Schur 引理与正规矩阵

定义：设 $\boldsymbol{A},\boldsymbol{B}\in \boldsymbol{C}^{n\times n}$ (或 $R^{n\times n}$)，若存在 $\boldsymbol{U}\in \boldsymbol{U}^{n\times n}$ (或 $E^{n\times n}$) 使得

$$\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{B} \quad (\text{或 } \boldsymbol{U}^{T}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{B})$$

则称 A 酉相似(或正交相似)于 B .

定理(Schur引理)：任何一个 n 阶复矩阵 A 酉相似于一个上(下)三角矩阵。

证明：用数学归纳法。A的阶数为 1 时定理显然成立。现设 A的阶数为 $k-1$时定理成立，考虑 A 的阶数为 k时的情况。

取 k 阶矩阵 A 的一个特征值 λ_1 ，对应的单位特征向量为α_1 ，构造以 α_1 为第一列的 k 阶酉矩阵，

$$U_1 = [\alpha_1, \alpha_2, \cdots, \alpha_k]$$

$$AU_1 = [A\alpha_1, A\alpha_2, \cdots, A\alpha_k]$$

$$= [\lambda_1\alpha_1, A\alpha_2, \cdots, A\alpha_k]$$

因为 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 构成 C^k 的一个标准正交基，

故 $A\alpha_i = \sum_{j=1}^{k} \bar{a}_{ij}\alpha_j (i = 2,3,\cdots,k)$， 因此

$$AU_1 = [\alpha_1, \alpha_2, \cdots, \alpha_k] \begin{bmatrix} \lambda_1 & \bar{a}_{21} & \bar{a}_{31} & \cdots & \bar{a}_{k1} \\ 0 & & & & \\ \vdots & & & A_1 & \\ 0 & & & & \end{bmatrix}$$

其中 A_1 是 $k-1$ 阶矩阵，根据归纳假设，存在 $k-1$ 阶酉矩阵 W 满足

$$W^H A_1 W = R_1 \quad (\text{上三角矩阵})$$

令

$$U_2 = \begin{bmatrix} \mathbf{1} & \\ & W \end{bmatrix} \in U^{k\times k}$$

那么

$$U_2^H U_1^H A U_1 U_2 = \begin{bmatrix} \lambda_1 & \boxed{\begin{matrix} b_{21} & & b_{k1} \end{matrix}} \\ 0 & \\ \vdots & R_1 \\ 0 & \end{bmatrix}$$

注意：等号右端的三角矩阵主对角线上的元素为矩阵 A 的**全部特征值**.

定理(Schur 不等式):

设 $A \in C^{n\times n}$， $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为矩阵 A 的特征值，那么

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

其中等号成立 $\Longleftrightarrow$ A 酉相似于对角矩阵。

正规矩阵

定义: 设 $A \in C^{n \times n}$，如果 A 满足

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

那么称矩阵 A 为一个正规矩阵.

设 $A\in R^{n\times n}$ ，如果 A 同样满足

$$AA^H = A^H A \quad 即 \quad AA^T = A^T A$$

那么称矩阵 A 为一个实正规矩阵.

例：(1) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ 为实正规矩阵。

(2)

$$\begin{bmatrix} a & b & c & d \\ b & -a & d & -c \\ c & -d & -a & b \\ d & c & -b & -a \end{bmatrix}$$

其中 a,b,c,d 是不全为零的实数，容易验证这是一个实正规矩阵.

(3)

$$\begin{bmatrix} 4+3i & 4i & -6-2i \\ -4i & 4-3i & -2-6i \\ 6+2i & -2-6i & 1 \end{bmatrix}$$

是一个正规矩阵.

(4) **H**-阵，反**H**-阵，正交矩阵，酉矩阵，对角矩阵都是正规矩阵.

正规矩阵的性质与结构定理

引理 1： 设 A 是一个正规矩阵， 则与 A 酉相似的矩阵一定是正规矩阵.

引理 2： 设 A 是一个正规矩阵， 且又是三角矩阵， 则 A 必为对角矩阵.

由上述引理可以得到正规矩阵的结构定理

定理： 设 $A \in C^{n\times n}$ ， 则 A 是正规矩阵的充要条件是存在一个酉矩阵 U 使得

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

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是矩阵A 的特征值.

$$\boldsymbol{U}^H\boldsymbol{A}^H\boldsymbol{U}=\mathrm{diag}(\bar{\lambda}_1,\bar{\lambda}_2,\cdots\bar{\lambda}_n)$$

推论 1： n 阶正规矩阵有 n 个线性无关的特征向量 . (必要但不充分）

举例说明：可对角化的矩阵不一定可酉对角化.

设 $\boldsymbol{X}, \boldsymbol{Y}$ 是两个线性无关但是不正交的向量，比如

取 $\boldsymbol{P}=[\boldsymbol{X}\ \ \boldsymbol{Y}]=\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$, $\quad \boldsymbol{D}=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$,

则 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}=\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

$=\begin{bmatrix} 1 & 0 \\ 4 & -1 \end{bmatrix}$ 可对角化，但不能酉对角化

推论 2： 正规矩阵属于不同特征值的特征向量彼此正交.

例 1： 设

$$A=\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

求正交矩阵 Q 使得 $Q^{-1}AQ$ 为对角矩阵.

解：先计算矩阵的特征值

$$\left|\lambda I - A\right| = (\lambda + 1)^2(\lambda - 8)$$

其特征值为

$$\lambda_1 = \lambda_2 = -1, \lambda_3 = 8$$

对于特征值 $\lambda_1 = -1$ 解线性方程组

$$(-I - A)X = 0$$

求得其一个基础解系

$$X_1 = \left[-1, 2, 0\right]^T, \quad X_2 = \left[-1, 0, 1\right]^T$$

现在将 X_1, X_2 正交化并单位化，得到两个标准正交向量

$$\eta_1 = \left[\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0 \right]^T, \quad \eta_2 = \left[\frac{-4}{3\sqrt{5}}, \frac{2}{2\sqrt{5}}, \frac{\sqrt{5}}{3} \right]^T$$

对于特征值 $\lambda_2 = 8$ 解线性方程组

$$(8I - A)X = 0$$

求得其一个基础解系

$$X_3 = [2,1,2]^T$$

将其单位化得到一个单位向量

$$\eta_3 = \left[\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right]^T$$

将这三个标准正交向量组成矩阵

$$\boldsymbol{Q} = \left[\eta_1, \eta_2, \eta_3\right] = \begin{bmatrix} \frac{-1}{\sqrt{5}} & \frac{-4}{3\sqrt{5}} & \frac{2}{3} \\ \frac{2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} & \frac{1}{3} \\ 0 & \frac{\sqrt{5}}{3} & \frac{2}{3} \end{bmatrix}$$

则矩阵 Q 即为所求正交矩阵且有

$$Q^{-1}AQ=\begin{bmatrix} -1 & & \\ & -1 & \\ & & 8 \end{bmatrix}$$

例 2： 设

$$A=\begin{bmatrix} 4+3i & 4i & -6-2i \\ -4i & 4-3i & -2-6i \\ 6+2i & -2-6i & 1 \end{bmatrix}$$

求酉矩阵 Q 使得 $Q^{H}AQ$ 为对角矩阵.

解：先计算矩阵的特征值

$$|\lambda I - A| = (\lambda^2 + 81)(\lambda - 9)$$

其特征值为

$$\lambda_1 = -9\mathrm{i}, \quad \lambda_2 = 9\mathrm{i}, \quad \lambda_3 = 9$$

对于特征值 $\lambda_1 = -9i$ 解线性方程组

$$(-9iI - A)X = 0$$

求得其一个基础解系

$$X_1 = [-i/2, 1, 1]^T$$

现在将 X_1 单位化，得到一个单位向量

$$\eta_1 = \left[\frac{-i}{3}, \frac{2}{3}, \frac{2}{3}\right]^T$$

对于特征值 $\lambda_2 = 9i$ 解线性方程组

$$(9iI - A)X = 0$$

求得其一个基础解系

$$X_2 = [-i, -1/2, 1]^T$$

将其单位化得到一个单位向量

$$\eta_2 = \left[\frac{2i}{3}, \frac{-1}{3}, \frac{2}{3}\right]^T$$

对于特征值 $\lambda_3 = 9$ 解线性方程组

$$(9I - A)X = 0$$

求得其一个基础解系

$$X_3 = \left[i, 1, -1/2\right]^T$$

将其单位化得到一个单位向量

$$\eta_3 = \left[\frac{2i}{3}, \frac{2}{3}, \frac{-1}{3}\right]^T$$

将这三个标准正交向量组成矩阵

$$Q=\left[\eta_1,\eta_2,\eta_3\right]=\begin{bmatrix} \frac{-i}{3} & \frac{2i}{3} & \frac{2i}{3} \\ \frac{2}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{-1}{3} \end{bmatrix}$$

则矩阵 Q 即为所求酉矩阵且有

$$Q^H A Q = \begin{bmatrix} -9i & & \\ & 9i & \\ & & 9 \end{bmatrix}$$

定理: 设 A 是正规矩阵, 则

(1) A 是**H**-阵的充要条件是 A 的特征值为实数.

(2) A 是反**H**-阵的充要条件是 A 的特征值的实部为零.

(3) A 是**U**-阵的充要条件是 A 的特征值的模长为**1**.

注意: 正规矩阵绝不仅此三类. 比如:

(1) A 是酉矩阵，$2A$也是正规矩阵;

(2) B是反H-阵, $I+B, B+B^2$ 也是正规矩阵.

例 3: 设 A是一个反 H-阵, 证明:

$$W=(A+I)(A-I)^{-1}$$

是U-阵.

证明：因为 A 是反H-阵，特征值为零或纯虚数，所以

$\lambda_i(\boldsymbol{A})\pm 1\neq 0$, 从而 $(\boldsymbol{A}+\boldsymbol{I})$ 与 $(\boldsymbol{A}-\boldsymbol{I})$ 都是可逆矩阵。

根据U-阵的定义

$$\boldsymbol{W}\boldsymbol{W}^H=(\boldsymbol{A}+\boldsymbol{I})(\boldsymbol{A}-\boldsymbol{I})^{-1}[(\boldsymbol{A}-\boldsymbol{I})^H]^{-1}(\boldsymbol{A}+\boldsymbol{I})^H$$

A 是反 H- 阵 $\Longrightarrow$ $(\boldsymbol{A}+\boldsymbol{I})^H=-\boldsymbol{A}+\boldsymbol{I}$ ，这样

$$[(\boldsymbol{A}-\boldsymbol{I})^H]^{-1}=-(\boldsymbol{A}+\boldsymbol{I})^{-1}$$

$$
\begin{aligned}
WW^H &= (A+I)(A-I)^{-1}[(A-I)^H]^{-1}(A+I)^H \\
&= -(A+I)(A-I)^{-1}(A+I)^{-1}(A+I)^H \\
&= -(A+I)[(A+I)(A-I)]^{-1}(-A+I) \\
&= -(A+I)[(A-I)(A+I)]^{-1}(-A+I) \\
&= (A+I)(A+I)^{-1}(A-I)^{-1}(A-I) \\
&= I
\end{aligned}
$$

这说明 W 为酉矩阵.

例 4： 设A是一个n阶**H-**阵且存在自然数k使得, $\boldsymbol{A}^k = \mathbf{0}$, 证明 $\boldsymbol{A} = 0$.

证明: 由于A是**H-**矩阵, 所以存在一个酉矩阵 $U \in U^{n\times n}$ 使得

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}\boldsymbol{U}^H, \quad \lambda_i \in \boldsymbol{R}$$

于是可得

$$\boldsymbol{A}^{k}=\boldsymbol{U}\begin{bmatrix}\lambda_1^{\,\boldsymbol{k}} & & & \\ & \lambda_2^{\,\boldsymbol{k}} & & \\ & & \ddots & \\ & & & \lambda_{\boldsymbol{n}}^{\,\boldsymbol{k}}\end{bmatrix}\boldsymbol{U}^{\boldsymbol{H}}=0$$

从而

$$\lambda_{\boldsymbol{i}}^{\boldsymbol{k}}=0,\quad \lambda_{\boldsymbol{i}}\in \boldsymbol{R}$$

$\Longrightarrow$ $\lambda_i=0,\quad i=1,2,\cdots,n$ 即 $A=0.$

Hermite二次型 (Hermite二次齐次多项式)

Hermite 矩阵的基本性质

引理**:** 设 $A \in C^{n\times n}$， $A = A^H$， 则

(1) $A + A^H, AA^H, A^H A$ 都是**H-**阵**.**

(2) $A - A^H$ 是反 **H-**阵**.**

(3) 如果 A 是**H-**阵**,** 那么 A^k 也是**H-**阵**,** k 为任意正整数**.**

(4) 如果 A 是可逆的 **H-**阵**,** 那么 A^{-1} 也是可逆的**H-**阵**.**

(5) 如果 A 是 **H-**阵**(**反**H-**阵**),** 那么 iA 是反**H-**矩阵**(H-**阵**),** 这里 i 为虚数单位**.**

(6) 如果 A, B 都是 **H-**阵, 那么 $kA + lB$ 也是**H-**阵, 这里 k, l 均为实数.

定理: **(1)** 设 $\boldsymbol{A} \in \boldsymbol{C}^{n \times n}$, 则 A 是 **H-**阵的充分必要条件是对于任意的 $\boldsymbol{X} \in \boldsymbol{C}^{n}$, $\boldsymbol{X}^{H}\boldsymbol{A}\boldsymbol{X}$ 是实数。

(2) A 是 **H-**阵的充分必要条件是对于任意的 n 阶方阵 $\boldsymbol{B}$, $\boldsymbol{B}^{H}\boldsymbol{A}\boldsymbol{B}$ 为**H-**阵.

H-阵的结构定理

定理：设 $\boldsymbol{A}\in \boldsymbol{C}^{n\times n}$ ，则 A 是 H-阵的充分必要条件是存在一个酉矩阵 $\boldsymbol{U}\in \boldsymbol{U}^{n\times n}$ 使得

$$\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},\quad \lambda_1,\lambda_2,\cdots,\lambda_n\in \boldsymbol{R}$$

H-阵酉相似于实对角矩阵

推论: 实对称阵正交相似于实对角矩阵.

例 : 设 A 为一个幂等 **H**-阵, 则存在酉矩阵 $\boldsymbol{U} \in \boldsymbol{U}^{n\times n}$ 使得

$$\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_r & 0 \\ 0 & 0 \end{bmatrix}$$

证明: A 为一个 **H**-阵, 所以存在酉矩阵 $\boldsymbol{W} \in \boldsymbol{U}^{n\times n}$

使得

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

A 为一个幂等 **H-**阵**,** $\Rightarrow$ $\lambda_i = 0$ 或 $\lambda_i = 1$

将 **1** 放在一起**,** **0** 放在一起**,** 那么可找到一个酉矩阵 $\boldsymbol{U} \in \boldsymbol{U}^{n\times n}$ 使得

$$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix},\quad r \text{ 为矩阵 } A \text{ 的秩.}$$

推论：设 A 为一个 n 阶矩阵，则

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

的充分必要条件是存在一个 $n \times r$ 型次酉矩阵使得 $\boldsymbol{U}_1 \in \boldsymbol{U}_r^{n\times r}$

$$\boldsymbol{A} = \boldsymbol{U}_1 \boldsymbol{U}_1^H$$

其中 $\boldsymbol{r} = \mathbf{rank}(\boldsymbol{A})$ 。

$$\boldsymbol{U}^H \boldsymbol{A} \boldsymbol{U} = \begin{bmatrix} \boldsymbol{I}_r & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \boldsymbol{A} = \underbrace{\boldsymbol{U} \begin{bmatrix} \boldsymbol{I}_r \\ 0 \end{bmatrix}}_{\boldsymbol{U}_1} \underbrace{\begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \end{bmatrix} \boldsymbol{U}^H}_{\boldsymbol{U}_1^H}$$

关于正交投影（了解）

设 χ 是 C^n 的子空间，矩阵 U_1 的列由 χ 的标准正交基构成， 矩阵

$$P_\chi = U_1 U_1^H,$$

则线性变换 $\sigma : C^n \to \chi$

$$\sigma(z) = P_\chi z, \qquad z \in C^n$$

是 C^n 到 χ 的正交投影变换. 并且

$$\left\| z - P_\chi z \right\|_2 = \min_{x \in \chi} \left\| z - x \right\|_2, \qquad \left(\left\| x \right\|_2 \triangleq \sqrt{x^H x} \right)$$