2.2 特殊矩阵的特征系统

本节将介绍理论上和特征系统计算上的非常重要的矩阵分解，即**Schur**分解。

定理 2.7 (Schur定理) 设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，则存在酉阵 $\boldsymbol{U} \in \mathbf{C}^{n\times n}$ 使得

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^H$$

其中 $\boldsymbol{R} \in \mathbf{C}^{n\times n}$ 为上三角矩阵。

$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^H$ 也称为**矩阵的Schur分解**

证：对矩阵的阶数$\boldsymbol{n}$用数学归纳法证明。

$\boldsymbol{n}$=1时， 定理显然成立。即 $\boldsymbol{A}=(a)=\boldsymbol{R}$, $\boldsymbol{U}=(1)$

设$n=k$时，定理成立， 即 对于 k 阶方阵 $\boldsymbol{A}$，存在k阶酉阵$\boldsymbol{U}$，使得 $\boldsymbol{A}=\boldsymbol{URU}^{H}$ 现在证明，当 $n=k+1$时，定理仍成立。

记 λ_1 为方阵 $A \in \mathbf{C}^{(k+1)\times(k+1)}$ 的一个特征值，于是存在 $\boldsymbol{u}_1 \in \mathbf{C}^{k+1}$ 为其对应于 λ_1 的特征向量。我们可以特别选取 $\boldsymbol{u}_1$，使得 $\|\boldsymbol{u}_1\| = 1$ 。再于 $\mathrm{Span}\{\boldsymbol{u}_1\} \subset \mathbf{C}^{k+1}$ 的正交补空间上选取标准正交基 $\boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1} \in \mathbf{C}^{k+1}$，使

$\boldsymbol{u}_1$ 构成一组标准正交基。

注

记 $U_1=(u_1,u_2,\cdots,u_{k+1})$, 显然 $U_1\in \mathbf{C}^{(k+1)\times(k+1)}$ 为酉阵，

由于 $Au_1=\lambda_1 u_1$，则

$$U_1^H A u_1=\begin{pmatrix} u_1^H\\ u_2^H\\ \vdots\\ u_{k+1}^H\end{pmatrix} A u_1=\begin{pmatrix} u_1^H\\ u_2^H\\ \vdots\\ u_{k+1}^H\end{pmatrix}\lambda_1 u_1=\lambda_1\begin{pmatrix}(u_1,u_1)\\ (u_2,u_1)\\ \vdots\\ (u_{k+1},u_1)\end{pmatrix}$$

由归纳法假设， $n=k$ 时结论成立， 即存在酉阵 $\tilde{\boldsymbol{U}}_2 \in \mathbf{C}^{k\times k}$, 使得

$$\tilde{\boldsymbol{U}}_2^H \boldsymbol{A}_1 \tilde{\boldsymbol{U}}_2 = \begin{pmatrix} \lambda_2 & r_{23} & \cdots & r_{2k+1} \\ 0 & \lambda_3 & \cdots & r_{3k+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{k+1} \end{pmatrix} = \boldsymbol{R}_1$$

上三角矩阵

即
$$\boldsymbol{A}_1 = \tilde{\boldsymbol{U}}_2 \boldsymbol{R}_1 \tilde{\boldsymbol{U}}_2^H \qquad (2\text{-}37)$$

其中 $\boldsymbol{R}_1 \in \mathbf{C}^{k\times k}$ 为上三角矩阵。

现取 $$\boldsymbol{U}_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{\boldsymbol{U}}_2 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \tilde{\boldsymbol{U}}_2 \end{pmatrix}$$ ， 则

$$U_2^H\left(U_1^H A U_1\right)U_2 = U_2^H \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1k+1} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} U_2$$

$$= \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1k+1} \\ 0 & & & \\ \vdots & & \tilde{U}_2^H A_1 \tilde{U}_2^H & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1k+1} \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{k+1} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \tilde{c}^T \\ 0 & R_1 \end{pmatrix} = R$$

取 $\boldsymbol{U}=\boldsymbol{U}_1\boldsymbol{U}_2$，显然 $\boldsymbol{U}\in\mathbf{C}^{(k+1)\times(k+1)}$ 为酉阵，且

$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U}=\boldsymbol{R}$， $\boldsymbol{R}\in\mathbf{C}^{(k+1)\times(k+1)}$ 为上三角矩阵， 则

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^H \qquad \textbf{(2-35)}$$

成立，即证明了 $n=k+1$ 时定理成立。 **证毕**

称（**2-35**）为矩阵的**Schur**分解。

下面给出Schur定理的几点注记

1. 在矩阵的Schur分解中，由于A和R是酉相似的($A=URU^H$, $U^HAU=R$)， 因此具有相同的特征值，而上三角矩阵的特征值即为其对角元，因此，**Schur**定理还可以表示为：

任意n阶方阵酉相似于一个以其特征值为对角元的上三角矩阵R

2. 通常称R为A的**Schur**标准型，在理论上我们得到了矩阵的特征值。但是，特征值的计算一般必须采用迭代法，通常我们无法在有限步内，准确地得到。

3. 由于实矩阵A的特征值可能是一个复数，如

$$A=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\quad \det(\lambda I-A)=\begin{vmatrix}\lambda & 1\\ -1 & \lambda\end{vmatrix}=\lambda^2+1=0, \Longrightarrow \begin{cases}\lambda_1=i\\ \lambda_2=-i\end{cases}$$

因此即使矩阵A是实矩阵，一般地， Schur分解中的U，R有可能是复的。（因为当 A为实矩阵, U为正交矩阵时，那么, $U^TAU=R$ 也为实矩阵, 若R为上三角矩阵，则R的对角元素（实的）为A的特征值，这与 A的特征值可能是一个复数相矛盾）。

4. U的列向量未必都是A的特征向量，尽管其第一列时A的特征向量。

矩阵的**Schur**分解还有许多应用, 在范数的性质的研究中, 用它可以证明如下定理。

定理2.8 设$\boldsymbol{A}$为$\boldsymbol{n}$阶方阵，$\varepsilon>0$, 则在 $\mathbf{C}^{n\times n}$ 中存在一种矩阵范数 $\|\cdot\|_M$（依赖矩阵**A**和常数 ε），满足 $\|\boldsymbol{I}_n\|_M=1$，并且

$$\|\boldsymbol{A}\|_M \le \rho(\boldsymbol{A})+\varepsilon \qquad \textbf{(2-41)}$$

证明:(构造性证明）根据**Schur**定理, 存在$\boldsymbol{n}$阶酉阵$\boldsymbol{U}$, 使得

$$\boldsymbol{R} = \boldsymbol{U}^H \boldsymbol{A} \boldsymbol{U}$$

为上三角矩阵, 其中对角元即为矩阵**A**的特征值。

记矩阵**R**的上三角元素为$r_{ij}\ (j \geq i)$，对任意的$\varepsilon > 0$

取

$$\delta = \min\left\{1, \frac{\varepsilon}{(n+1)\max\limits_{1\leq i<j\leq n}|r_{ij}|}\right\}, \quad \boldsymbol{D} = \begin{pmatrix} 1 & & & & \\ & \delta & & & \\ & & \delta^2 & & \\ & & & \ddots & \\ & & & & \delta^{n-1} \end{pmatrix}$$

$$
\boldsymbol{R}=\begin{pmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1n} \\ & r_{22} & r_{23} & \cdots & r_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & r_{n-1n} \\ & & & & r_{nn} \end{pmatrix},\ \boldsymbol{RD}=\begin{pmatrix} r_{11} & \delta r_{12} & \delta^2 r_{13} & \cdots & \delta^{n-1} r_{1n} \\ & \delta r_{22} & \delta^2 r_{23} & \cdots & \delta^{n-1} r_{2n} \\ & & \delta^2 r_{33} & \cdots & \delta^{n-1} r_{3n} \\ & & & \ddots & \vdots \\ & & & & \delta^{n-1} r_{nn} \end{pmatrix}
$$

$$
\boldsymbol{D}^{-1}\boldsymbol{RD}=\begin{pmatrix} 1 & & & & \\ & \frac{1}{\delta} & & & \\ & & \frac{1}{\delta^2} & & \\ & & & \ddots & \\ & & & & \frac{1}{\delta^{n-1}} \end{pmatrix}
$$

$$
=\begin{pmatrix} r_{11} & r_{12}\delta & r_{13}\delta^2 & \cdots & r_{1n}\delta^{n-1} \\ & r_{22} & r_{23}\delta & \cdots & r_{2n}\delta^{n-2} \\ & & r_{33} & \cdots & \vdots \\ & & & \ddots & r_{n-1n}\delta \\ & & & & r_{nn} \end{pmatrix}
$$

于是，

$$\left\|\boldsymbol{D}^{-1}\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}\boldsymbol{D}\right\|_1 = \left\|\boldsymbol{D}^{-1}\boldsymbol{R}\boldsymbol{D}\right\|_1 = \max_{1\le i\le n}\sum_{i\le j}^{n}\left|r_{ij}\delta^{i-1}\right| \le \max_{1\le i\le n}\left(\left|r_{ii}\right| + \sum_{i<j}^{n}\delta^{i-1}\left|r_{ij}\right|\right)$$

$$\le \max_{1\le i\le n}\left|\lambda_i\right| + \max_{1\le i<j\le n}\left|r_{ij}\right|(1+\delta+\ldots+\delta^{n-2})\delta$$

$$\le \rho(\boldsymbol{A}) + (n-1)\max_{1\le i<j\le n}\left|r_{ij}\right|\delta \le \rho(\boldsymbol{A}) + \varepsilon$$

现记 $\left\|\boldsymbol{A}\right\|_M = \left\|\boldsymbol{D}^{-1}\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}\boldsymbol{D}\right\|_1 = \left\|(\boldsymbol{U}\boldsymbol{D})^{-1}\boldsymbol{A}(\boldsymbol{U}\boldsymbol{D})\right\|_1 = \left\|\boldsymbol{M}\boldsymbol{A}\boldsymbol{M}^{-1}\right\|_1$

其中 $M = (UD)^{-1}$，$\boldsymbol{A}$为任意$\boldsymbol{n}$阶方阵, 且满足

$$\left\|\boldsymbol{I}\right\|_M = \left\|\boldsymbol{M}\,\boldsymbol{I}\,\boldsymbol{M}^{-1}\right\|_1 = \left\|\boldsymbol{M}\,\boldsymbol{M}^{-1}\right\|_1 = \left\|\boldsymbol{I}\right\|_1 = 1 \text{。}$$

我们已经验证 $\|\cdot\|_M$ 为$\boldsymbol{n}$阶方阵的一种矩阵范数（见习题）

Schur定理还可以表示为：任意n阶方阵酉相似于一个以其特征值为对角元的上三角矩阵$\boldsymbol{R}$。$\boldsymbol{R}$ 通常称为$\boldsymbol{A}$的Schur**标准型**。

定义2.5 设 $A \in \mathbf{C}^{n\times n}$，若 $A^H A = AA^H$ 则称矩阵A为**正规矩阵**。

常见正规矩阵的有：

Hermite阵：$A^H = A$　　实对称矩阵：$A^T = A$

斜Hermite阵：$A^H = -A$　实反对称矩阵：$A^T = -A$

酉阵：$A^H A = AA^H = I$

正交矩阵：$A^T A = AA^T = I$

以上矩阵均为正规矩阵。

推论 2.1 设 A为n阶方阵， 则A为正规矩阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中$D \in \mathbf{C}^{n\times n}$ 为对角矩阵。

证明：充分性 $(\Leftarrow)$

由于$A=UDU^H$，则

$$A^H A=\left(UDU^H\right)^H\left(UDU^H\right)$$

$$=UD^H U^H UDU^H=U\left(D^H D\right)U^H$$

$$AA^H=\left(UDU^H\right)\left(UDU^H\right)^H$$

$$=UDU^H UD^H U^H=U\left(DD^H\right)U^H$$

而

$$D^H D = \begin{pmatrix} \bar{d}_1 & & \\ & \ddots & \\ & & \bar{d}_n \end{pmatrix} \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} = \begin{pmatrix} |d_1|^2 & & \\ & \ddots & \\ & & |d_n|^2 \end{pmatrix}$$

$$= \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \begin{pmatrix} \bar{d}_1 & & \\ & \ddots & \\ & & \bar{d}_n \end{pmatrix} = DD^H$$

故

$$A^H A = AA^H$$

即A为正规矩阵.

必要性（$\Rightarrow$）由Schur分解定理知，$A = URU^H$

$U \in \mathbf{C}^{n\times n}$ 为酉阵, R为上三角阵。那么，由假设知A为正规矩阵，即 $A^H A = AA^H \Rightarrow R^H R = RR^H$，即

R为正规矩阵。而上三角阵R正规矩阵$\Leftrightarrow$ R为对角矩阵。

事实上, 设

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix} \quad R^H = \begin{pmatrix} \bar{r}_{11} & & & \\ \bar{r}_{12} & \bar{r}_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar{r}_{1n} & \bar{r}_{2n} & \cdots & \bar{r}_{nn} \end{pmatrix}$$

再注意到,

$$\boldsymbol{R}^H\boldsymbol{R}=\begin{pmatrix} |r_{11}|^2 & * & \cdots & * \\ * & |r_{22}|^2+|r_{12}|^2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & \sum_{i=1}^{n}|r_{in}|^2 \end{pmatrix} = \boldsymbol{R}\boldsymbol{R}^H=\begin{pmatrix} \sum_{i=1}^{n}|r_{1i}|^2 & * & \cdots & * \\ * & \sum_{i=2}^{n}|r_{2i}|^2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & |r_{nn}|^2 \end{pmatrix}$$

从而可得:

$$|r_{11}|^2=|r_{11}|^2+|r_{12}|^2+\cdots+|r_{1n}|^2 \Rightarrow r_{1j}=\bar{r}_{1j}=0,\quad j=2,\cdots,n$$

$$|r_{12}|^2+|r_{22}|^2=|r_{22}|^2+|r_{23}|^2+\cdots+|r_{2n}|^2 \Rightarrow r_{2j}=\bar{r}_{2j}=0,\ j=3,\cdots,n$$

总之有: $r_{ij}=\bar{r}_{ij}=0, 1\le i<j\le n$。 即$\boldsymbol{R}$为对角矩阵。

推论 2.2 设 A为n阶方阵，则A为Hermite矩阵的充分必要条件是存在n阶酉阵U，使得

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{H}$$

其中$\boldsymbol{D} \in \mathrm{R}^{n \times n}$为对角矩阵。

证明：由推论2.1，存在 n 阶酉阵 U，使得

$A=UDU^H$ 其中D为n阶对角阵。 而 $A^H=A$，则可得

$D^H=D$，即D的对角元素均为实数。注意到，由

$$d_i = a + ib = a - ib = \overline{d_i} \Rightarrow b = 0,$$

即

$$d_i = \overline{d_i} = a$$

从而D为n阶实对角阵。

推论 2.2’ 设 A为n阶方阵，则A为斜Hermite矩阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中D是对角矩阵，且对角元素为纯虚数。

证明：由推论2.1,存在 n 阶酉阵 U，使得

$A=UDU^H$ 其中D为n阶对角阵。 而 $A^H=-A$，则可得 $D^H=-D$，即D的对角元素均为实数。注意到，由

$$d_i=a+ib=-(a-ib)=-\bar{d}_i\Rightarrow a=0,$$

即

$$d_i=\bar{d}_i=ib$$

从而D为n阶纯虚对角阵。

推论 2.3 设 A为n阶方阵， 则A为酉阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中$D \in \mathbf{C}^{n\times n}$ 为对角矩阵，其对角元的模均为1。

证明：由推论2.1，存在n阶酉阵U使得 $A=UDU^H$

而 A为酉阵，则有 $A^H A = AA^H = I$

$$\Rightarrow U\left(DD^H\right)U^H = U\left(D^H D\right)U^H = I$$

$$\Rightarrow \quad DD^H = D^H D = I$$

即

$$D^H D = DD^H = \begin{pmatrix} |d_{11}|^2 & & \\ & \ddots & \\ & & |d_{nn}|^2 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \Rightarrow |d_{ii}| = 1$$

$$i = 1, \cdots, n$$

即D为n阶对角阵，其对角元的模均为1.

矩阵的基本分类

Ⅰ 在正规矩阵的集合中，特征值均为实数的子集为Hermite矩阵的集合；矩阵的特征值的模均为1的子集为酉阵的集合

Ⅱ 一般矩阵 $\supset$ 可对角化矩阵

$\supset$ 正规矩阵

$$\supset \left\{ \begin{array}{r} \text{Hermite矩阵} \supset \text{实对称矩阵} \\ \text{酉矩阵} \supset \text{实正交矩阵} \end{array} \right\}$$

定理 设A为n阶方阵，则存在n阶酉阵U和V，使得

$$\|UA\|_F=\|AV\|_F=\|UAV\|_F=\|A\|_F$$

称之为F-范数的酉不变性。

$$\|UA\|_F^2=\mathrm{tr}\left((UA)^H(UA)\right)=\mathrm{tr}\left(A^HU^HUA\right)=\mathrm{tr}\left(A^HA\right)=\|A\|_F^2$$

$$\|AV\|_F^2=\mathrm{tr}\left((AV)^H(AV)\right)=\mathrm{tr}\left((VA)(AV)^H\right)$$

$$=\mathrm{tr}\left(AV^HVA^H\right)=\mathrm{tr}\left(AA^H\right)=\|A\|_F^2$$

最后，

$$\|UAV\|_F=\|AV\|_F=\|A\|_F$$

证明 Schur不等式:

$$\sum_{i=1}^{n}|\lambda_i| \le \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2$$

其中 λ_i 为A的特征值, 并且Schur不等式等号成立的充分必要条件是A为正规矩阵。

证: 根据**Schur**定理, 存在n阶酉阵U使得 $A = URU^H$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 = \|A\|_F^2 = \|URU^H\|_F^2 = \|R\|_F^2 \ge \sum_{i=1}^{n}|r_{ii}|^2 = \sum_{i=1}^{n}|\lambda_i|^2 = \sum_{i=1}^{n}\sum_{j=i+1}^{n}|r_{ij}|^2$$

要使得等号成立, 只需 $r_{ij} = 0, 1 \le i < j \le n$ 即D为阶对角阵, 则由推论2. 1, 可知其充分必要条件是A为正规矩阵。

返回本节