4$是 V 的标准正交基，则$S = \mathrm{span}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$与$T = \mathrm{span}\{\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$是正交的.

定理 3.2.5 设 S, T 是 n 维酉空间 V 的两个正交子空间，则

(1) $S \cap T = \{\boldsymbol{0}\}$;

(2) $\dim(S + T) = \dim S + \dim T$.

证：(1) 设$\boldsymbol{x} \in S \cap T$，由$\boldsymbol{x} \in S$，故对任意$\boldsymbol{y} \in T$，都有$(\boldsymbol{x}, \boldsymbol{y}) = 0$，又因$\boldsymbol{x} \in T$，因此取$\boldsymbol{y} = \boldsymbol{x}$，则$(\boldsymbol{x}, \boldsymbol{x}) = 0$，于是$\boldsymbol{x} = \boldsymbol{0}$. 根据$\boldsymbol{x}$的任意性可得$S \cap T = \{\boldsymbol{0}\}$.

(2) 根据维数公式，并由(1)立刻可得.

定义 3.2.4 若子空间 S, T 正交，则$S + T$称为 S 与 T 的**正交和**，记为 $S \mathbin{\textcircled{\perp}} T$.

定理 3.2.6 设$\boldsymbol{A} \in \mathbb{C}^{m\times n}$(或$\mathbb{R}^{m\times n}$)，则

(1) $N(\boldsymbol{A}) \mathbin{\textcircled{\perp}} R(\boldsymbol{A}^{\mathrm{H}}) = \mathbb{C}^n$(或$\mathbb{R}^n$);

(2) $R(\boldsymbol{A}) \mathbin{\textcircled{\perp}} N(\boldsymbol{A}^{\mathrm{H}}) = \mathbb{C}^m$（或$\mathbb{R}^m$）.

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

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

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

又

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

因此

$$N(\boldsymbol{A}) \mathbin{\textcircled{\perp}} R(\boldsymbol{A}^{\mathrm{H}}) = \mathbb{C}^n(\text{或}\mathbb{R}^n).$$

类似可得(2). ■

定义 3.2.5 设 n 维酉空间 V 的子空间 S, T 满足$S \mathbin{\textcircled{\perp}} T = V$，则称 S 为 T 的

正交补，记为$T_{\perp}$，或

$$S = T_{\perp} = \{\boldsymbol{\alpha} | (\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0, \forall \boldsymbol{\beta} \in T\}.$$

显然，若 S 是 T 的正交补，则 T 也是 S 的正交补.

定理 3.2.7　设 T 是 n 维酉空间 V 的子空间，则存在唯一的子空间 S，使得

$$S \oplus T = V.$$

例 3.2.2　已知$\boldsymbol{\alpha}_1 = (1,0,1,1)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (0,1,1,2)^{\mathrm{T}}, T = \mathrm{span}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$，求 T 的正交补.

分析：可利用定理 3.2.6 的结论求解.

解：　取$\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$，则

$$\boldsymbol{A}^{\mathrm{H}} = \begin{bmatrix} \boldsymbol{\alpha}_1^{\mathrm{H}} \\ \boldsymbol{\alpha}_2^{\mathrm{H}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix},$$

可求得线性方程组$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x} = \boldsymbol{0}$的基础解系为

$$\boldsymbol{\xi}_1 = (-1,-1,1,0)^{\mathrm{T}},\ \boldsymbol{\xi}_2 = (-1,-2,0,1)^{\mathrm{T}},$$

则$S = \mathrm{span}\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2\}$就是 T 的正交补.

定义 3.2.6　设$S \oplus T = V$，若对 V 中任何向量$\boldsymbol{\alpha} = \boldsymbol{x} + \boldsymbol{y}$，其中$\boldsymbol{x} \in S$，$\boldsymbol{y} \in T$，线性变换$\sigma$：$V \to V$由下式确定

$$\sigma(\boldsymbol{\alpha}) = \boldsymbol{x},$$

则称σ是由 V 到 S 的**正交投影**.

显然σ是 V 沿 T 到 S 的投影，正交投影是特殊的投影变换，由于 S 的正交补 T 是唯一的，所以 V 到 S 的正交投影就不必指出是沿 T 的正交投影.

请注意直和、正交和概念的区别，以及分别与斜投影、正交投影的对应关系.

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σ是 Euclid 空间 V 的一个线性变换，如果对任意的$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$都有$(\sigma(\boldsymbol{\alpha}), \boldsymbol{\beta}) = \big(\boldsymbol{\alpha}, \sigma(\boldsymbol{\beta})\big)$，则称$\sigma$为 V 的一个**对称变换**.

定义 3.3.2　设σ是 Euclid 空间 V 的一个线性变换，如果对任意的$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$都

有$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=-(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta}))$，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$\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，设

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\ \boldsymbol{\alpha}_1\in W,\boldsymbol{\alpha}_2\in W_{\perp},$$
$$\boldsymbol{\beta}=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2,\ \boldsymbol{\beta}_1\in W,\boldsymbol{\beta}_2\in W_{\perp},$$

由正交投影的定义可知$P(\boldsymbol{\alpha})=\boldsymbol{\alpha}_1,P(\boldsymbol{\beta})=\boldsymbol{\beta}_1$，那么

$$(P(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)=(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2)=(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1),$$
$$\big(\boldsymbol{\alpha},P(\boldsymbol{\beta})\big)=(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)=(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1)+(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)=(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1),$$

于是$(P(\boldsymbol{\alpha}),\boldsymbol{\beta})=\big(\boldsymbol{\alpha},P(\boldsymbol{\beta})\big)$，故$P$是一个对称变换.　■

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

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

$$(\sigma(\boldsymbol{\beta}),\boldsymbol{\alpha})=\big(\boldsymbol{\beta},\sigma(\boldsymbol{\alpha})\big)=0.$$

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

同理可得定理 3.3.2.

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

定理 3.3.3　Euclid 空间 V 的线性变换σ是对称变换的充要条件是σ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$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$，$\sigma$在该标准正交基下对应的矩阵为$\boldsymbol{A}=(a_{ij})_{n\times n}$，则有

$$\sigma(\boldsymbol{\alpha}_i)=a_{1i}\boldsymbol{\alpha}_1+a_{2i}\boldsymbol{\alpha}_2+\cdots+a_{ni}\boldsymbol{\alpha}_n,\quad \big(\sigma(\boldsymbol{\alpha}_i),\boldsymbol{\alpha}_j\big)=a_{ji},$$
$$\sigma(\boldsymbol{\alpha}_j)=a_{1j}\boldsymbol{\alpha}_1+a_{2j}\boldsymbol{\alpha}_2+\cdots+a_{nj}\boldsymbol{\alpha}_n,\quad \big(\sigma(\boldsymbol{\alpha}_j),\boldsymbol{\alpha}_i\big)=a_{ij}.$$

由于σ是对称变换，所以$a_{ji}=\big(\sigma(\boldsymbol{\alpha}_i),\boldsymbol{\alpha}_j\big)=\Big(\boldsymbol{\alpha}_j,\sigma(\boldsymbol{\alpha}_i)\Big)=\big(\sigma(\boldsymbol{\alpha}_j),\boldsymbol{\alpha}_i\big)=a_{ij}$

这表明$\boldsymbol{A}$是对称矩阵.

②充分性

设线性变换σ在 V 的一个标准正交基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的矩阵$\boldsymbol{A}$是对称矩阵，任取$\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，且在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标分别为$\boldsymbol{X}$和$\boldsymbol{Y}$，即

$$\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{X},$$
$$\boldsymbol{\beta}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{Y}.$$

于是有

$$\sigma(\boldsymbol{\alpha})=\sigma(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{X}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{X},$$
$$\sigma(\boldsymbol{\beta})=\sigma(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{Y}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{Y},$$

则有

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y}=\big(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta})\big).$$

这表明σ是 V 的对称变换. ■

同理可得定理 3.3.4.

定理 3.3.4 Euclid 空间 V 的线性变换σ是反对称变换的充要条件是σ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$\boldsymbol{A}$，一定有n阶正交矩阵$\boldsymbol{Q}$，使$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$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 Euclid 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$\boldsymbol{u}$为过直角坐标系原点的平面 π 的单位法矢量，变换 T 定义为

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

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

$$T(k\boldsymbol{\alpha}+l\boldsymbol{\beta})=kT(\boldsymbol{\alpha})+lT(\boldsymbol{\beta}),$$
$$(T(\boldsymbol{\alpha}),T(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta}),$$
$$(T(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},T(\boldsymbol{\beta})).$$

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$\boldsymbol{A}\in\mathbb{C}^{\mathrm{n}\times\mathrm{n}}$, 若$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$, 则称$\boldsymbol{A}$为 Hermite 矩阵, 若$\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$, 则称$\boldsymbol{A}$为**反 Hermite 矩阵**.

对 n 阶矩阵$\boldsymbol{A}$, 有:

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

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

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

$$\boldsymbol{A}=\begin{bmatrix}1 & 2+3i\\ 2-3i & 4\end{bmatrix}\qquad \boldsymbol{B}=\begin{bmatrix}4 & 1-\mathrm{i} & 2+3\mathrm{i}\\ 1+\mathrm{i} & 6 & 5+\mathrm{i}\\ 2-3\mathrm{i} & 5-\mathrm{i} & 8\end{bmatrix}$$

是 Hermite 矩阵.

矩阵

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

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. 通常 Hermite 矩阵不对称, 除非所有元素均为实数.

Hermite 矩阵的性质:

(1) 若$\boldsymbol{A}$和$\boldsymbol{B}$是 Hermite 矩阵, 则$\boldsymbol{A}+\boldsymbol{B}$也是 Hermite 矩阵.

(2) 若$\boldsymbol{A}$和$\boldsymbol{B}$是 Hermite 矩阵且$\boldsymbol{AB}=\boldsymbol{BA}$, 则$\boldsymbol{AB}$也是 Hermite 矩阵.

(3) 若$\boldsymbol{A}$是 Hermite 矩阵且可逆, 则其逆矩阵$\boldsymbol{A}^{-1}$也是 Hermite 矩阵.

(4) 若$\boldsymbol{A}$是 Hermite 矩阵, 对$n\in\mathbb{Z}^+$, 则$\boldsymbol{A}^n$也是 Hermite 矩阵.

(5) 方阵$\boldsymbol{A}$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.

(7) 若 Hermite 矩阵$\boldsymbol{A}$的元素都是实数, 则$\boldsymbol{A}$就是实对称矩阵.

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$\boldsymbol{A}$是 n 阶 Hermite 矩阵，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$是任意一个 n 维复向量，称

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

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

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

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

定义 3.3.5 给定 Hermite 二次型

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

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

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

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

(1) $f(\boldsymbol{x})$是正定的，$\boldsymbol{A}$是正定的.

(2) 对于任意 n 阶可逆矩阵 $\boldsymbol{P}$，都有$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}$为正定矩阵.

(3) $\boldsymbol{A}$的 n 个特征值全大于零.

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

(1) $f(\boldsymbol{x})$是半正定的，$\boldsymbol{A}$是半正定的.

(2) 对于任意 n 阶可逆矩阵 $\boldsymbol{P}$，都有$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}$为半正定矩阵.

(3) $\boldsymbol{A}$的 n 个特征值全是非负实数.

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

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta})),$$

则称σ是V的一个 **Hermite 变换**.

由定义可见，酉空间中的 Hermite 变换与 Euclid 空间中的对称变换形式相同.

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

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=-(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta})),$$

则称σ是V的一个**反 Hermite 变换**.

由定义可见，酉空间中的反 Hermite 变换与 Euclid 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σ是 Hermite 变换的充要条件是σ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$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$，$\sigma$在该标准正交基下对应的矩阵为$\boldsymbol{A}=(a_{ij})_{n\times n}$，则有

$$\sigma(\boldsymbol{\alpha}_i)=a_{1i}\boldsymbol{\alpha}_1+a_{2i}\boldsymbol{\alpha}_2+\cdots+a_{ni}\boldsymbol{\alpha}_n,\quad \left(\sigma(\boldsymbol{\alpha}_i),\boldsymbol{\alpha}_j\right)=a_{ji},$$

$$\sigma\left(\boldsymbol{\alpha}_j\right)=a_{1j}\boldsymbol{\alpha}_1+a_{2j}\boldsymbol{\alpha}_2+\cdots+a_{nj}\boldsymbol{\alpha}_n,\quad \left(\sigma\left(\boldsymbol{\alpha}_j\right),\boldsymbol{\alpha}_i\right)=a_{ij}.$$

由于σ是 Hermite 变换，所以

$$a_{ji}=\left(\sigma(\boldsymbol{\alpha}_i),\boldsymbol{\alpha}_j\right)=\left(\boldsymbol{\alpha}_i,\sigma\left(\boldsymbol{\alpha}_j\right)\right)=\overline{\left(\sigma\left(\boldsymbol{\alpha}_j\right),\boldsymbol{\alpha}_i\right)}=\overline{a_{ij}}.$$

这表明$\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，即$\boldsymbol{A}$是 Hermite 矩阵.

②充分性

设线性变换σ在 V的一个标准正交基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的矩阵$\boldsymbol{A}$是 Hermite 矩阵，即$\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，任取$\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，且在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标分别为$\boldsymbol{X}$和$\boldsymbol{Y}$，即

$$\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{X},$$

$$\boldsymbol{\beta}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{Y}.$$

于是有

$$\sigma(\boldsymbol{\alpha})=\sigma(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{X}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{X},$$

$$\sigma(\boldsymbol{\beta})=\sigma(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{Y}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{Y},$$

则有

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{X}=(\boldsymbol{A}\boldsymbol{Y})^{\mathrm{H}}\boldsymbol{X}=\left(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta})\right).$$

这表明σ是 V的 Hermite 变换. ■

定理 3.3.8 酉空间 V的线性变换σ是反 Hermite 变换的充要条件是σ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$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，若

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A},$$

则称$\boldsymbol{A}$为**正规矩阵**.

定义 3.4.2　设$\boldsymbol{A} \in \mathbb{R}^{n\times n}$，若

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},$$

则称$\boldsymbol{A}$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$\boldsymbol{A},\boldsymbol{B} \in \mathbb{C}^{n\times n}$，若存在$\boldsymbol{U} \in U^{n\times n}$，使得

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

则称$\boldsymbol{A}$酉相似于$\boldsymbol{B}$.

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

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

则称$\boldsymbol{A}$正交相似于$\boldsymbol{B}$.

酉相似（正交相似）符合矩阵相似的定义，且这里的$\boldsymbol{U}$是酉矩阵（正交矩阵）.

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

引理 3.4.2　设$\boldsymbol{A}$是正规矩阵，则与$\boldsymbol{A}$酉相似的矩阵都是正规矩阵.

引理 3.4.3　设$\boldsymbol{A}$是正规矩阵，且$\boldsymbol{A}$是三角矩阵，则$\boldsymbol{A}$是对角阵.

证：　设

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

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

$$
\begin{gathered}
a_{11}\bar{a}_{11}+a_{12}\bar{a}_{12}+\cdots+a_{1n}\bar{a}_{1n}=\bar{a}_{11}a_{11}\\
a_{22}\bar{a}_{22}+\cdots+a_{2n}\bar{a}_{2n}=\bar{a}_{12}a_{12}+\bar{a}_{22}a_{22}\\
\vdots\\
a_{nn}\bar{a}_{nn}=\bar{a}_{1n}a_{1n}+\bar{a}_{2n}a_{2n}+\cdots+\bar{a}_{nn}a_{nn}
\end{gathered}
$$

由第一个等式，有$a_{12}\bar{a}_{12}+\cdots+a_{1n}\bar{a}_{1n}=0$，又$a_{ij}\bar{a}_{ij}=\left|a_{ij}\right|^2\geq 0$，可得$a_{12}\bar{a}_{12}=\cdots=a_{1n}\bar{a}_{1n}=0$, 将此结果再代入第二式, 以此类推直至第 n 式, 可得: 当$i\neq j$时, $a_{ij}=0$, 即$\boldsymbol{A}$是对角阵. ■

定理 3. 4. 1（正规矩阵的结构定理） 设$\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 则$\boldsymbol{A}$是正规矩阵的充要条件是存在$\boldsymbol{U}\in U^{n\times n}$, 使得

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

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

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

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

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

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

推论 3. 4. 1 设$\boldsymbol{A}$是正规矩阵, λ_i是$\boldsymbol{A}$的特征值, 对应的特征向量是$\boldsymbol{x}$, 则$\overline{\lambda_i}$是$\boldsymbol{A}^{\mathrm{H}}$的特征值, 其对应的特征向量是$\boldsymbol{x}$.

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$\boldsymbol{A}$有n个线性无关的特征向量.

证: 设$\boldsymbol{A}$是正规矩阵，$\lambda_i(i=1,2,\cdots,n)$是$\boldsymbol{A}$的特征值，则存在$\boldsymbol{U}=[\boldsymbol{U}_1,\boldsymbol{U}_2,\cdots,\boldsymbol{U}_n]\in U^{n\times n}$, 使得

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

也即

$$\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix}.$$

因此

$$\boldsymbol{A}[\boldsymbol{U}_1,\boldsymbol{U}_2,\cdots,\boldsymbol{U}_n]=[\boldsymbol{U}_1,\boldsymbol{U}_2,\cdots,\boldsymbol{U}_n]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}=[\lambda_1\boldsymbol{U}_1,\lambda_2\boldsymbol{U}_2,\cdots,\lambda_n\boldsymbol{U}_n],$$

展开得

$$\boldsymbol{A}\boldsymbol{U}_1=\lambda_1\boldsymbol{U}_1,$$
$$\boldsymbol{A}\boldsymbol{U}_2=\lambda_2\boldsymbol{U}_2,$$
$$\vdots$$
$$\boldsymbol{A}\boldsymbol{U}_n=\lambda_n\boldsymbol{U}_n.$$

由于$\boldsymbol{U}\in U^{n\times n}$，$\boldsymbol{U}_1,\boldsymbol{U}_2,\cdots,\boldsymbol{U}_n$是标准正交向量组，证毕. ■

特征向量加上零向量生成的子空间称为**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$\boldsymbol{A}$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$\boldsymbol{A}$的全部特征值. 类似地，现在讨论当$\boldsymbol{A}$是n阶正规矩阵时，如何求n阶酉矩阵$\boldsymbol{U}$，使$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}$为一个对角阵. $\boldsymbol{U}$称为**相似因子**.

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

(1) 求$\boldsymbol{A}$的特征值，即求$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$的根$\lambda_1,\lambda_2$，$\cdots$，$\lambda_n$;

(2) 对每一个相异特征值λ_i，$i=1,2,\cdots,s$，求λ_i的特征向量，即求$(\lambda_i\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$的非零解;

(3) 用 Schmidt 方法求$\lambda_i(i=1,2,\cdots,s)$的特征向量生成的空间的标准正交基$\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{in_i}$;

(4) 令$\boldsymbol{U}=(\boldsymbol{\alpha}_{11},\boldsymbol{\alpha}_{12},\cdots,\boldsymbol{\alpha}_{1n_1},\boldsymbol{\alpha}_{21},\boldsymbol{\alpha}_{22},\cdots,\boldsymbol{\alpha}_{2n_2},\cdots,\boldsymbol{\alpha}_{s1},\boldsymbol{\alpha}_{s2},\cdots,\boldsymbol{\alpha}_{sn_s})$，则得到的酉矩阵$\boldsymbol{U}$满足$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

例 3.4.1 已知$\boldsymbol{A}=\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$，验证$\boldsymbol{A}$是正规矩阵，并求酉矩阵$\boldsymbol{U}$，使$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}$为对角矩阵.

解: 由于$\boldsymbol{A}^{\mathrm{H}}=\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$，可算得

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A},$$

故$\boldsymbol{A}$是正规矩阵.

$\boldsymbol{A}$的特征多项式

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1 & 1\\ -1 & \lambda-1\end{vmatrix}=\lambda^2-2\lambda+2,$$

故$\boldsymbol{A}$的特征值是：$\lambda_1=1+\mathrm{i}$，$\lambda_2=1-\mathrm{i}$.

当$\lambda_1=1+\mathrm{i}$ 时，特征矩阵

$$\lambda_1\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}\mathrm{i} & 1\\ -1 & \mathrm{i}\end{bmatrix}\rightarrow\begin{bmatrix}-1 & \mathrm{i}\\ 0 & 0\end{bmatrix},$$

故$x_1=\mathrm{i}x_2$.

取x_2为自由未知量，可得属于$\lambda_1=1+\mathrm{i}$的单位特征向量$\boldsymbol{\alpha}_1=\left(\frac{\mathrm{i}}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$.

当$\lambda_2=1-\mathrm{i}$ 时，特征矩阵

$$\lambda_2\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}-\mathrm{i} & 1\\ -1 & -\mathrm{i}\end{bmatrix}\rightarrow\begin{bmatrix}1 & \mathrm{i}\\ 0 & 0\end{bmatrix},$$

故$x_1=-\mathrm{i}x_2$.

取x_2为自由未知量，可得属于$\lambda_2=1-\mathrm{i}$的单位特征向量$\boldsymbol{\alpha}_2=\left(\frac{-\mathrm{i}}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$.

令

$$\boldsymbol{U}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)=\begin{bmatrix}\frac{\mathrm{i}}{\sqrt{2}} & -\frac{\mathrm{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix},$$

则$\boldsymbol{U}$是酉矩阵，且满足

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}1+\mathrm{i} & 0\\ 0 & 1-\mathrm{i}\end{bmatrix}.$$

3.4.3 伴随变换和正规变换

定义 3.4.5 设V是一个酉空间(Euclid 空间)，σ是V的一个线性变换，如果存在V的一个线性变换σ^{H}，使得

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\sigma^{\mathrm{H}}(\boldsymbol{\beta}))\qquad\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V,$$

则称σ有一个**伴随变换**σ^{H}.

例 3.4.2 设σ是 Euclid 空间 V 的一个对称变换，则根据对称变换的定义

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=\big(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta})\big)\qquad\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V,$$

则有$\sigma^{\mathrm{H}}=\sigma$.

例 3.4.3 设σ是酉空间 V 的一个 Hermite 变换，则根据 Hermite 变换的定义

$$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=\big(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta})\big)\qquad\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V,$$

则有$\sigma^{\mathrm{H}} = \sigma$. 因此, Hermite 变换也被称为**自伴随变换**.

例 3.4.4 设σ是 Euclid 空间 V 的一个正交变换, 则根据正交变换的定义

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

则有$\sigma^{\mathrm{H}} = \sigma^{-1}$. 这是因为

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

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

定义 3.4.6 设 V 是酉空间（Euclid 空间）, σ是 V 的一个线性变换, 如果σ满足

$$\sigma^{\mathrm{H}}\sigma = \sigma\sigma^{\mathrm{H}},$$

则称σ是**正规变换**.

Euclid 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（Euclid 空间）, σ是 V 的一个线性变换, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$是 V 的一个标准正交基, 且σ在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$下对应的矩阵为$\boldsymbol{A} = (a_{ij})_{n\times n}$, 则$\boldsymbol{A}$的伴随变换$\sigma^{\mathrm{H}}$在基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$下的矩阵$\boldsymbol{B}$为$\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$.

证: 由伴随变换定义可知

$$\big(\sigma(\boldsymbol{\alpha}_i), \boldsymbol{\alpha}_j\big) = (\boldsymbol{\alpha}_i, \sigma^{\mathrm{H}}(\boldsymbol{\alpha}_j)) \ ,\quad i, j = 1,2, \cdots, n,$$

又$\big(\sigma(\boldsymbol{\alpha}_i), \boldsymbol{\alpha}_j\big) = a_{ji}$, 设$\boldsymbol{B} = (b_{ij})_{n\times n}$, 则有

$$\big(\sigma(\boldsymbol{\alpha}_i), \boldsymbol{\alpha}_j\big) = \Big(\boldsymbol{\alpha}_i, \sigma^{\mathrm{H}}\big(\boldsymbol{\alpha}_j\big)\Big) = \overline{\big(\sigma^{\mathrm{H}}\big(\boldsymbol{\alpha}_j\big), \boldsymbol{\alpha}_i\big)} = \overline{b_{ij}},$$

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

定理 3.4.3 设 V 是一个 n 维酉空间（Euclid 空间）, σ是 V 的一个线性变换, σ是正规变换当且仅当σ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$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$, 设$\sigma$在这个基下对应的矩阵为$\boldsymbol{A}$, 那么$\sigma^{\mathrm{H}}$在这个基下对应的矩阵为$\boldsymbol{A}^{\mathrm{H}}$, 又

$$\sigma\sigma^{\mathrm{H}} = \sigma^{\mathrm{H}}\sigma \ \Leftrightarrow \ \boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A},$$

所以, σ是 V 的正规变换当且仅当$\boldsymbol{A}$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σ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σ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$\boldsymbol{\alpha} \in \mathbb{C}^n$，且$\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\alpha} = 1$，则酉矩阵

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

所代表的酉变换称为 **Householder**(豪斯霍尔德)**镜像变换**.

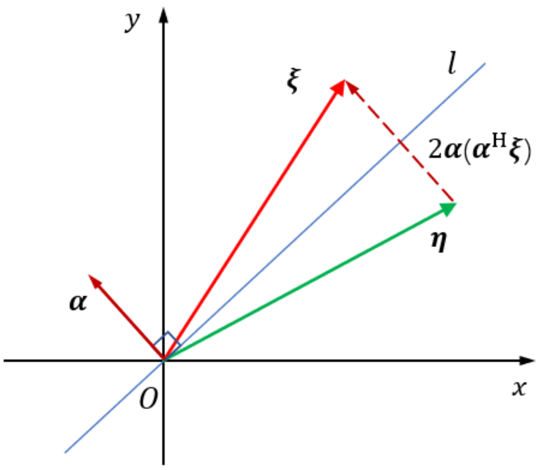


图 3.5.1　Householder 变换示意图

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$\boldsymbol{\xi}$和$\boldsymbol{\eta}$，而$\boldsymbol{\alpha}$是与$\boldsymbol{\xi}-\boldsymbol{\eta}$同方向的单位向量(向量$\boldsymbol{\alpha}$的长度是$\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})} = \sqrt{\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\alpha}} = 1$), l 轴(无方向)与向量$\boldsymbol{\alpha}$垂直.

向量$\boldsymbol{\xi}$在$\boldsymbol{\alpha}$上的投影长度是

$$(\boldsymbol{\xi},\boldsymbol{\alpha}) = \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\xi},$$

因此

$$\boldsymbol{\xi} - \boldsymbol{\eta} = 2\boldsymbol{\alpha}(\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\xi}),$$

所以

$$\boldsymbol{\eta} = \boldsymbol{\xi} - 2\boldsymbol{\alpha}(\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\xi}) = (\boldsymbol{E} - 2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}})\boldsymbol{\xi} = \boldsymbol{H}\boldsymbol{\xi}.$$

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

综上，矩阵$\boldsymbol{H}=\boldsymbol{E}-2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}$作用在原始向量$\boldsymbol{\xi}$上，代表对$\boldsymbol{\xi}$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$\boldsymbol{\alpha}$垂直的平面(向量、平面都要过坐标原点)，$\boldsymbol{\eta}=\boldsymbol{H}\boldsymbol{\xi}$即为变换所得向量，因此, Householder 变换又称为**镜像变换**，酉矩阵$\boldsymbol{H}$也称为 **Householder 矩阵**，镜面的法向量$\boldsymbol{\alpha}$也称为 **Householder 向量**.

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 Euclid 空间 V 的一个线性子空间，对于任意向量$\boldsymbol{\alpha}\in V$，$\boldsymbol{\alpha}_1\in W$是$\boldsymbol{\alpha}$在 W 上的正交投影的充要条件是

$$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\le|\boldsymbol{\alpha}-\boldsymbol{\beta}|,\ \forall\boldsymbol{\beta}\in W.$$

证：①必要性　设$\boldsymbol{\alpha}_1\in W$是$\boldsymbol{\alpha}$在 W 上的正交投影，则对$\forall\boldsymbol{\beta}\in W$，有

$$(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1)\perp(\boldsymbol{\alpha}_1-\boldsymbol{\beta}),$$

$$|\boldsymbol{\alpha}-\boldsymbol{\beta}|^2=|(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1)+(\boldsymbol{\alpha}_1-\boldsymbol{\beta})|^2=|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|^2+|\boldsymbol{\alpha}_1-\boldsymbol{\beta}|^2\ge|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|^2,$$

即$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\le|\boldsymbol{\alpha}-\boldsymbol{\beta}|$.

②充分性　设$\boldsymbol{\delta}$是$\boldsymbol{\alpha}$在 W 上的正交投影，则对于$\boldsymbol{\alpha}_1\in W$，有

$$(\boldsymbol{\alpha}-\boldsymbol{\delta})\perp(\boldsymbol{\delta}-\boldsymbol{\alpha}_1),$$

$$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|^2=|(\boldsymbol{\alpha}-\boldsymbol{\delta})+(\boldsymbol{\delta}-\boldsymbol{\alpha}_1)|^2=|\boldsymbol{\alpha}-\boldsymbol{\delta}|^2+|\boldsymbol{\delta}-\boldsymbol{\alpha}_1|^2\ge|\boldsymbol{\alpha}-\boldsymbol{\delta}|^2,$$

即$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\ge|\boldsymbol{\alpha}-\boldsymbol{\delta}|$. 又由于$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\le|\boldsymbol{\alpha}-\boldsymbol{\beta}|$，$\forall\boldsymbol{\beta}\in W$，有$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\le|\boldsymbol{\alpha}-\boldsymbol{\delta}|$，所以$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|=|\boldsymbol{\alpha}-\boldsymbol{\delta}|$，$\forall\boldsymbol{\alpha}\in V$，所以$\boldsymbol{\alpha}_1=\boldsymbol{\delta}$. ■

定义 3.5.2　设 W 是 Euclid 空间 V 的一个线性子空间，对于任意向量$\boldsymbol{\alpha}\in V$，若存在$\boldsymbol{\alpha}_1\in W$，使得

$$|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\le|\boldsymbol{\alpha}-\boldsymbol{\beta}|,\ \forall\boldsymbol{\beta}\in W,$$

则称$\boldsymbol{\alpha}_1$是$\boldsymbol{\alpha}$在 W 上的**最佳逼近元**.

$\boldsymbol{\alpha}$在 W 上的最佳逼近元也就是$\boldsymbol{\alpha}$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$(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$，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$y_i \approx a_0 + a_1x_i$，$i = 1,2,\cdots,n$. 由于不存在a_0和a_1使得$y_i = a_0 + a_1x_i$，因此方程组

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

当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

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

其中，

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

该问题是求参数向量$\boldsymbol{a}$使$\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{a}$长度最小，若将$\boldsymbol{Y}$看成$\mathbb{R}^n$空间中的一个向量，则问题转化为求$\boldsymbol{a}$使$\boldsymbol{A}\boldsymbol{a}$是向量$\boldsymbol{Y}$在系数矩阵$\boldsymbol{A}$的列向量$(1,1,\cdots,1)^{\mathrm{T}}$和$(x_1, x_2, \cdots, x_n)^{\mathrm{T}}$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$e_i = y_i - (a_0 + a_1x_i)$，$i = 1,2,\cdots,n$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$\min\limits_{a_0,a_1} \max\limits_{1\le i\le n} |e_i|$;

② 使误差的绝对值之和最小，即$\min\limits_{a_0,a_1} \sum_{i=1}^{n} |e_i|$;

③ 使误差的平方之和最小，即$\min\limits_{a_0,a_1} \sum_{i=1}^{n} e_i^2$.

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$x_1, x_2, \cdots, x_n$成线性关系时，拟合函数可表示为:

$$\varphi(\boldsymbol{x}) = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n.$$

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

$$\begin{cases} a_0 + a_1 x_{11} + a_2 x_{21} + \cdots + a_n x_{n1} = y_1 \\ a_0 + a_1 x_{12} + a_2 x_{22} + \cdots + a_n x_{n2} = y_2 \\ \qquad\vdots \\ a_0 + a_1 x_{1m} + a_2 x_{2m} + \cdots + a_n x_{nm} = y_m \end{cases},$$

其矩阵形式为

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

其中,

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

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$a_0, a_1, \cdots, a_n$的原则应是使误差$e_i = y_i - \varphi(\boldsymbol{x}_i)$，$i = 1,2, \cdots, m$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:

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

其中，$a_0, a_1, \cdots, a_n$为待定参数. 设总共进行了 m 次观测，确定待定参数$a_0, a_1, \cdots, a_n$的原则应是使误差$e_i = y_i - \varphi(x_i)$，$i = 1,2, \cdots, m$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

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

则拟合函数可写为

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

从而转化为多变量拟合问题进行处理.

对于最小二乘法，有

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

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

$$J=\sum_{i=1}^{n} e_i^2 = (\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{a})^{\mathrm{T}}(\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{a}) = \boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{\delta}$$

$$=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y}-\boldsymbol{a}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}-\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{a}+\boldsymbol{a}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{a},$$

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

$$\frac{\partial J}{\partial \boldsymbol{a}} = -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}+(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{a}=\boldsymbol{0},$$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{a}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y},$$

当$m\times n$矩阵$\boldsymbol{A}$列满秩时，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$非奇异，可得

$$\boldsymbol{a}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}.$$

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$\boldsymbol{Y}$时，按最小二乘法可得相应的预报值(拟合值)向量$\widehat{\boldsymbol{Y}}$，即

$$\widehat{\boldsymbol{Y}}=\boldsymbol{A}\boldsymbol{a}=\boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{L}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y},\quad (\boldsymbol{L}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}),$$

称拟合误差$\boldsymbol{e}=\boldsymbol{Y}-\widehat{\boldsymbol{Y}}$为残差向量，并记$\boldsymbol{P}=\boldsymbol{A}\boldsymbol{L}^{-1}\boldsymbol{A}^{\mathrm{T}}$，则有

$$\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P};\quad \boldsymbol{P}^2=\boldsymbol{P},$$

$$(\boldsymbol{E}-\boldsymbol{P})^{\mathrm{T}}=\boldsymbol{E}-\boldsymbol{P};\quad (\boldsymbol{E}-\boldsymbol{P})^2=\boldsymbol{E}-2\boldsymbol{P}+\boldsymbol{P}^2=\boldsymbol{E}-\boldsymbol{P},$$

即$\boldsymbol{P}$和$\boldsymbol{E}-\boldsymbol{P}$都是幂等矩阵.

$$\widehat{\boldsymbol{Y}}=\boldsymbol{A}\boldsymbol{a}=\boldsymbol{A}\boldsymbol{L}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{P}\boldsymbol{Y}, \tag{1}$$

$$\boldsymbol{e}=\boldsymbol{Y}-\widehat{\boldsymbol{Y}}=\boldsymbol{Y}-\boldsymbol{P}\boldsymbol{Y}=(\boldsymbol{E}-\boldsymbol{P})\boldsymbol{Y}, \tag{2}$$

$$\boldsymbol{P}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{L}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A};\quad (\boldsymbol{E}-\boldsymbol{P})\boldsymbol{A}=\boldsymbol{0},$$

$$\left(\widehat{\boldsymbol{Y}},\boldsymbol{e}\right)=(\boldsymbol{P}\boldsymbol{Y})^{\mathrm{T}}(\boldsymbol{E}-\boldsymbol{P})\boldsymbol{Y}=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{E}-\boldsymbol{P})\boldsymbol{Y}=0, \tag{3}$$

即$\widehat{\boldsymbol{Y}}\perp\boldsymbol{e}$，由式(1)~(3)，向量之间的关系如图 3.5.2 所示.

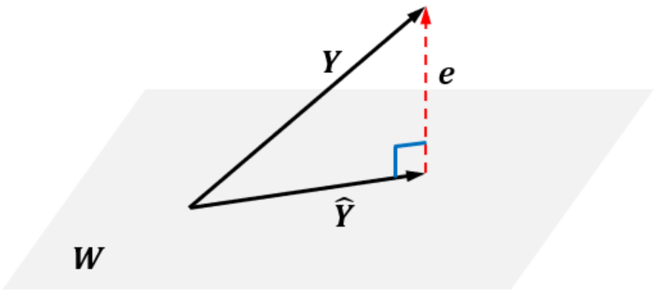


图 3.5.2　向量关系示意图

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

综上，由最小二乘法求得的估计值$\hat{\boldsymbol{Y}}$是$\boldsymbol{Y}$在子空间 W 上的最佳逼近元，因此，采用目标函数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.

附表 1 总结了本章内容.

附表 1　典型矩阵与变换总结

<table>
<tr><th></th><th>矩阵</th><th>变换</th><th>特征值</th><th>特征向量</th><th>对角化</th></tr>
<tr><td rowspan="7">正规矩阵</td><td>酉矩阵
$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{E}$</td><td>酉变换
$(\sigma(\boldsymbol{\alpha}),\sigma(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta})$</td><td>模等于 1</td><td rowspan="2">特征向量互相正交</td><td rowspan="2">可以</td></tr>
<tr><td>正交矩阵
$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{E}$</td><td>正交变换
$(\sigma(\boldsymbol{\alpha}),\sigma(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta})$</td><td>$\pm1$</td></tr>
<tr><td>Hermite 矩阵$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$</td><td>Hermite 变换
$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta}))$</td><td>实数</td><td rowspan="2">不同特征值对应的特征向量互相正交</td><td rowspan="2">可以</td></tr>
<tr><td>对称矩阵
$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$</td><td>对称变换
$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta}))$</td><td>实数</td></tr>
<tr><td>反 Hermite 矩阵
$\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$</td><td>反 Hermite 变换
$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=-(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta}))$</td><td>0 或虚数</td><td rowspan="2">特征值之和不为零的两个特征向量互相正交</td><td rowspan="2">可以</td></tr>
<tr><td>反对称矩阵
$\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$</td><td>反对称变换
$(\sigma(\boldsymbol{\alpha}),\boldsymbol{\beta})=-(\boldsymbol{\alpha},\sigma(\boldsymbol{\beta}))$</td><td>0 或虚数</td></tr>
<tr><td>正规矩阵
$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$</td><td>正规变换
$\sigma^{\mathrm{H}}\sigma=\sigma\sigma^{\mathrm{H}}$</td><td>$\boldsymbol{A}^{\mathrm{H}}$和$\boldsymbol{A}$的特征值互为共轭</td><td>$n$ 阶矩阵$\boldsymbol{A}$有 n 个线性无关的特征向量</td><td>可以</td></tr>
<tr><td rowspan="2">幂等矩阵</td><td>幂等矩阵$\boldsymbol{A}^2=\boldsymbol{A}$</td><td>投影变换$\sigma(\boldsymbol{x}+\boldsymbol{y})=\boldsymbol{x}$</td><td>0 或 1</td><td>无明显规律</td><td rowspan="2">可以</td></tr>
<tr><td>$\boldsymbol{A}^2=\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$</td><td>正交投影变换$\sigma(\boldsymbol{x}+\boldsymbol{y})=\boldsymbol{x}$，其中$\boldsymbol{x}\perp\boldsymbol{y}$</td><td>0 或 1</td><td>不同特征值对应的特征向量互相正交</td></tr>
</table>

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$\boldsymbol{A} \in U^{n\times n}$，试证明:

(1)　$|\det\boldsymbol{A}| = 1$;

(2)　$\boldsymbol{A}^{\mathrm{T}} \in U^{n\times n}$;

(3)　$\boldsymbol{A}^2 \in U^{n\times n}$.

3-2　试写出:

(1)　$\mathbb{R}^2$空间中将向量$\boldsymbol{\alpha}$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$\boldsymbol{\alpha}$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θ角的旋转变换的矩阵$\boldsymbol{R}_y(\theta)$;

(2)　三维空间中的向量绕z轴旋转θ角的旋转变换的矩阵$\boldsymbol{R}_z(\theta)$.

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$\boldsymbol{\alpha}$关于y轴对称;

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

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

3-6　求正交矩阵$\boldsymbol{Q}$，使$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}$为对角矩阵，已知

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

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

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

3-8　填空:

(1)　设σ是 Euclid 空间 V 的一个反对称变换，则其伴随变换$\sigma^{\mathrm{H}}=$ ________.

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

3-9　试求矩阵$\boldsymbol{P}$，使$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{E}$ (或$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{E}$)，已知

(1)　$\boldsymbol{A}=\begin{bmatrix}1 & i & 1+i\\ -i & 0 & 1\\ 1-i & 1 & 2\end{bmatrix}$

(2)　$\boldsymbol{A}=\begin{bmatrix}2 & 2 & -2\\ 2 & 5 & -4\\ -2 & -4 & 5\end{bmatrix}$

3-10　验证下列矩阵是否是正规矩阵，并求酉矩阵$\boldsymbol{U}$，使$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}$为对角矩阵.

(1)　$\boldsymbol{A}=\begin{bmatrix}0 & -1 & i\\ 1 & 0 & 0\\ i & 0 & 0\end{bmatrix}$

(2)　$\boldsymbol{A}=\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$

3-11　设$\boldsymbol{A}$、$\boldsymbol{B}$均是 Hermite 矩阵，试证：$\boldsymbol{A}$与$\boldsymbol{B}$酉相似的充要条件是$\boldsymbol{A}$与$\boldsymbol{B}$的特征值相同.

3-12　设$\boldsymbol{A}$、$\boldsymbol{B}$均是 Hermite 矩阵，试证：$\boldsymbol{AB}$也是 Hermite 矩阵的充要条件是$\boldsymbol{AB}=\boldsymbol{BA}$.

3-13　设$\boldsymbol{A}$、$\boldsymbol{B}$均是正规矩阵，试证：$\boldsymbol{A}$与$\boldsymbol{B}$酉相似的充要条件是$\boldsymbol{A}$与$\boldsymbol{B}$的特征值相同.

3-14　设$\boldsymbol{A}$、$\boldsymbol{B}$均是正规矩阵，且$\boldsymbol{AB}=\boldsymbol{BA}$，试证：$\boldsymbol{AB}$与$\boldsymbol{BA}$均是正规矩阵.

3-15　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.

3-16　已知$\boldsymbol{A}=\begin{bmatrix}0 & i & -1\\ -i & 0 & i\\ -1 & -i & 0\end{bmatrix}$，验证$\boldsymbol{A}$是正规矩阵，并求酉矩阵$\boldsymbol{U}$，使$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}$为对角矩阵.