§5. 酉（正交）相似下的标准形

一、预备知识

1. 向量的内积

定义 设 $\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}\in\mathbf{C}^n$， $\boldsymbol{y}=(\eta_1,\eta_2,\cdots,\eta_n)^{\mathrm{T}}\in\mathbf{C}^n$，令

$$(\boldsymbol{x},\boldsymbol{y})=\xi_1\overline{\eta_1}+\xi_2\overline{\eta_2}+\cdots+\xi_n\overline{\eta_n}=\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x}$$

称 $(\boldsymbol{x},\boldsymbol{y})$ 为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的**内积**。

例 已知 $\boldsymbol{x}=(3,4,5\mathrm{i})^{\mathrm{T}}$， $\boldsymbol{y}=(2\mathrm{i},0,\mathrm{i})^{\mathrm{T}}$，求 $(\boldsymbol{x},\boldsymbol{y})$ 和 $(\boldsymbol{x},\boldsymbol{x})$。

解 $(\boldsymbol{x},\boldsymbol{y})=3\times(-2\mathrm{i})+4\times0+5\mathrm{i}\times(-\mathrm{i})=5-6\mathrm{i}$，

$(\boldsymbol{x},\boldsymbol{x})=3\times3+4\times4+5\mathrm{i}\times(-5\mathrm{i})=50$

由内积的定义容易证明下列性质：

1） $(\boldsymbol{x},\boldsymbol{y})=\overline{(\boldsymbol{y},\boldsymbol{x})}$；

2） $(k\boldsymbol{x},\boldsymbol{y})=k\,(\boldsymbol{x},\boldsymbol{y})$； （但 $(\boldsymbol{x},k\boldsymbol{y})=\bar{k}\,(\boldsymbol{x},\boldsymbol{y})$）；

3） $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{z})+(\boldsymbol{y},\boldsymbol{z})$；

4） $(\boldsymbol{x},\boldsymbol{x})\geq0$，仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时，$(\boldsymbol{x},\boldsymbol{x})=0$；

5） **Cauchy-Schwarz 不等式**

$$(\boldsymbol{x},\boldsymbol{y})\,(\boldsymbol{y},\boldsymbol{x})\leq(\boldsymbol{x},\boldsymbol{x})(\boldsymbol{y},\boldsymbol{y}) \quad \text{或} \quad |(\boldsymbol{x},\boldsymbol{y})|^2\leq(\boldsymbol{x},\boldsymbol{x})(\boldsymbol{y},\boldsymbol{y})$$

证 当 $\boldsymbol{y}=\boldsymbol{0}$ 时，等式成立。下设 $\boldsymbol{y}\neq\boldsymbol{0}$，则对任意 $t\in\mathbf{C}$，有

$$0\leq(\boldsymbol{x}-t\boldsymbol{y},\boldsymbol{x}-t\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x})-t(\boldsymbol{y},\boldsymbol{x})-\bar{t}(\boldsymbol{x},\boldsymbol{y})+t\bar{t}(\boldsymbol{y},\boldsymbol{y})$$

取 $t=\dfrac{(\boldsymbol{x},\boldsymbol{y})}{(\boldsymbol{y},\boldsymbol{y})}$，代人上式得

$$0\leq(\boldsymbol{x},\boldsymbol{x})-\frac{(\boldsymbol{x},\boldsymbol{y})}{(\boldsymbol{y},\boldsymbol{y})}(\boldsymbol{y},\boldsymbol{x})-\frac{(\boldsymbol{y},\boldsymbol{x})}{(\boldsymbol{y},\boldsymbol{y})}(\boldsymbol{x},\boldsymbol{y})+\frac{(\boldsymbol{x},\boldsymbol{y})}{(\boldsymbol{y},\boldsymbol{y})}\frac{(\boldsymbol{y},\boldsymbol{x})}{(\boldsymbol{y},\boldsymbol{y})}(\boldsymbol{y},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x})-\frac{(\boldsymbol{x},\boldsymbol{y})(\boldsymbol{y},\boldsymbol{x})}{(\boldsymbol{y},\boldsymbol{y})}$$

整理得 $$0\leq(\boldsymbol{x},\boldsymbol{x})(\boldsymbol{y},\boldsymbol{y})-(\boldsymbol{x},\boldsymbol{y})(\boldsymbol{y},\boldsymbol{x})$$ **证毕**

2．向量的长度

定义　非负实数$\sqrt{(\boldsymbol{x},\boldsymbol{x})}$称为向量$\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}\in\mathbf{C}^n$的**长度**（或**模**，**2-范数**），记为 $|\boldsymbol{x}|$ 或$\|\boldsymbol{x}\|_2$ ，即

$$\|\boldsymbol{x}\|_2=\sqrt{(\boldsymbol{x},\boldsymbol{x})}=\sqrt{\sum_{i=1}^{n}\xi_i\overline{\xi_i}}=\sqrt{\sum_{i=1}^{n}|\xi_i|^2}$$

由定义可推知长度的如下性质：

1）非负性 若$\boldsymbol{x}\neq\boldsymbol{0}$，则$\|\boldsymbol{x}\|_2>0$，仅当时$\boldsymbol{x}=\boldsymbol{0}$，$\|\boldsymbol{x}\|_2=0$；

2）齐次性 $\|k\boldsymbol{x}\|_2=|k|\|\boldsymbol{x}\|_2$，　$k\in\mathbf{C}$；

证　$\|k\boldsymbol{x}\|_2=\sqrt{(k\boldsymbol{x},k\boldsymbol{x})}=\sqrt{k\bar{k}(\boldsymbol{x},\boldsymbol{x})}=|k|\sqrt{(\boldsymbol{x},\boldsymbol{x})}=|k|\|\boldsymbol{x}\|_2$　**证毕**

3）三角不等式 $\|\boldsymbol{x}+\boldsymbol{y}\|_2\le\|\boldsymbol{x}\|_2+\|\boldsymbol{y}\|_2$。

证　$\|\boldsymbol{x}+\boldsymbol{y}\|_2^2=(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{x})+(\boldsymbol{x},\boldsymbol{y})+(\boldsymbol{y},\boldsymbol{x})+(\boldsymbol{y},\boldsymbol{y})$

$$=(\boldsymbol{x},\boldsymbol{x})+2\,\mathrm{Re}(\boldsymbol{x},\boldsymbol{y})+(\boldsymbol{y},\boldsymbol{y})\le(\boldsymbol{x},\boldsymbol{x})+2|(\boldsymbol{x},\boldsymbol{y})|+(\boldsymbol{y},\boldsymbol{y})$$

$$\le\|\boldsymbol{x}\|_2^2+2\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2+\|\boldsymbol{y}\|_2^2=(\|\boldsymbol{x}\|_2+\|\boldsymbol{y}\|_2)^2$$　**证毕**

定义　若$\|\boldsymbol{x}\|_2=1$（或$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=1$），称$\boldsymbol{x}$为**单位向量**。如果$\boldsymbol{x}\neq\boldsymbol{0}$，则$\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}$为单位向量，这一过程称为将$\boldsymbol{x}$**单位化**。

例　已知向量$\boldsymbol{x}=(3+\mathrm{i},-\mathrm{i},2,1-\mathrm{i})^{\mathrm{T}}$，试将其单位化。

解　因为 $\|\boldsymbol{x}\|_2=\sqrt{|3+\mathrm{i}|^2+|-\mathrm{i}|^2+2^2+|1-\mathrm{i}|^2}=\sqrt{17}$，所以

$$\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}=(\frac{3+\mathrm{i}}{\sqrt{17}},\frac{-\mathrm{i}}{\sqrt{17}},\frac{2}{\sqrt{17}},\frac{1-\mathrm{i}}{\sqrt{17}})^{\mathrm{T}}$$

3．向量的夹角与正交

定义　非零向量$\boldsymbol{x}$和$\boldsymbol{y}$的夹角$\langle\boldsymbol{x},\boldsymbol{y}\rangle$规定为

$$\cos^2\langle\boldsymbol{x},\boldsymbol{y}\rangle=\frac{(\boldsymbol{x},\boldsymbol{y})(\boldsymbol{y},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})(\boldsymbol{y},\boldsymbol{y})}=\frac{|(\boldsymbol{x},\boldsymbol{y})|^2}{\|\boldsymbol{x}\|_2^2\|\boldsymbol{y}\|_2^2}\qquad 0\le\langle\boldsymbol{x},\boldsymbol{y}\rangle\le\pi$$

（对于实向量为 $\cos\langle\boldsymbol{x},\boldsymbol{y}\rangle=\dfrac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2}$，　$0\le\langle\boldsymbol{x},\boldsymbol{y}\rangle\le\pi$ ）。

定义 设 $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{C}^n$，若 $(\boldsymbol{x}, \boldsymbol{y})=0$，则称 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ **垂直**或**正交**，记为 $\boldsymbol{x} \perp \boldsymbol{y}$。如果一组非零向量两两正交，则称之为**正交向量组**。

关于正交向量与正交向量组有如下几个性质：

性质 1 若向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 正交，则 $\|\boldsymbol{x}+\boldsymbol{y}\|_2^2 = \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2$（称之为**商高定理**或**勾股弦定理**）。

证 $\|\boldsymbol{x}+\boldsymbol{y}\|_2^2 = (\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{x},\boldsymbol{x}) + (\boldsymbol{x},\boldsymbol{y}) + (\boldsymbol{y},\boldsymbol{x}) + (\boldsymbol{y},\boldsymbol{y}) = \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2$。**证毕**

更进一步，若 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 是两两正交的向量，则

$$\|\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_m\|_2^2 = \|\boldsymbol{x}_1\|_2^2 + \|\boldsymbol{x}_2\|_2^2 + \cdots + \|\boldsymbol{x}_m\|_2^2$$

性质 2 正交向量组必线性无关。

证 设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 是正交向量组，若有一组数 $k_1, k_2, \cdots, k_m$ 使

$$k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_m\boldsymbol{x}_m = \mathbf{0}$$

两边与 $\boldsymbol{x}_i$ 作内积，并利用正交性得 $k_i(\boldsymbol{x}_i, \boldsymbol{x}_i) = 0 \quad (i = 1,2,\cdots,m)$

由于 $\boldsymbol{x}_i \neq \mathbf{0}$，故 $(\boldsymbol{x}_i, \boldsymbol{x}_i) > 0$，从而 $k_i = 0 \ (i = 1,2,\cdots,m)$。**证毕**

4. Gram-Schmidt 正交化过程

已知一组线性无关的向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \in \mathbf{C}^n$，按照下述公式

$$\boldsymbol{y}_1 = \boldsymbol{x}_1, \quad \boldsymbol{y}_k = \boldsymbol{x}_k - \frac{(\boldsymbol{x}_k, \boldsymbol{y}_1)}{(\boldsymbol{y}_1, \boldsymbol{y}_1)}\boldsymbol{y}_1 - \cdots - \frac{(\boldsymbol{x}_k, \boldsymbol{y}_{k-1})}{(\boldsymbol{y}_{k-1}, \boldsymbol{y}_{k-1})}\boldsymbol{y}_{k-1} \quad (k = 2,3,\cdots,m)$$

即构造出正交向量组 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m$，这一过程称为 **Gram-Schmidt 正交化过程**。

例 试把向量组 $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ \mathrm{i} \\ 0 \end{pmatrix}$，$\boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix}$，$\boldsymbol{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ 正交化。

解 $\boldsymbol{y}_1 = \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ \mathrm{i} \\ 0 \end{pmatrix}$，$\boldsymbol{y}_2 = \boldsymbol{x}_2 - \frac{(\boldsymbol{x}_2, \boldsymbol{y}_1)}{(\boldsymbol{y}_1, \boldsymbol{y}_1)}\boldsymbol{y}_1 = \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ \mathrm{i} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{\mathrm{i}}{2} \\ \mathrm{i} \end{pmatrix}$，

$$\boldsymbol{y}_3 = \boldsymbol{x}_3 - \frac{(\boldsymbol{x}_3, \boldsymbol{y}_1)}{(\boldsymbol{y}_1, \boldsymbol{y}_1)}\boldsymbol{y}_1 - \frac{(\boldsymbol{x}_3, \boldsymbol{y}_2)}{(\boldsymbol{y}_2, \boldsymbol{y}_2)}\boldsymbol{y}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{0}{2}\boldsymbol{y}_1 - \frac{-\mathrm{i}}{\frac{3}{2}}\begin{pmatrix} \frac{1}{2} \\ -\frac{\mathrm{i}}{2} \\ i \end{pmatrix} = \begin{pmatrix} \frac{\mathrm{i}}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

则 $\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3$ 是正交向量组。

5. 酉矩阵

定义　若 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 满足 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{I}$ （或 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{H}}$，或 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{I}$），则称 $\boldsymbol{A}$ 为**酉矩阵**。

性质 1　设 $\boldsymbol{A}$ 是酉矩阵，则 $|\det\boldsymbol{A}|=1$。

证　因 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{I}$，两边取行列式得

$$1=\det\boldsymbol{I}=\det(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\det\boldsymbol{A}^{\mathrm{H}}\det\boldsymbol{A}=\overline{\det\boldsymbol{A}}\det\boldsymbol{A}=|\det\boldsymbol{A}|^2$$ **证毕**

性质 2　设 $\boldsymbol{A}$ 是酉矩阵，则 $\boldsymbol{A}^{-1}$ 也是酉矩阵。

证　$(\boldsymbol{A}^{-1})^{\mathrm{H}}\boldsymbol{A}^{-1}=(\boldsymbol{A}^{\mathrm{H}})^{-1}\boldsymbol{A}^{-1}=(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1}=\boldsymbol{I}^{-1}=\boldsymbol{I}$ **证毕**

性质 3　设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶酉矩阵，则 $\boldsymbol{AB}$ 也是酉矩阵。

证　$(\boldsymbol{AB})^{\mathrm{H}}(\boldsymbol{AB})=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{AB}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}=\boldsymbol{I}$ **证毕**

性质 4 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 是酉矩阵，则对任意 $\boldsymbol{x}\in\mathbf{C}^n$ 有 $\|\boldsymbol{Ax}\|_2=\|\boldsymbol{x}\|_2$。

证　$\|\boldsymbol{Ax}\|_2=\sqrt{(\boldsymbol{Ax},\boldsymbol{Ax})}=\sqrt{(\boldsymbol{Ax})^{\mathrm{H}}(\boldsymbol{Ax})}=\sqrt{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Ax}}=\sqrt{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}}=\|\boldsymbol{x}\|_2$ **证毕**

（称这一性质为 2-范数的**酉不变性**）。

性质 5　n 阶方阵 $\boldsymbol{A}$ 是酉矩阵的充要条件是，$\boldsymbol{A}$ 的列向量是两两正交的单位向量。

证　设 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n)$，即将 $\boldsymbol{A}$ 按列分块，则由

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{a}_1^{\mathrm{H}}\\ \boldsymbol{a}_2^{\mathrm{H}}\\ \vdots\\ \boldsymbol{a}_n^{\mathrm{H}}\end{pmatrix}(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n)=\begin{pmatrix}\boldsymbol{a}_1^{\mathrm{H}}\boldsymbol{a}_1 & \boldsymbol{a}_1^{\mathrm{H}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1^{\mathrm{H}}\boldsymbol{a}_n\\ \boldsymbol{a}_2^{\mathrm{H}}\boldsymbol{a}_1 & \boldsymbol{a}_2^{\mathrm{H}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2^{\mathrm{H}}\boldsymbol{a}_n\\ \vdots & \vdots & & \vdots\\ \boldsymbol{a}_n^{\mathrm{H}}\boldsymbol{a}_1 & \boldsymbol{a}_n^{\mathrm{H}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n^{\mathrm{H}}\boldsymbol{a}_n\end{pmatrix}=$$

$$\begin{pmatrix}(\boldsymbol{a}_1,\boldsymbol{a}_1) & (\boldsymbol{a}_2,\boldsymbol{a}_1) & \cdots & (\boldsymbol{a}_n,\boldsymbol{a}_1)\\ (\boldsymbol{a}_1,\boldsymbol{a}_2) & (\boldsymbol{a}_2,\boldsymbol{a}_2) & \cdots & (\boldsymbol{a}_n,\boldsymbol{a}_2)\\ \vdots & \vdots & & \vdots\\ (\boldsymbol{a}_1,\boldsymbol{a}_n) & (\boldsymbol{a}_2,\boldsymbol{a}_n) & \cdots & (\boldsymbol{a}_n,\boldsymbol{a}_n)\end{pmatrix}$$

即知 $$\boldsymbol{A}\text{ 是酉矩阵}\Leftrightarrow(\boldsymbol{a}_i,\boldsymbol{a}_j)=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$$ **证毕**

性质 5 给出了判定和构造酉矩阵的一种方法。

例 矩阵$\boldsymbol{A}=\begin{pmatrix}1&1&0\\ \mathrm{i}&0&0\\0&\mathrm{i}&1\end{pmatrix}$是否酉矩阵？若不是，试利用其列向量构造一个酉矩阵。

解 法 1．因为$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{pmatrix}2&1&0\\1&2&\mathrm{i}\\0&-\mathrm{i}&1\end{pmatrix}\neq\boldsymbol{I}$，所以$\boldsymbol{A}$不是酉矩阵。

法 2．设$\boldsymbol{a}_1=\begin{pmatrix}1\\ \mathrm{i}\\0\end{pmatrix}$，$\boldsymbol{a}_2=\begin{pmatrix}1\\0\\ \mathrm{i}\end{pmatrix}$，$\boldsymbol{a}_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$，因为$\|\boldsymbol{a}_1\|_2=\sqrt{2}$，所以$\boldsymbol{A}$不是酉矩阵。

利用 Gram-Schmidt 正交化过程构造正交向量组

$$\boldsymbol{p}_1=\begin{pmatrix}1\\ \mathrm{i}\\0\end{pmatrix},\quad \boldsymbol{p}_2=\begin{pmatrix}\frac{1}{2}\\-\frac{\mathrm{i}}{2}\\ \mathrm{i}\end{pmatrix},\quad \boldsymbol{p}_3=\begin{pmatrix}\frac{\mathrm{i}}{3}\\ \frac{1}{3}\\ \frac{1}{3}\end{pmatrix},$$

单位化得

$$\boldsymbol{q}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{\mathrm{i}}{\sqrt{2}}\\0\end{pmatrix},\quad \boldsymbol{q}_2=\begin{pmatrix}\frac{1}{\sqrt{6}}\\-\frac{\mathrm{i}}{\sqrt{6}}\\ \frac{2\mathrm{i}}{\sqrt{6}}\end{pmatrix},\quad \boldsymbol{q}_3=\begin{pmatrix}\frac{\mathrm{i}}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\end{pmatrix},$$

故 $\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&\frac{\mathrm{i}}{\sqrt{3}}\\ \frac{\mathrm{i}}{\sqrt{2}}&-\frac{\mathrm{i}}{\sqrt{6}}&\frac{1}{\sqrt{3}}\\0&\frac{2\mathrm{i}}{\sqrt{2}}&\frac{1}{\sqrt{3}}\end{pmatrix}$是一个酉矩阵。

二、 Schur 定理

定理（**Schur**） 设$\boldsymbol{A}\in\mathbf{C}^{n\times n}$，则$\boldsymbol{A}$可酉相似于上三角矩阵$\boldsymbol{T}$，即存在$n$阶酉矩阵$\boldsymbol{U}$，使得$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{T}$。

证 对阶数n用归纳法。$n=1$时，$\boldsymbol{A}$本身就是一个上三角阵，取$\boldsymbol{U}=[1]$ 即知结论成立。设对$n-1$阶矩阵结论成立，考虑n阶矩阵$\boldsymbol{A}$。设$\boldsymbol{A}\boldsymbol{u}_1=\lambda_1\boldsymbol{u}_1$且$\|\boldsymbol{u}_1\|_2=1$（$\lambda_1$是$\boldsymbol{A}$的特征值，$\boldsymbol{u}_1$是对应$\lambda_1$的特征向量，可差一倍数），以$\boldsymbol{u}_1$为第一列构造$n$阶酉矩阵$\boldsymbol{U}_1=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n)$，其中 $\|\boldsymbol{u}_i\|_2=1$,且$\boldsymbol{u}_i{}^{\mathrm{H}}\boldsymbol{u}_j=0\,(i\neq j)$，则

$$\boldsymbol{U}_1{}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1=\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{H}}\\ \boldsymbol{u}_2^{\mathrm{H}}\\ \vdots\\ \boldsymbol{u}_n^{\mathrm{H}}\end{pmatrix}\boldsymbol{A}(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n)=\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_1 & \boldsymbol{u}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_n\\ \boldsymbol{u}_2^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_1 & \boldsymbol{u}_2^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_2^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_n\\ \vdots & \vdots & & \vdots\\ \boldsymbol{u}_n^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_1 & \boldsymbol{u}_n^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n^{\mathrm{H}}\boldsymbol{A}\boldsymbol{u}_n\end{pmatrix}=$$

$$\left(\begin{array}{c|ccc}\lambda_1 & * & \cdots & *\\ \hline 0 & & & \\ \vdots & & \boldsymbol{A}_1 & \\ 0 & & & \end{array}\right),\qquad \text{其中 } \boldsymbol{A}_1\in\mathbf{C}^{(n-1)\times(n-1)}\text{。}$$

由归纳假设知，存在$n-1$阶酉矩阵$\tilde{\boldsymbol{U}}_2$使 $\tilde{\boldsymbol{U}}_2{}^{\mathrm{H}}\boldsymbol{A}_1\tilde{\boldsymbol{U}}_2=\begin{pmatrix}\lambda_2 & & *\\ & \ddots & \\ & & \lambda_{n-1}\end{pmatrix}$，令

$\boldsymbol{U}_2=\begin{pmatrix}1 & \\ & \tilde{\boldsymbol{U}}_2\end{pmatrix}$，易知$\boldsymbol{U}_2$是$n$阶酉矩阵，且

$$\boldsymbol{U}_2^{\mathrm{H}}(\boldsymbol{U}_1{}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1)\boldsymbol{U}_2=\begin{pmatrix}1 & \\ & \tilde{\boldsymbol{U}}_2{}^{\mathrm{H}}\end{pmatrix}\left(\begin{array}{c|ccc}\lambda_1 & * & \cdots & *\\ \hline 0 & & & \\ \vdots & & \boldsymbol{A}_1 & \\ 0 & & & \end{array}\right)\begin{pmatrix}1 & \\ & \tilde{\boldsymbol{U}}_2\end{pmatrix}=$$

$$\left(\begin{array}{c|ccc}\lambda_1 & * & \cdots & *\\ \hline 0 & & & \\ \vdots & & \tilde{\boldsymbol{U}}_2{}^{\mathrm{H}}\boldsymbol{A}_1\tilde{\boldsymbol{U}}_2 & \\ 0 & & & \end{array}\right)=\left(\begin{array}{c|ccc}\lambda_1 & & * & \\ \hline & \lambda_2 & & *\\ & & \ddots & \\ & & & \lambda_n\end{array}\right)=\boldsymbol{T}$$

令 $\boldsymbol{U}=\boldsymbol{U}_1\boldsymbol{U}_2$，则它是$n$阶酉矩阵，且 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{T}$。**证毕**

三、 酉相似于对角阵的矩阵

由 Schur 定理自然会想到，什么样的矩阵可以酉相似于对角阵呢？

1．正规矩阵

定义 若$\boldsymbol{A}\in\mathbf{C}^{n\times n}$满足$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$，则称$\boldsymbol{A}$为**正规矩阵**。

容易验证：**实对称矩阵**（$\boldsymbol{A}\in\mathbf{R}^{n\times n}$且$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$）；**实反对称矩阵**（$\boldsymbol{A}\in\mathbf{R}^{n\times n}$且$\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$）；**Hermite 矩阵**（$\boldsymbol{A}\in\mathbf{C}^{n\times n}$且$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$，也称**共轭对称阵**）；**反 Hermite 矩阵**（$\boldsymbol{A}\in\mathbf{C}^{n\times n}$且$\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$）；**酉矩阵**（$\boldsymbol{A}\in\mathbf{C}^{n\times n}$且$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{I}$)；**正交矩阵**（$\boldsymbol{A}\in\mathbf{R}^{n\times n}$且$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$) 等均是正规矩阵。这说明正规矩阵包括了很大一类常用的矩阵。

定理 矩阵 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 酉相似于对角阵的充要条件是 $\boldsymbol{A}$ 为正规矩阵。

证 必要性。已知存在酉矩阵 $\boldsymbol{U}$，使 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{\varLambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，即 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\varLambda}\boldsymbol{U}^{\mathrm{H}}$，从而

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\overline{\lambda_1}&&\\&\ddots&\\&&\overline{\lambda_n}\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}=$$

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

即 $\boldsymbol{A}$ 是正规矩阵。

充分性。已知 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$，由 Schur 定理知，存在酉矩阵 $\boldsymbol{U}$，使 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{T}$，其中 $\boldsymbol{T}=(t_{ij})$ 是上三角阵，于是

$$\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}=\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}$$

可见 $\boldsymbol{T}$ 也是正规矩阵。下面证明正规三角矩阵必是对角矩阵。因为

$$\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}=\begin{pmatrix}\overline{t_{11}}&&&\\\overline{t_{12}}&\overline{t_{22}}&&\\\vdots&\vdots&\ddots&\\\overline{t_{1n}}&\overline{t_{2n}}&\cdots&\overline{t_{nn}}\end{pmatrix}\begin{pmatrix}t_{11}&t_{12}&\cdots&t_{1n}\\&t_{22}&\cdots&t_{2n}\\&&\ddots&\vdots\\&&&t_{nn}\end{pmatrix}=$$

$$\begin{pmatrix}|t_{11}|^2&*&\cdots&*\\ *&|t_{12}|^2+|t_{22}|^2&\cdots&*\\\vdots&\vdots&\ddots&\vdots\\ *&*&\cdots&|t_{1n}|^2+|t_{2n}|^2+\cdots+|t_{nn}|^2\end{pmatrix}$$

$$\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}=\begin{pmatrix}t_{11}&t_{12}&\cdots&t_{1n}\\&t_{22}&\cdots&t_{2n}\\&&\ddots&\vdots\\&&&t_{nn}\end{pmatrix}\begin{pmatrix}\overline{t_{11}}&&&\\\overline{t_{12}}&\overline{t_{22}}&&\\\vdots&\vdots&\ddots&\\\overline{t_{1n}}&\overline{t_{2n}}&\cdots&\overline{t_{nn}}\end{pmatrix}=$$

$$\begin{pmatrix}|t_{11}|^2+|t_{12}|^2+\cdots+|t_{1n}|^2&*&\cdots&*\\ *&|t_{22}|^2+\cdots+|t_{2n}|^2&\cdots&*\\\vdots&\vdots&\ddots&\vdots\\ *&*&\cdots&|t_{nn}|^2\end{pmatrix}$$

由 $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}=\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}$ 得（只比较对角元素）

$$\begin{cases} |t_{11}|^2=|t_{11}|^2+|t_{12}|^2+\cdots+|t_{1n}|^2 & \Rightarrow t_{12}=\cdots=t_{1n}=0 \\ |t_{12}|^2+|t_{22}|^2=|t_{22}|^2+\cdots+|t_{2n}|^2 & \Rightarrow t_{23}=\cdots=t_{2n}=0 \\ \cdots\cdots & \Rightarrow\cdots\cdots \\ |t_{1n}|^2+\cdots+|t_{nn}|^2=|t_{nn}|^2 & \Rightarrow t_{n-1,n}=0 \end{cases}$$

故 $\boldsymbol{T}$ 为对角阵。**证毕**

推论 1 Hermite 矩阵的特征值全为实数；反 Hermite 矩阵的特征值为零或纯虚数。

证 设 $\boldsymbol{A}$ 是 Hermite 矩阵，从而是正规矩阵，于是存在酉矩阵 $\boldsymbol{U}$ 使得

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

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

$$=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

得 $\lambda_i=\overline{\lambda_i}\quad(i=1,2,\cdots,n)$，故 λ_i 均为实数。

如果 $\boldsymbol{A}$ 是反 Hermite 矩阵，即 $\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$，同上可推得 $\overline{\lambda_i}=-\lambda_i\quad(i=1,2,\cdots,n)$，从而 λ_i 为零或纯虚数。**证毕**

推论 2 实对称阵的特征值全为实数；实反对称阵的特征值为零或纯虚数。

推论 3 设 $\boldsymbol{A}$ 是 n 阶正规矩阵，λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 是对应的特征向量，则 $\boldsymbol{A}^{\mathrm{H}}$ 的特征值是 $\overline{\lambda}$，对应的特征向量仍是 $\boldsymbol{x}$。

证 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，取共轭转置得

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

记 $\boldsymbol{U}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n)$，则有 $\boldsymbol{A}\boldsymbol{u}_j=\lambda_j\boldsymbol{u}_j$ 和 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{u}_j=\overline{\lambda_j}\boldsymbol{u}_j$。**证毕**

推论 4 正规矩阵的不同特征值对应的特征向量正交。

证 设 $\boldsymbol{A}$ 是正规矩阵，$\boldsymbol{A}\boldsymbol{x}=\lambda\,\boldsymbol{x}$，$\boldsymbol{A}\boldsymbol{y}=\mu\,\boldsymbol{y}$，$\boldsymbol{x}\neq\boldsymbol{0}$，$\boldsymbol{y}\neq\boldsymbol{0}$，$\lambda\neq\mu$，则有 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}=\overline{\lambda}\boldsymbol{x}^{\mathrm{H}}$，右乘 $\boldsymbol{y}$ 得 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{y}=\overline{\lambda}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$。注意到 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{y}=\overline{\mu}\boldsymbol{y}$，得 $\overline{\mu}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}=\overline{\lambda}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$，即 $\overline{(\lambda-\mu)}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}=0$。由 $\lambda\neq\mu$. 得 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}=0$，即 $\boldsymbol{x}\perp\boldsymbol{y}$。**证毕**

2. 正规矩阵酉相似于对角阵的计算

正规矩阵 $\boldsymbol{A}$ 酉相似于对角阵的计算步骤如下：

1） 求 $\boldsymbol{A}$ 的特征值和特征向量；
2） 将属于重特征值的线性无关特征向量正交化（单根时绕过这一步）；
3） 将全部特征向量单位化；
4）将特征值排成对角阵 $\boldsymbol{\Lambda}$，相应的特征向量排成矩阵 $\boldsymbol{U}$，则有 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{\Lambda}$。

例 矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & 0\end{pmatrix}$ 是否正规矩阵？若是，试将其酉相似对角化。

解 因为 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$，即 $\boldsymbol{A}$ 是实反对称阵，所以 $\boldsymbol{A}$ 是正规矩阵。又

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda & -1 & 0\\ 1 & \lambda & -1\\ 0 & 1 & \lambda\end{vmatrix}=\lambda(\lambda^2+2),$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=0, \lambda_2=\sqrt{2}\,\mathrm{i}, \lambda_3=-\sqrt{2}\,\mathrm{i}$。可求得对应的特征向量为

$$\boldsymbol{x}_1=\begin{pmatrix}1\\0\\1\end{pmatrix},\quad \boldsymbol{x}_2=\begin{pmatrix}1\\ \sqrt{2}\,\mathrm{i}\\ -1\end{pmatrix},\quad \boldsymbol{x}_3=\begin{pmatrix}1\\ -\sqrt{2}\,\mathrm{i}\\ -1\end{pmatrix}$$

它们已正交；单位化得 $\begin{pmatrix}\frac{1}{\sqrt{2}}\\ 0\\ \frac{1}{\sqrt{2}}\end{pmatrix}$, $\begin{pmatrix}\frac{1}{2}\\ \frac{\mathrm{i}}{\sqrt{2}}\\ -\frac{1}{2}\end{pmatrix}$, $\begin{pmatrix}\frac{1}{2}\\ -\frac{\mathrm{i}}{\sqrt{2}}\\ -\frac{1}{2}\end{pmatrix}$,

故酉矩阵 $\boldsymbol{U}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2}\\ 0 & \frac{\mathrm{i}}{\sqrt{2}} & -\frac{\mathrm{i}}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & -\frac{1}{2}\end{pmatrix}$，使 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\mathrm{diag}(0,\sqrt{2}\,\mathrm{i},-\sqrt{2}\,\mathrm{i})$。

定理 实对称矩阵必可正交相似于对角阵。

例 已知实对称阵 $\boldsymbol{A}=\begin{pmatrix}3 & 1 & 0 & -1\\ 1 & 3 & -1 & 0\\ 0 & -1 & 3 & 1\\ -1 & 0 & 1 & 3\end{pmatrix}$，求正交阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ 为对角阵。

解 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-1)(\lambda-3)^2(\lambda-5)$，

$\boldsymbol{A}$ 的特征值为 $\lambda_1=1, \lambda_2=\lambda_3=3, \lambda_4=5$。

对应 $\lambda_1=1$ 的特征向量为 $(1,-1,-1,1)^{\mathrm{T}}$，单位化得 $(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2})^{\mathrm{T}}$；

对应 $\lambda_2=\lambda_3=3$ 的特征向量为 $(1,0,1,0)^{\mathrm{T}}$，$(0,1,0,1)^{\mathrm{T}}$，它们已正交，单位化得

$$(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}, 0)^{\mathrm{T}}, \quad (0, \tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}})^{\mathrm{T}}$$

对应于$\lambda_4 = 5$的特征向量为 $(1, 1, -1, -1)^{\mathrm{T}}$，单位化得 $(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2})^{\mathrm{T}}$，

故正交矩阵 $\boldsymbol{Q} = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix}$，使 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \begin{pmatrix} 1 & & & \\ & 3 & & \\ & & 3 & \\ & & & 5 \end{pmatrix}$。