2.2 正规矩阵及Schur分解

定理1(*Schur*引理)： 设$A \in C^{n\times n}$, 则存在酉矩阵U, 使得

$$U^H A U = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

即任一复方阵相似于一个上三角阵, 其对角元为A的特征值.

(实方阵*Schur*引理)设$A \in R^{n\times n}$, 且A的特征值均为实数, 则存在正交矩阵Q, 使得

$$Q^T AQ = Q^{-1}AQ = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

$\lambda_1, \cdots, \lambda_{\mathrm{n}}$为$A$的特征值.

定义 设$A\in C^{n\times n}$, 若$A^HA=AA^H$, 则称A为正规矩阵(规范阵).

$$
\text{正规矩阵}\begin{cases}
\text{实对称矩阵}(A^T=A, A\in R^{n\times n})\\
\text{实反对称矩阵}(A^T=-A, A\in R^{n\times n})\\
Hermite\text{阵}(A^T=A, A\in C^{n\times n})\\
\text{反}Hermite\text{阵}(A^T=-A, A\in C^{n\times n})\\
\text{正交矩阵}(A^T=A^{-1}, A\in R^{n\times n})\\
\text{酉矩阵}(A^H=A^{-1}, A\in C^{n\times n})
\end{cases}
$$

还有其他正规矩阵吗？

例: $B=U^H\begin{pmatrix}3+5i & 0\\ 0 & 2-i\end{pmatrix}U$, U为任一二阶酉阵.

由$B^H=U^H\begin{pmatrix}3-5i & 0\\ 0 & 2+i\end{pmatrix}U$易得$B^HB=BB^H$.

引理： 正规三角矩阵是对角矩阵.

证明： 设$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$, 则

$$A^H=\begin{pmatrix} \overline{a_{11}} & 0 & \cdots & 0 \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & 0 \\ & & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{pmatrix}.$$

$A^H A =$

$$\begin{pmatrix} \overline{a_{11}} & 0 & \cdots & 0 \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & 0 \\ & & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} :=$$

$(c_{ij})_{n\times n}.$

$c_{11} = \overline{a_{11}} a_{11} = |a_{11}|^2.$

而$AA^H =$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} \overline{a_{11}} & 0 & \cdots & 0 \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & 0 \\ & & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{pmatrix},$$

$c_{11} = \overline{a_{11}}a_{11} + \overline{a_{12}}a_{12} + \ldots + \overline{a_{1n}}a_{1n} = |a_{11}|^2 + \ldots + |a_{1n}|^2 \Rightarrow a_{12} = a_{13} = \ldots = a_{1n} = 0.$

同理可证$a_{ij} = 0(i \neq j)$, 即A为对角阵.

定理2: 设$A \in C^{n\times n}$, 则A是正规矩阵当且仅当A酉相似于一个对角阵, 即

$$A^H A = AA^H \Leftrightarrow \exists \text{酉阵} U \text{使得} U^H AU = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}.$$

证明:$\Longrightarrow$:由条件$A^H A = AA^H$, 由$Schur$引理, ∃酉阵U, 使$U^H AU = K$(上三角阵), 即$A = UKU^H$, 因此$A^H A = UK^H U^H UKU^H = UK^H KU^H = AA^H = UKU^H UK^H U^H = UKK^H U^H$. 所以$K^H K = KK^H$, 因为$K$为上三角阵, 由引理得$K$为对角阵.

$\Leftarrow$:由$U^HAU=\mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}$得$U^HA^HU=\mathrm{diag}\{\overline{\lambda_1},\ldots,\overline{\lambda_n}\}$, 所以$U^HAUU^HA^HU=U^HA\,A^HU=\mathrm{diag}\{|\lambda_1|^2,\ldots,|\lambda_n|^2\}U^HA^HU\,U^HAU=U^HA^HAU$, 所以$AA^H=A^HA$. 证毕.

注: 正规矩阵是单纯矩阵, 反之不然. 如$A=\begin{pmatrix}1 & 2\\ 1 & 1\end{pmatrix}$, $A^H=\begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix}$, 可知$|\lambda I-A|$无重根, A为单纯矩阵, 但$AA^H\neq A^HA$.

推论1: A为正规矩阵$\Leftrightarrow$ A有n个特征向量构成C^n的一组标基, 且A的不同特征值的特征向量正交.

证明: A正规$\Rightarrow$存在酉阵U使$U^HAU =$ $\text{diag}\{\lambda_1, \dots, \lambda_n\} \Rightarrow$ $AU = U\text{diag}\{\lambda_1, \dots, \lambda_n\}$, 令$U =$ $(x_1, \cdots, x_n)$, 则$Ax_i = \lambda_i x_i$, 即$x_i(i = 1, \dots, n)$是A的n个正交的特征向量(U是酉矩阵), 它们构成C^n的一组标基.以上过程可逆.

推论2： 设$A \in R^{n\times n}$, 则

1) A为实对称阵$\Leftrightarrow$ A的特征值$\lambda_1, \dots, \lambda_n$为实数, 且存在$Q \in R^{n\times n}$使得

$$Q^T AQ = Q^{-1}AQ = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}.$$

2) A为正交矩阵$\Leftrightarrow$ A的每个特征值λ_i的模$|\lambda_i| = 1$, 且存在酉矩阵U使得$U^H AU = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$.

证明： 1)⇒由定理2知, 存在酉阵U使$U^H AU = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, 转秩得$U^H AU = \text{diag}\{\overline{\lambda_1}, \dots, \overline{\lambda_n}\}$, 所以$\lambda_i = \overline{\lambda_i}$, λ_i为实数.

再由实数的Schur引理知, 存在正交矩阵Q, 使

$$Q^T AQ = Q^{-1}AQ = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = K,$$

因此$A = QKQ^T$, $A^T = QK^TQ^T$, 从而$K = K^T$, 即K为对角阵. 也就是$Q^T AQ = \text{diag}\{\lambda_1, \dots, \lambda_n\}$.

$\Leftarrow$由$Q^TAQ=\mathrm{diag}\{\lambda_1,\dots,\lambda_n\}$转秩得$Q^TA^TQ=\mathrm{diag}\{\lambda_1,\dots,\lambda_n\}$, 所以$Q^TAQ=Q^TA^TQ$, 因为$Q$满秩, 所以$A=A^T$.

证明: 2)$\Rightarrow$设$Ax=\lambda x$, 共轭转秩有$x^H A^H=\bar{\lambda}\, x^H$, 故

$x^H A^H A x=\bar{\lambda}\, x^H \lambda x=|\lambda|^2 x^H x$, 因为$A^H A=I$, 所以

$|\lambda|^2=1$. (因为$x^H x=\|x\|^2>0$)

$\Leftarrow$:由条件$U^H A U=\mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}$共轭转秩得

$U^H A^T U=\mathrm{diag}\{\overline{\lambda_1},\ldots,\overline{\lambda_n}\}$, 所以$U^H A\, A^T U=$

$\mathrm{diag}\{|\lambda_1|^2,\ldots,|\lambda_n|^2\}=I_n$, 所以$AA^T=I_n$.

推论2： 设$A \in R^{n\times n}$, 则

1) A为实对称阵$\Leftrightarrow$ A的特征值$\lambda_1, \dots, \lambda_n$为实数, 且存在$Q \in R^{n\times n}$使得$Q^T A Q = Q^{-1} A Q = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$.

2) A为正交矩阵$\Leftrightarrow$ A的每个特征值λ_i的模$|\lambda_i| = 1$, 且存在U使得$U^H A U = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$.

注1： 设$A \in C^{n\times n}$, 则

1) A为$Hermite$阵$\Leftrightarrow$ A的特征值$\lambda_1, \ldots, \lambda_n$为实数, 酉相似于diag$\{\lambda_1, \ldots, \lambda_n\}$.

2) A为酉矩阵$\Leftrightarrow$ A的每个特征值λ_i的模$|\lambda_i| = 1$, 且酉相似于对角阵.

注2: 对于实对称阵A, 正交相似于对角阵, 求正交阵步骤:

1) 求出A的全部互异特征值$\lambda_1, \ldots, \lambda_s$, λ_i的重数为r_i, $\sum_{i=1}^{s} r_i = n$;

2) 对每个λ_i, 求$(\lambda_i I - A)X = 0$的一个基础解系: $a_i^1, \ldots, a_i^{r_i}$, 将其标准化为$\eta_i^1, \ldots, \eta_i^{r_i}$;

3) 令$Q = (\eta_1^1, \ldots, \eta_1^{r_i}, \ldots, \eta_s^1, \ldots, \eta_s^{r_s})$, 显然$Q$为正交阵,

且$Q^T A Q = \mathrm{diag}\{\underbrace{\lambda_1, \ldots, \lambda_1}_{r_1}, \ldots, \underbrace{\lambda_s, \ldots, \lambda_s}_{r_s}\}$.

例: 设$B=\begin{pmatrix}-1 & -3 & 3 & -3\\ -3 & -1 & -3 & 3\\ 3 & -3 & -1 & -3\\ -3 & 3 & -3 & -1\end{pmatrix}$, 求正交阵$Q$, 使$Q^TBQ$为对角阵.

解: $|\lambda I-B|=(\lambda-8)(\lambda+4)^3$, 所以$\lambda=8,\ -4$ (三重).

对$\lambda=8$, 得特征向量为$a_1=(-1,1,-1,1)^T$.

对$\lambda=-4$, 得特征向量为$a_2=(-1,1,0,0)^T, a_3=(1,0,-1,0)^T, a_4=(1,0,0,1)^T$.

所以$E(8)=\mathrm{span}\{a_1\}, E(-4)=\mathrm{span}\{a_2,a_3,a_4\}$,

且$E(8) \perp E(-4)$. 标准化得

$$\eta_1 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2})^T,$$

$$\eta_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)^T,$$

$$\eta_3 = (-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0)^T,$$

$$\eta_4 = (\frac{1}{\sqrt{12}}, -\frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{3}{\sqrt{12}})^T.$$

令$Q = (\eta_1, \eta_2, \eta_3, \eta_4)$, 则$Q^T B Q = \mathrm{diag}\{8, -4, -4, -4\}$.

注: 此方法可将二次型$f(x_1,\cdots,x_n)=\sum_{i,j=1}^{n} a_{ij}\,x_i\,x_j$

化为标准型, 利用$Y=Q^TX$.

设$X=(x_1,\cdots,x_n)^T$, 则$f(X)=\sum_{i,j=1}^{n} a_{ij}\,x_i x_j=$

X^TBX. 因为$B=Q\mathrm{diag}\{8,-4,-4,-4\}Q^T$, 所以

$f(X)=X^TQ\mathrm{diag}\{8,-4,-4,-4\}Q^TX$, 令$Y=Q^TX$,

则$f(X)$变成$g(Y)=Y^T\mathrm{diag}\{8,-4,-4,-4\}Y$.

例: 设$A=\begin{pmatrix}0 & i & -1\\ -i & 0 & i\\ -1 & -i & 0\end{pmatrix}$, 验证$A$是正规矩阵, 求酉矩阵$U$, 使$U^HAU$为对角阵.

解: $A^H=\begin{pmatrix}0 & i & -1\\ -i & 0 & i\\ -1 & -i & 0\end{pmatrix}=A$, A是Hermite矩阵, 故正规.

$|\lambda I-A|=\begin{pmatrix}\lambda & -i & 1\\ i & \lambda & -i\\ 1 & i & \lambda\end{pmatrix}=(\lambda+1)^2(\lambda-2)$, 特征值为$\lambda_1=\lambda_2=-1, \lambda_3=2$.

属于特征值-1的特征向量为$a_1=(1,0,1)^T, a_2=(0,1,i)^T$，标准化为$\eta_1=(\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}})^T, \eta_2=(\frac{-i}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{i}{\sqrt{6}})^T$.

属于特征值2的特征向量为$a_3=(-1,i,1)^T$，标准化为$\eta_3=(\frac{-1}{\sqrt{3}},\frac{i}{\sqrt{3}},\frac{1}{\sqrt{3}})^T$.

令$U=(\eta_1,\eta_2,\eta_3)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{6}} & \frac{-1}{\sqrt{3}}\\ 0 & \frac{2}{\sqrt{6}} & \frac{i}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix}$，则$U$是酉矩阵，

且$U^H AU = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

例: 设$f(x) = c_0 + c_1 x + \ldots + c_m x^m$, A是正规矩阵,

则$f(A)$也正规. 特别地$cI \pm A$正规.

例: 设$f(x)=c_0+c_1x+\ldots+c_mx^m$, A是正规矩阵,

则$f(A)$也正规. 特别地$cI\pm A$正规.

证明: 利用公式如果$AB=BA$, 则

$$f(A)g(B)=g(B)f(A).$$

设$f(A)=c_0I+c_1A+\ldots+c_mA^m$

$$\Rightarrow (f(A))^H=\overline{c_0}I+\overline{c_1}A^H+\ldots+\overline{c_n}(A^H)^m.$$

例: 设$f(x)=c_0+c_1x+\ldots+c_mx^m$, A是正规矩阵,

则$f(A)$也正规. 特别地$cI\pm A$正规.

证明: 令$B=A^H$,则$AB=BA$. 设

$$g(B)=\overline{c_0}I+\overline{c_1}B+\ldots+\overline{c_m}B^m=\big(f(A)\big)^H$$

$$\Rightarrow f(A)\big(f(A)\big)^H=\big(f(A)\big)^Hf(A)$$

$\Rightarrow f(A)$正规.