第五章　特征值的估计及对称矩阵的极性

5.1　特征值的估计

5.2　广义特征值问题

5.3　对称矩阵特征值的极性

5.4　矩阵的直积及其应用

5.1 特征值的估计

5.1.1 特征值的界

定理5.1 设$A=(a_{ij})\in \mathbf{R}^{n\times n}$，令

$$M=\max_{1\le r,s\le n}\frac{1}{2}\left|a_{rs}-a_{sr}\right|.$$

若λ表示A 的任一特征值，则λ 的虚部$\mathrm{Im}(\lambda)$满足不等式

$$\left|\mathrm{Im}(\lambda)\right|\le M\sqrt{\frac{n(n-1)}{2}}$$

证 设 $x=(\xi_1,\ldots,\xi_n)^T$为A 的属于特征值λ 的一个特征向量，即 $Ax=\lambda x$，假定 $x^H x=1$, 于是

由 $$x^H A x=\lambda x^H x=\lambda$$

两端取共轭转置，得

$$\bar{\lambda}=(x^H A x)^H=x^H A^H x=x^H A^T x$$

则有

$$\begin{aligned}
2i\,\mathrm{Im}(\lambda)&=\lambda-\bar{\lambda}=x^H(A-A^T)x\\
&=\frac{1}{2}[x^H(A-A^T)x+x^H(A-A^T)x]\\
&=\frac{1}{2}[x^H(A-A^T)x+x^T(A^T-A)\bar{x}]\\
&=\frac{1}{2}\left[\sum_{r,s=1}^{n}(a_{rs}-a_{sr})\bar{\xi}_r\xi_s+\sum_{r,s=1}^{n}(a_{sr}-a_{rs})\xi_r\bar{\xi}_s\right]\\
&=\sum_{r,s=1}^{n}(a_{rs}-a_{sr})\frac{\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s}{2}
\end{aligned}$$

$$\therefore \quad 2|\mathrm{Im}(\lambda)| \le \sum_{r,s=1}^{n}\left(\frac{1}{2}|a_{rs}-a_{sr}|\right)|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s|$$

$$\le M\sum_{r,s=1}^{n}|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s| = M\sum_{\substack{r\neq s\\ r,s=1}}^{n}|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s|$$

由于任意*m*个实数 $\eta_1,\ldots,\eta_m$恒满足

$$m(\eta_1^2+\cdots+\eta_m^2)-(\eta_1+\cdots+\eta_m)^2 = \sum_{1\le i<k\le m}(\eta_i-\eta_k)^2 \ge 0$$

$$\therefore \quad (\eta_1+\cdots+\eta_m)^2 \le m(\eta_1^2+\cdots+\eta_m^2)$$

可得 $4[\mathrm{Im}(\quad)]^2 \quad M^2 \sum_{\substack{r\quad s\\ r,s\quad 1}}^{n} \left|\bar{}_{r}\ {}_{s} \qquad {}_{r}\ \bar{}_{s}\right|$

$$\le n(n-1)M^2\sum_{\substack{r\neq s\\ r,s=1}}^{n}|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s|^2$$

又由 $\left|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s\right|^2=(\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s)\overline{(\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s)}$

$$=2\left|\xi_r\right|^2\left|\xi_s\right|^2-\xi_r^2\bar{\xi}_s^2-\bar{\xi}_r^2\xi_s^2$$

可得 $\displaystyle\sum_{\substack{r\neq s\\ r,s=1}}^{n}\left|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s\right|^2=\sum_{r,s=1}^{n}\left|\bar{\xi}_r\xi_s-\xi_r\bar{\xi}_s\right|^2$

$$=2\sum_{r,s=1}^{n}\left|\xi_r\right|^2\left|\xi_s\right|^2-\sum_{r,s=1}^{n}(\xi_r^2\bar{\xi}_s^2+\bar{\xi}_r^2\xi_s^2)$$

$$=2\sum_{r=1}^{n}\left|\xi_r\right|^2\bullet\sum_{s=1}^{n}\left|\xi_s\right|^2-2\sum_{r=1}^{n}\bar{\xi}_r^2\bullet\sum_{s=1}^{n}\xi_s^2$$

$$=2-2\left|\sum_{r=1}^{n}\xi_r^2\right|^2\leq 2$$

$\therefore\quad 4[\mathrm{Im}(\lambda)]^2\leq 2n(n-1)M^2$ 即 $\left|\mathrm{Im}(\quad)\right|\quad M\sqrt{\dfrac{n(n\quad 1)}{2}}$

证毕

推论 实对称矩阵的特征值都是实数.

引理 1 设 $B\in C^{n\times n}$, 列向量 $y\in C^n$ 满足 $\|y\|_2=1$, 则

$$|y^H By| \le \|B\|_{m_\infty}.$$

证 设 $B=(b_{ij})_{n\times n}$, $y=(\eta_1,\ldots,\eta_n)^T$, 于是有

$$|y^H By| = \left|\sum_{i,j=1}^n b_{ij}\overline{\eta}_i\eta_j\right| \le \max_{i,j}|b_{ij}| \bullet \sum_{i,j=1}^n |\eta_i||\eta_j|$$

$$\le \max_{i,j}|b_{ij}| \bullet \frac{1}{2}\sum_{i,j=1}^n (|\eta_i|^2+|\eta_j|^2)$$

$$= \max_{i,j}|b_{ij}| \bullet \frac{1}{2}(n+n) = \|B\|_{m_\infty}$$

证毕

定理5.2 设 $A\in C^{n\times n}$，则A 的任一特征值λ 满足

$$|\lambda|\le\|A\|_{m_\infty},$$

$$|\mathrm{Re}(\lambda)|\le\frac{1}{2}\|A+A^H\|_{m_\infty},\ |\mathrm{Im}(\lambda)|\le\frac{1}{2}\|A-A^H\|_{m_\infty}$$

证 $x=(\xi_1,\xi_2,\ldots,\xi_n)^T$是$A$ 的属于特征值λ的单位特征向量，即 $Ax=\lambda x$. 两端左乘x^H，可得 $\lambda=x^HAx$，则有 $\bar{\lambda}=x^HA^Hx$. 根据引理1，有

$$|\lambda|=|x^HAx|\le\|A\|_{m_\infty}$$

$$|\mathrm{Re}(\lambda)|=\frac{1}{2}|\lambda+\bar{\lambda}|=\frac{1}{2}|x^H(A+A^H)x|\le\frac{1}{2}\|A+A^H\|_{m_\infty}$$

$$|\mathrm{Im}(\lambda)|=\frac{1}{2}|\lambda-\bar{\lambda}|=\frac{1}{2}|x^H(A-A^H)x|\le\frac{1}{2}\|A-A^H\|_{m_\infty}$$

证毕

推论 Hermite矩阵的特征值都是实数，反Hermite矩阵的特征值为零或纯虚数.

例5.1 估计矩阵

$$A=\begin{pmatrix} 1 & -0.8 \\ 0.5 & 0 \end{pmatrix}$$

的特征值的上界.

解 应用**定理5.2** $|\lambda| \le 2, |\mathrm{Re}(\lambda)| \le 2, |\mathrm{Im}(\lambda)| \le 1.3$

应用**定理5.1**

$$|\mathrm{Im}(\lambda)| \le M\sqrt{\frac{2(2-1)}{2}} = 0.65$$

定义5.1 设 $A=(a_{ij})\in C^{n\times n}, R_r(A)=\sum_{\substack{s=1\\s\neq r}}^{n}|a_{rs}|$

（简写为 R_r），$r=1,2,\cdots,n$. 如果 $|a_{rr}|>R_r$ $(r=1,2,\cdots,n)$，则称矩阵A按行**严格对角占优.**

如果 $|a_{rr}|\geq R_r(r=1,2,\cdots,n)$,且有 $1\leq r_0\leq n$

使得 $|a_{r_0r_0}|>R_{r_0}$ 成立，则称矩阵A 按行**(弱)对角占优.**

定义 5.2 设 $A\in C^{n\times n}$，如果 A^T 按行严格对角占优，则称A按列严格对角占优；如果A^T按行（弱）对角占优，则称矩阵A 按列（弱）对角占优.

定理5.3 设 $A=(a_{ij})\in C^{n\times n}, M_r=|a_{rr}|+\sum_{s=r+1}^{n}|a_{rs}|$

$m_r=|a_{rr}|-\sum_{s=r+1}^{n}|a_{rs}|$. 如果$A$按行严格对角占优，

则
$$0<\prod_{r=1}^{n}m_r\le|\det A|=\prod_{r=1}^{n}\lambda_r(A)\le\prod_{r=1}^{n}M_r$$

且当$a_{rs}=0(s>r)$时，等号成立.

证 由于A按行严格对角占优，所以 $\det A\neq 0$,

考虑方程组

$$\begin{pmatrix}a_{21}\\ \vdots\\ a_{n1}\end{pmatrix}+A_1\begin{pmatrix}\xi_2\\ \vdots\\ \xi_n\end{pmatrix}=O,\quad A_1=\begin{pmatrix}a_{22} & \cdots & a_{2n}\\ \vdots & & \vdots\\ a_{2n} & \cdots & a_{nn}\end{pmatrix}$$

有唯一解，设解为 $x^{(1)}=\left(\xi_2^{(1)},\xi_3^{(1)},\cdots,\xi_n^{(1)}\right)$

则必有 $\left|\xi_k^{(1)}\right|=\max\left\{\left|\xi_2^{(1)}\right|,\left|\xi_3^{(1)}\right|,\cdots,\left|\xi_n^{(1)}\right|\right\}<1$

由于 $a_{k1}+\sum_{s=2}^{n} a_{ks}\xi_s^{(1)}=0$，则有 $-a_{kk}=\dfrac{1}{\xi_k^{(1)}}a_{k1}+\sum_{\substack{s=2\\ s\neq k}}^{n} a_{ks}\dfrac{\xi_s^{(1)}}{\xi_k^{(1)}}$

因为 $$\det A=\det\left[A\begin{pmatrix}1 & O^T\\ x^{(1)} & I\end{pmatrix}\right]$$

$$=\det\begin{pmatrix}b_{11} & a_{12}\cdots a_{1n}\\ O & A_1\end{pmatrix}=b_{11}\det A_1$$

其中 $$b_{11}=a_{11}+\sum_{s=2}^{n} a_{1s}\xi_s^{(1)},\quad \left|\xi_s^{(1)}\right|<1(s=2,\cdots,n)$$

类似地， $\det \boldsymbol{A}_1 \quad b_{22} \det \boldsymbol{A}_2$,

其中 $b_{22} \quad a_{22} \quad \overset{n}{\underset{s\ \ 3}{\quad}} a_{2s} \ {}_{s}^{(2)}, \ \left| \ {}_{s}^{(2)} \right| \quad 1(s \quad 3, \quad , n)$

$$A_2 \quad \begin{matrix} a_{33} & & a_{3n} \\ & & \\ a_{n3} & & a_{nn} \end{matrix}$$

于是 $\det A = b_{11} b_{22} \det A_2$ ，重复上述过程，最后可得

$$\det A = \prod_{r=1}^{n} \left(a_{rr} + \sum_{s=r+1}^{n} a_{rs} \xi_s^{(r)} \right)$$

$$\prod_{r=1}^{n} \left| a_{rr} + \sum_{s=r+1}^{n} a_{rs} \xi_s^{(r)} \right| \le \prod_{r=1}^{n} \left(|a_{rr}| + \sum_{s=r+1}^{n} |a_{rs}| \right) = \prod_{r=1}^{n} M_r$$

$$\prod_{r=1}^{n}\left|a_{rr}+\sum_{s=r+1}^{n}a_{rs}\xi_s^{(r)}\right| \ge \prod_{r=1}^{n}\left(|a_{rr}|-\sum_{s=r+1}^{n}|a_{rs}|\right)=\prod_{r=1}^{n}m_r$$

故 $$0<\prod_{r=1}^{n}m_r \le |\det A| = \prod_{r=1}^{n}\lambda_r(A) \le \prod_{r=1}^{n}M_r$$

特别地，当 $a_{rs}=0(s>r)$ 时，恒有 $m_r = M_r$

故得 $$\overset{n}{\underset{r\quad 1}{}} m_r \qquad |\det \boldsymbol{A}| \qquad \overset{n}{\underset{r\quad 1}{}} M_r$$

证毕

例5.2 估计矩阵 A $\begin{matrix} 1 & 0.8 \\ 0.5 & 1 \end{matrix}$ 按模最小特征值的上界.

解 因为 $M_1 \quad 1.8, M_2 \quad 1,$ 所以有

$$|\lambda(A)|_{\min} \le \left[\prod_{r=1}^{2} |\lambda_r(A)|\right]^{\frac{1}{2}} \le \left[\prod_{r=1}^{2} M_r\right]^{\frac{1}{2}} = \sqrt{1.8}$$

实际上，A 的两个特征值是 $\lambda_{1,2} = 1 \pm i\sqrt{0.4},$

从而 $|\lambda_{1,2}| = \sqrt{1.4}.$

定理5.4 设 $A \quad (a_{ij}) \quad C^{n \quad n}$，则有

$$\prod_{r=1}^{n} \left| \quad _{r} \quad \boldsymbol{A} \right| \quad |\det \boldsymbol{A}| \quad \prod_{s=1}^{n} \quad \sum_{r=1}^{n} |a_{rs}|^{2} \quad ^{\frac{1}{2}}$$

且式中等号成立 $\Leftrightarrow$ 某 $a_{s_0} \quad 0$ 或者 $(a_r, a_s) \quad 0 \ (r \neq s)$

其中 $a_1, \cdots, a_n$ 表示A 的n 个列向量.

证 (1)如果向量组$a_1, \cdots, a_n$线性相关，则 $\det A \quad 0$

(2)如果向量组 $a_1, \cdots, a_n$线性无关，构造向量组 $b_1, \cdots, b_n$两两正交，且满足

$$\begin{cases} a_1 = b_1 \\ a_2 = b_2 + \lambda_{21} b_1 \\ \vdots \\ a_n = b_n + \lambda_{n1} b_1 + \cdots + \lambda_{n,n-1} b_{n-1} \end{cases}$$

其中 $\lambda_{sr} = (a_s, b_r) / \|b_r\|^2 \quad (s > r)$

令 $B \quad (b_1, \quad , b_n)$，则 $$A \quad B \begin{matrix} 1 & {}_{21} & & {}_{n1} \\ 0 & 1 & & {}_{n2} \\ & & & \\ 0 & 0 & & 1 \end{matrix}$$

所以 $\det A \quad \det B$

又 $$\|a_s\|^2 = \|b_s + \lambda_{s1} b_1 + \cdots + \lambda_{s,s-1} b_{s-1}\|^2$$
$$= \|b_s\|^2 + |\lambda_1|^2 \|b_1\|^2 + \cdots + |\lambda_{s,s-1}|^2 \|b_{s-1}\|^2 \geq \|b_s\|^2$$

$$|\det B|^2 = \det B^H \det B = \prod_{s=1}^{n} \|b_s\|^2 = \left(\prod_{s=1}^{n} \|b_s\| \right)^2 .$$

因此 $$|\det A| = |\det B| = \prod_{s=1}^{n} \|b_s\| \leq \prod_{s=1}^{n} \|a_s\| = \left[\prod_{s=1}^{n} \left(\sum_{r=1}^{n} |a_{rs}|^2 \right) \right]^{\frac{1}{2}} .$$

充分性. 当 $a_{s_0} \quad 0$ 时，两端都为零，所以等号成立. 当 $(a_r, a_s) = 0(r \neq s)$ 时，有

$$|\det A|^2 = \det A^H \det A = \prod_{s=1}^{n} \|a_s\|^2 = \prod_{s=1}^{n}\left(\sum_{r=1}^{n} |a_{rs}|^2\right)$$

所以等号成立.

必要性.（反证法）若 $a_s \quad 0(s \quad 1, \quad , n)$且存在最小指标 s_0满足 $(a_{s_0}, a_{r_0}) \neq 0(r_0 < s_0)$, 则有

$$\begin{cases} a_1 = b_1 \\ \vdots \\ a_{s_0-1} = b_{s_0-1} \\ a_{s_0} = b_{s_0} + \cdots + \lambda_{s_0,r_0} b_{r_0} + \cdots \\ \vdots \end{cases}$$

其中 $\lambda_{s_0,r_0} = (a_{s_0}, b_{r_0}) / \|b_{r_0}\|^2 = (a_{s_0}, a_{r_0}) / \|a_{r_0}\|^2 \neq 0$

所以 $\|a_{s_0}\|^2 = \|b_{s_0} + \cdots + \lambda_{s_0,r_0} b_{r_0} + \cdots\|^2 > \|b_{s_0}\|^2$

因而

$$|\det \boldsymbol{A}| \quad |\det \boldsymbol{B}| \quad \mathop{}_{s\ \ 1}^{n} \|b_s\| \quad \mathop{}_{s\ \ 1}^{n} \|a_s\| \quad \mathop{}_{s\ \ 1}^{n} \mathop{}_{r\ \ 1}^{n} |a_{rs}|^2 \quad ^{\frac{1}{2}}.$$

与条件矛盾. 所以必有某 $a_{s_0} \quad 0$

或者 $(a_r, a_s) = 0 \ (r \neq s)$ 证毕

定理5.5 设 $A \quad (a_{ij}) \quad C^{n \quad n}$的特征值为 $\lambda_1, \cdots, \lambda_n$,
则有

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

证 因为存在酉矩阵 U，使得 $A = UTU^H$, 其中 T 是上三角矩阵. 而

$$\sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i=1}^{n} |t_{ii}|^2 \le \sum_{i,j=1}^{n} |t_{ij}|^2 = \|T\|_F^2$$

又 $\|A\|_F^2 = \|T\|_F^2$ 所以

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

证毕

例 5.3 在**定理5.5**中等号成立 $\Leftrightarrow \boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^H.$

例5.4 已知矩阵

$$\boldsymbol{A} \quad \begin{matrix} 3 \quad i & 2 \quad 3i & 2i \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{matrix}$$

的一个特征值是2，估计另外两个特征值的上界.

解 因为 $\|A\|_F^2 - 4 = 25,$ 所以 $|\lambda_{2,3}| \le 5.$

5.1.2 特征值的包含区域

定义5.3 设 $\boldsymbol{A} = (a_{ij})\ ?\ C^{n' n}$, 称由不等式

$$|z \quad a_{ii}| \quad R_i$$

在复平面上确定的区域为矩阵***A*** 的第***i***个**Gerschgorin**圆（盖尔圆），并用记号 G_i 来表示. 其中

$$R_i \quad R_i(\boldsymbol{A}) \quad \overset{n}{\underset{\substack{j\ 1\\ j\ i}}{}} |a_{ij}|$$ 称为盖尔圆 G_i 的半径$(i = 1, \cdots, n)$.

定理5.6（Gerschgorin theorem) 矩阵 $\boldsymbol{A} = (a_{ij}) \in C^{n\times n}$ 的一切特征值都在它的***n***个盖尔圆的并集之内.

证 设 λ 是A 的任一特征值，$x=(\xi_1,\cdots\xi_n)$为属于 λ 的特征向量. 令 $\left|\xi_{i_0}\right|=\max\limits_{i}\left|\xi_i\right|$，则 $\xi_{i_0}\neq 0.$

由于$Ax=\lambda x$，所以 $\sum\limits_{j=1}^{n} a_{i_0 j}\xi_j=\lambda\xi_{i_0}$

即 $\quad a_{i_0 j}\xi_j \quad \lambda \quad a_{i_0 i_0} \quad \xi_{i_0}$，从而可得

$j \quad i_0$

$$\left|\lambda-a_{i_0 i_0}\right|=\left|\sum_{j\neq i_0} a_{i_0 j}\frac{\xi_j}{\xi_{i_0}}\right|\leq\sum_{j\neq i_0}\left|a_{i_0 j}\right|\left|\frac{\xi_j}{\xi_{i_0}}\right|\leq R_{i_0},$$

所以 $\lambda\in G_{i_0}.$

证毕

例5.5 估计矩阵

$$A \quad \begin{matrix} 1 & 0.1 & 0.2 & 0.3 \\ 0.5 & 3 & 0.1 & 0.2 \\ 1 & 0.3 & 1 & 0.5 \\ 0.2 & 0.3 & 0.1 & 4 \end{matrix}$$

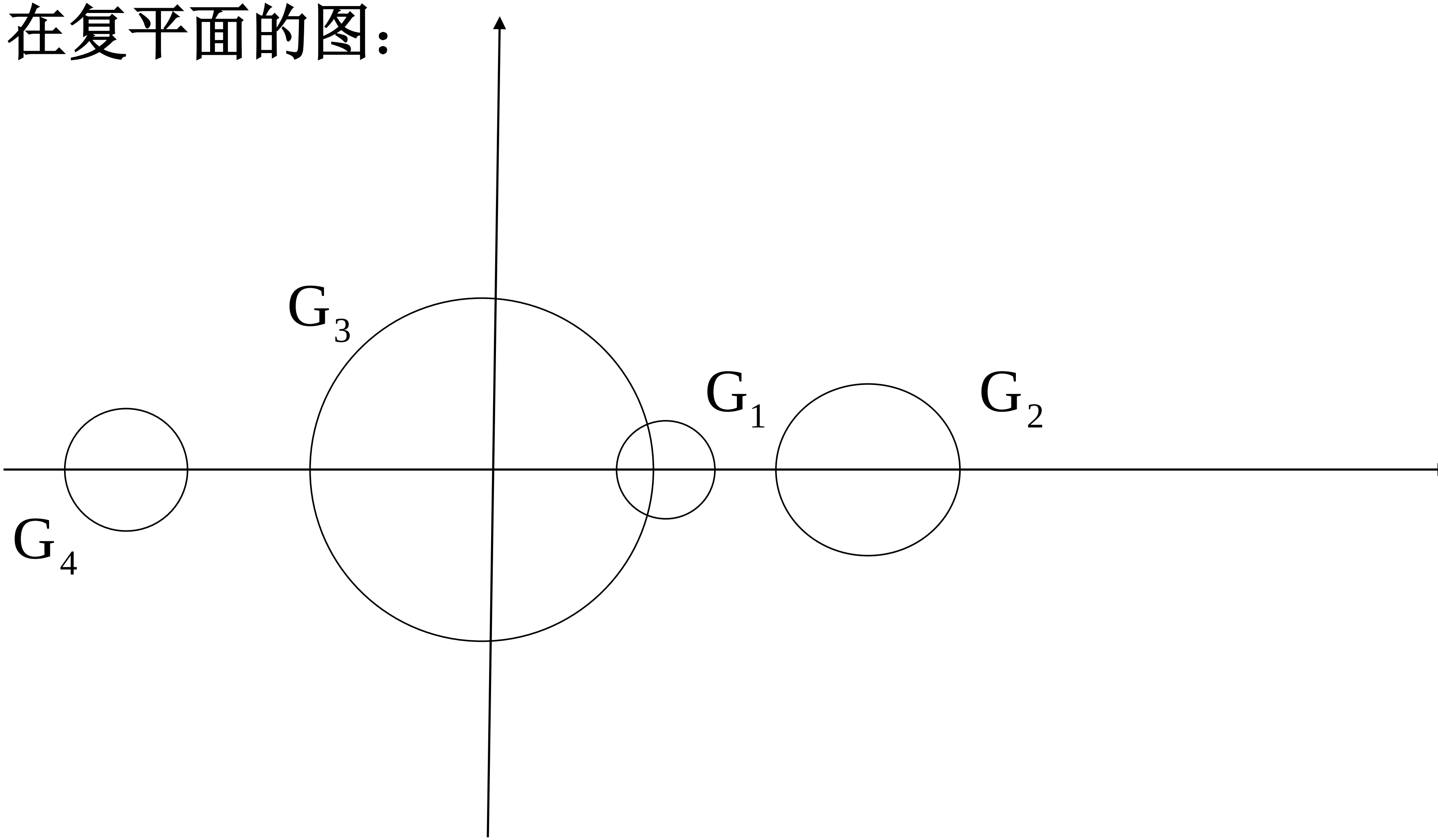
的特征值的范围.

解 A的4 个盖尔圆为

$$|z-1| \le 0.6, \quad |z-3| \le 0.8$$

$$|z+1| \le 1.8, \quad |z+4| \le 0.6$$

在复平面的图：



那么，$\boldsymbol{A}$的全部特征值就在这四个盖尔圆并起来的区域之中.

连通区域：区域中的任意两点都可以用位于该区域内的一条折线连接起来的区域.

连通部分：交结为一起的盖尔圆所构成的最大连通区域.

定理 5.7（**Gerschgorin theorem**）在矩阵*A*所有盖尔圆组成的任一连通部分中，含有*A*的特征值的个数等于该连通部分的盖尔圆的个数.

证 考虑矩阵

$$A(u) \quad \begin{matrix} a_{11} & ua_{12} & ua_{1n} \\ ua_{21} & a_{22} & ua_{2n} \\ & & \\ ua_{n1} & ua_{n2} & a_{nn} \end{matrix}$$

因为 $A(u)$ 的特征值 $\lambda_i(u)(i=1,\cdots,n)$ 连续依赖于 $\boldsymbol{u}$，考虑 $u\in[0,1]$, $\lambda_i(0)=a_{ii}$ 是 $A(0)$ 的特征值，$\lambda_i(1)$ 是 $\boldsymbol{A}$ 的特征值. 因此 $\lambda_i(u)$ 在复平面上的图形是以 $\lambda_i(0)$ 为起点，以 $\lambda_i(1)$ 为终点的连续曲线.

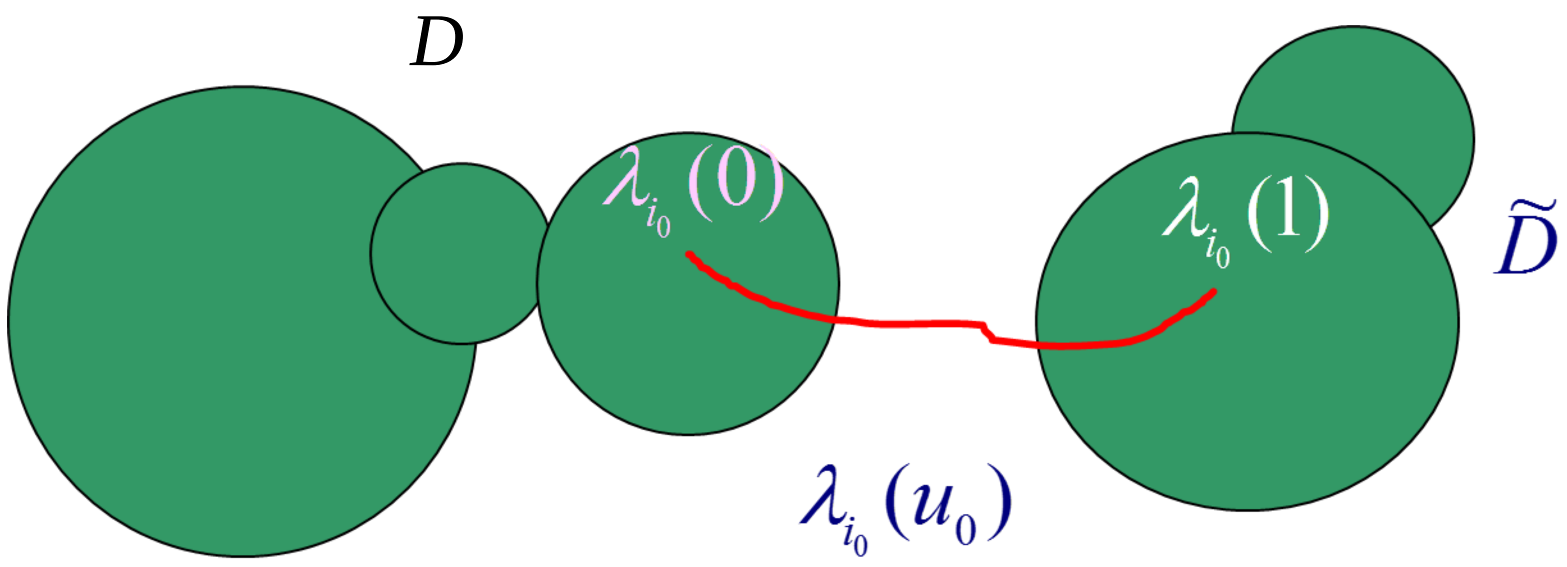
设 **A(1)=A** 的一个连通部分由 ***k*** 个盖尔圆构成，记作D. 因此，**A(0)** 的 ***k*** 个特征值必在其中，如果 D 中没有 ***A*** 的 ***k*** 个特征值，则至少有一个 i_0 ,使得点 $\lambda_{i_0}(0)$ 连续地变动到 $\lambda_{i_0}(1)$ 且 $\lambda_{i_0}(1)$ 在 ***D*** 之外.而必在 ***A*** 的另一连通区域 D 中.因此这条曲线必定有一部分既不在 ***D*** 中,又不在 $\tilde{D}$ 中，也不在 ***A*** 的其它连通部分之中. 即存在 $u_0\in(0,1)$，使得 $\lambda_{i_0}(u_0)$ 不在 ***A*** 的所有盖尔

圆的并集中. 但因$\lambda_{i_0}(u_0)$是 $A(u_0)$ 的特征值， 则它必在盖尔圆

$$|z-a_{ii}| \le \sum_{j\ne i} |u_0 a_{ij}| = u_0 R_i \quad (i=1,\cdots,n)$$

的并集中.与 $\lambda_{i_0}(u_0)$不在A 的所有盖尔圆矛盾.

所以，A 在D 中的特征值的个数等于构成D的盖尔圆个数. 证毕



例5.6 讨论矩阵 $A = \begin{pmatrix} 1 & 0.8 \\ 0.5 & 0 \end{pmatrix}$ 的特征值分布状况.

盖尔圆定理的**应用**：矩阵特征值的隔离.

设 $A = (a_{ij})_{n \times n}$，构造对角矩阵

$$D = diag(\alpha_1, \alpha_2, \cdots, \alpha_n)$$

其中 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 都是正数. 由于

$$B = DAD^{-1}$$

所以 B 与 A 的特征值相同.

推论 若将 A 的 R_i 改作 $r_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \dfrac{\alpha_i}{\alpha_j}$，则**定理5.6**与**5.7**的结论成立.

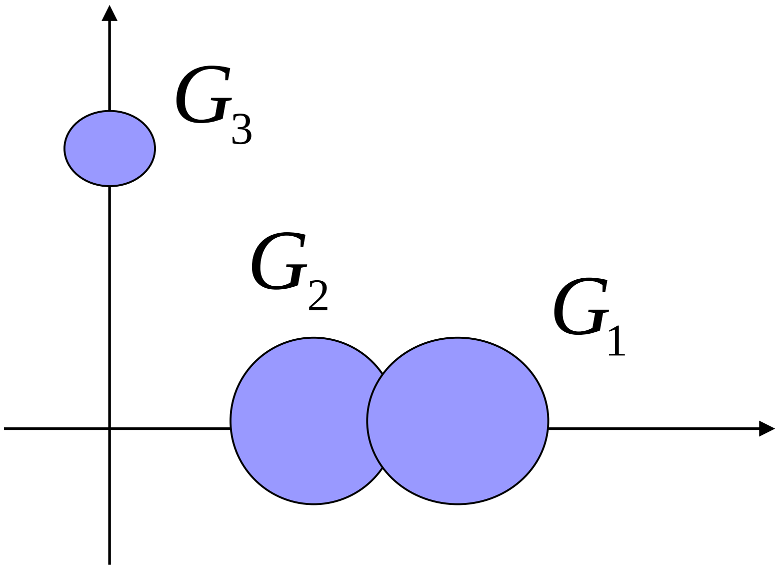
例5.7 隔离矩阵 $A=\begin{bmatrix}20 & 5 & 0.8\\ 4 & 10 & 1\\ 1 & 2 & 10i\end{bmatrix}$ 的特征值.

解 A 的3个盖尔圆为

$$G_1: |z-20|\le 5.8,$$

$$G_2: |z-10|\le 5;$$

$$G_3: |z-10i|\le 3$$



选取 $D=diag$（1，1，2）

则

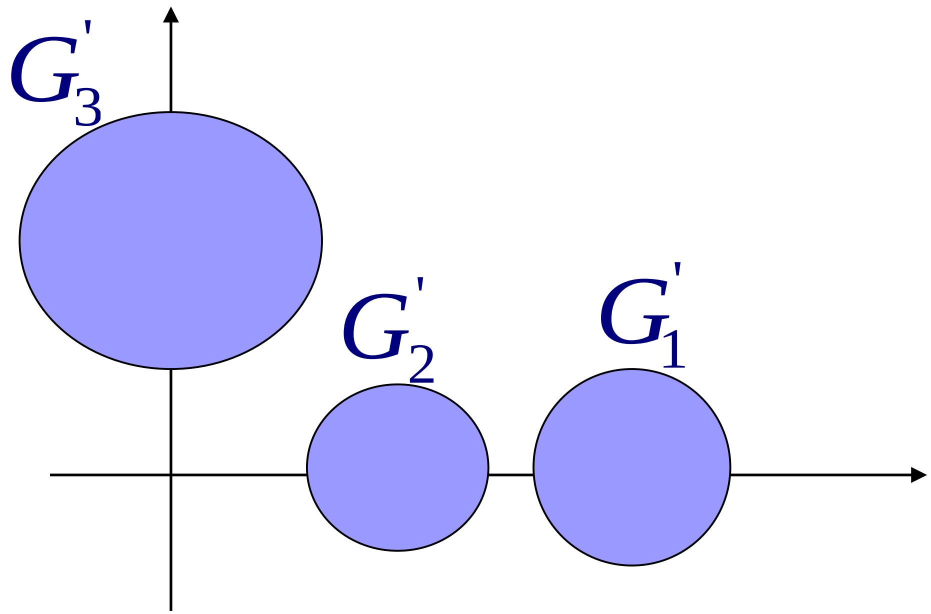
$$B = DAD^{-1} = \begin{pmatrix} 20 & 5 & 0.4 \\ 4 & 10 & 0.5 \\ 2 & 4 & 10i \end{pmatrix}$$

的3个盖尔圆为

G_1': $|z-20| \le 5.4,$

G_2': $|z-10| \le 4.5;$

G_3': $|z-10i| \le 6$



因为 5.4+4.5=9.9< 10， 4.5+6=10.5 $< 10\sqrt{2}$

5.4+6=11.4 $< 10\sqrt{5}$ ， 所以 G_1', G_2'与G_3' 互不相交，因而3个特征值在分别在该3个盖尔圆中.

例5.8 如果矩阵 $A=(a_{ij})_{n\times n}$，按行（列）严格对角占优，则 $\det A \neq 0.$

证 设 λ 是A的任一特征值，则存在 i使 $\lambda \in G_i$，

于是可得 $|\lambda - a_{ii}| \le R_i = \sum\limits_{j\neq i} |a_{ij}|$

如果 $\lambda = 0$，则有 $|a_{ii}| \le \sum\limits_{j\neq i} |a_{ij}|$

这与A按行严格对角占优矛盾，故应有 $\lambda \neq 0$,

所以 $\det A \neq 0.$

定理5.8 设不可约矩阵 $\boldsymbol{A}=(a_{ij})_{n' \ n}$,的一个特征值在其$\boldsymbol{n}$个盖尔圆 $|\lambda \quad a_{ii}| \quad R_i(i \quad 1, \quad ,n)$ 并集的边界上，则$\boldsymbol{n}$个圆周 $|\lambda \quad a_{ii}| \quad R_i(i \quad 1, \quad ,n)$ 都过点 λ.

定理5.9 如果 $A=(a_{ij})_{n\times n}$,不可约，且存在 i_0 使得

$$\sum_{j=1}^{n}|a_{i_0 j}|<\|A\|_\infty，\text{则有}\rho(A)<\|A\|_\infty.$$

证 对 $\forall z\in\bigcup_{i=1}^{n}G_i, \exists i_1$ 使 $|z-a_{i_1 i_1}|\le R_{i_1}$,

所以有 $|z|\le|a_{i_1 i_1}|+R_{i_1}=\sum_{j=1}^{n}|a_{i_1 j}|\le\|A\|_\infty$

设 $\bar{S} \quad z \,|\, |z| \quad \|\boldsymbol{A}\|$ ，则有 $\overset{n}{\underset{i \quad 1}{}} G_i \quad \bar{S}$

假设 $\rho(\boldsymbol{A}) \quad \|\boldsymbol{A}\|$,则$\boldsymbol{A}$至少有一个特征值 λ_0

满足 $$|\lambda_0| = \rho(A) = \|A\|_\infty .$$

所以 λ_0 在 $\overline{S}$ 的边界上，从而 λ_0 在**A** 的**n**个盖尔圆并集的边界上. 由定理5.8可得

$$|\lambda_0 - a_{ii}| = R_i \quad (i = 1, \cdots, n)$$

特别地， $$|\lambda_0 - a_{i_0 i_0}| = R_{i_0}$$

于是 $$|\lambda_0| \le R_{i_0} + |a_{i_0 i_0}| = \sum_{j=1}^{n} |a_{i_0 j}| < \|A\|_\infty .$$

这与假设矛盾，故应有

$\rho(A) \neq \|A\|_\infty$ 所以 $\rho(A) < \|A\|_\infty .$ 证毕

定理5.10 设 $\boldsymbol{A} \quad (a_{ij}) \quad C^{n \quad n}, \boldsymbol{B} \quad (b_{ij}) \quad R^{n \quad n}$,

如果 $b_{ij} \ge |a_{ij}| \ (i, j = 1, \cdots, n)$, 则对$\boldsymbol{A}$的任一特征值 λ，

必有 $\boldsymbol{i}$，使 $|\lambda \quad a_{ii}| \quad \rho(\boldsymbol{B}) \quad b_{ii}$.

证 由 $b_{ij} \ge |a_{ij}| \ (i, j = 1, \cdots, n)$, 有 $b_{ij} \ge 0$

（1）如果 $b_{ij} > 0$ 则 $\exists x = (\xi_1, \cdots, \xi_n)^T$ 满足

$\xi_i > 0$ 且使 $\boldsymbol{Bx} = \rho(\boldsymbol{B})\boldsymbol{x}$

令 $\boldsymbol{D} \quad diag(\xi_1, \quad , \xi_n), \boldsymbol{C} \quad \boldsymbol{D}^{\ 1}\boldsymbol{AD} \quad (c_{ij})$

对$\boldsymbol{C}$的任一特征值 λ 必有$\boldsymbol{i}$使

$$|\lambda - c_{ii}| \le \sum_{j \ne i} |c_{ij}|$$

所以 $\left|\lambda_0 \quad a_{ii}\right| \qquad \underset{j \quad i}{} \left|\xi_i{}^{1} a_{ij}\xi_j\right| \quad \xi_i{}^{1} \underset{j \quad i}{} b_{ij}\xi_j$

$$\xi_i{}^{1} \overset{n}{\underset{j \quad 1}{}} b_{ij}\xi_j \quad b_{ii}\xi_i \qquad \xi_i{}^{1} \ \rho(\boldsymbol{B})\xi_i \quad b_{ii}\xi_i$$

$$\rho(\boldsymbol{B}) \quad b_{ii}$$

（2）如果 $b_{ij} \geq 0$, 令

$$B_k = (b_{ij}^{(k)})_{n\times n}, \quad b_{ij}^{(k)} = b_{ij} + \frac{1}{k}$$

则 $b_{ij}^{(k)} > 0$ 且有 $b_{ij}^{(k)} > \left|a_{ij}\right|$

所以对$\boldsymbol{A}$ 的任一特征值λ，必有 i_k 使

$$\left|\lambda - a_{i_k i_k}\right| \leq \rho(B_k) - b_{i_k i_k}^{(k)}$$

由于 $i_k \in \{1,\cdots,n\}$，所以无穷序列 $\{i_k\}$ 中必有一个无穷子列 $\{i_{k_m}\}$ 满足 $i_{k_m} = i \in \{1,\cdots,n\}$ $(m = 1,2,3,\cdots)$

于是
$$|\lambda - a_{ii}| = |\lambda - a_{i_{k_m} i_{k_m}}| \le \rho(B_{k_m}) - b_{i_{k_m} i_{k_m}}^{(k_m)}$$
$$= \rho(B_{k_m}) - b_{ii}^{(k_m)}$$

因为 $\lim\limits_{k} B_k \quad B$,所以 $\lim\limits_{k_\mathrm{m}} B_{k_m} \quad B$,从而
$$\lim_{k_m \to \infty} b_{ii}^{(k_m)} = b_{ii}, \quad \lim_{k_m \to \infty} \rho(B_{k_m}) = \rho(B)$$

因此
$$|\lambda_0 - a_{ii}| \le \rho(B) - b_{ii}$$

证毕

例5.9 估计矩阵 $A \quad \begin{matrix} 1 & 0.8 \\ 0.5 & 0 \end{matrix}$ 的特征值范围

解 取 $B \quad \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$,则$b_{ij} \quad |a_{ij}|$.因为$\rho(\boldsymbol{B}) \quad 2$,

所以A的特征值λ至少满足下面二不等式之一,

$$|\lambda - 1| \le 2 - 1 = 1, \quad |\lambda| \le 2 - 1 = 1$$

引理2 设 σ, τ 是两个非负实数, $0 \le \alpha \le 1$ ，则有

$$\tau^{\alpha} \sigma^{1-\alpha} \le \alpha \tau + (1 - \alpha) \sigma$$

定理5.11（Ostrowski theorm）设

$A \quad (a_{ij}) \quad C^{n \quad n}, 0 \quad \alpha \quad 1$, λ是A的任一特征值.则存

在i，使得 $\left|\lambda - a_{ii}\right| \le \left[R_i(A)\right]^{\alpha}\left[R_i(A^T)\right]^{1-\alpha}$.

证 当 $\alpha \quad 0$或1时，显然成立.

当 $0 < \alpha < 1$时，设 $x = (\xi_1, \cdots, \xi_n)^T$ 是A的属于

特征值 λ 的特征向量，即有 $Ax = \lambda x$.于是

$$\left|\lambda - a_{ii}\right|\left|\xi_i\right| \le \left|\sum_{j \ne i} a_{ij}\xi_j\right| \le \sum_{j \ne i}\left|a_{ij}\right|\left|\xi_j\right| \quad (i = 1, \cdots, n)$$

用反证法. 假设不成立，则对于 $0 < \alpha < 1$

有 $\left|\lambda - a_{ii}\right| > \left[R_i(A)\right]^{\alpha}\left[R_i(A^T)\right]^{1-\alpha}, (i = 1, \cdots, n).$

且$R_i(A)$ 0. 由于x 0,所以 i_0使得ξ_{i_0} 0.则有

$$\left[R_{i_0}(A)\right]^{\alpha}\left[R_{i_0}(A^T)\right]^{1-\alpha}\left|\xi_{i_0}\right| < \left|\lambda - a_{i_0 i_0}\right|\left|\xi_{i_0}\right|$$

$$\le \sum_{j\ne i_0}\left|a_{i_0 j}\right|\left|\xi_j\right| = \sum_{j\ne i_0}\left|a_{i_0 j}\right|^{\alpha}\left|a_{i_0 j}\right|^{1-\alpha}\left|\xi_j\right|$$

$$\le \left(\sum_{j\ne i_0}\left|a_{i_0 j}\right|\right)^{\alpha}\left(\sum_{j\ne i_0}\left|a_{i_0 j}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}\right)^{1-\alpha}$$

$$= \left[R_{i_0}(A)\right]^{\alpha}\left(\sum_{j\ne i_0}\left|a_{i_0 j}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}\right)^{1-\alpha}$$

所以 $$\left[R_{i_0}(A^T)\right]^{1-\alpha}\left|\xi_{i_0}\right| < \left(\sum_{j\ne i_0}\left|a_{i_0 j}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}\right)^{1-\alpha}$$

即 $$\left[R_{i_0}(A^T)\right]\left|\xi_{i_0}\right|^{\frac{1}{1-\alpha}} < \sum_{j\neq i_0}\left|a_{i_0 j}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}$$

因为当 $\xi_i = 0$ 时，恒有

$$\left[R_i(A^T)\right]\left|\xi_i\right|^{\frac{1}{1-\alpha}} \leq \sum_{j\neq i}\left|a_{ij}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}$$

所以 $$\sum_{i=1}^{n}\left[R_i(A^T)\right]\left|\xi_i\right|^{\frac{1}{1-\alpha}} < \sum_{i=1}^{n}\sum_{j\neq i}\left|a_{ij}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}$$

又

$$\sum_{i=1}^{n}\left[R_i(A^T)\right]\left|\xi_i\right|^{\frac{1}{1-\alpha}} = \sum_{i=1}^{n}\sum_{j\neq i}\left|a_{ji}\right|\left|\xi_i\right|^{\frac{1}{1-\alpha}} = \sum_{j=1}^{n}\sum_{i\neq j}\left|a_{ij}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n}\left|a_{ij}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}} - \sum_{j=1}^{n}\left|a_{jj}\right|\left|\xi_j\right|^{\frac{1}{1-\alpha}}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}||\xi_j|^{\frac{1}{1-\alpha}}-\sum_{i=1}^{n}|a_{ii}||\xi_i|^{\frac{1}{1-\alpha}}=\sum_{i=1}^{n}\sum_{j\neq i}|a_{ij}||\xi_j|^{\frac{1}{1-\alpha}}$$

与上式矛盾，故假设错误，所以不等式对于$0<\alpha<1$亦成立. 证毕

推论1 在**定理5.11** 的条件下，存在 $\boldsymbol{i}$ 使得

$$|\lambda \quad a_{ii}| \quad \alpha R_i(\boldsymbol{A}) \quad (1 \quad \alpha)R_i(\boldsymbol{A}^T)$$

推论2 如果$\boldsymbol{A}$ 奇异，取 $0 \quad \alpha \quad 1$，则存在$\boldsymbol{i}$使得

$$(1)\ |a_{ii}| \quad R_i(\boldsymbol{A})^{\alpha}\ R_i(\boldsymbol{A}^T)^{1 \quad \alpha};$$

$$(2)\ |a_{ii}| \quad \alpha R_i(\boldsymbol{A}) \quad (1 \quad \alpha)R_i(\boldsymbol{A}^T).$$

推论3 对于 $0 \quad \alpha \quad 1$, 恒成立

$$(1)\ \rho(A) \quad \max_i \ |a_{ii}| \quad R_i(A)^{\alpha} \ R_i(A^T)^{1 \quad \alpha} \ ;$$

$$(2)\ \rho(A) \quad \max_i \ |a_{ii}| \quad \alpha R_i(A) \quad (1 \quad \alpha) R_i(A^T) \ .$$

推论4 设 $\rho_i(A) \quad {}_{j\ 1}^{\ n} |a_{ij}|$, 取 $0 \le \alpha \le 1$, 则有

$$\rho(A) \quad \max_i \ \rho_i(A)^{\alpha} \ \rho_i(A^T)^{1 \quad \alpha}$$

推论5 （Farnbll A B）

$$\rho(A) £ \max_i \text{轻臌}(A)\rho_i(A^T)^{\frac{1}{2}}$$

推论6 （BrauerA）

$$\rho(A) £ \min\left\{\max_i \rho_i(A), \max_i \rho_i(A^T)\right\}$$

推论7 （**Parker W V**）

$$\rho(A) ? \frac{1}{2}\max_i\left\{\rho_i(A) \quad \rho_i(A^T)\right\}$$

推论8（**Browue E T**）

$$\rho(A) ? \frac{1}{2}\max_i \rho_i(A) \quad \max_i \rho_i(A^T)$$

定理5.12（**Ostrowski theorem2**）设

$A=\left(a_{ij}\right)\in C^{n\times n}(n\geq 2)$，则对于$A$的任一特征值$\lambda$，

存在 **i,j (i≠j)** ，使 λ 属于

$$\Omega_{ij}(A)=\left\{z \,\middle|\, z\in C, |z-a_{ii}|\,|z-a_{jj}|\leq R_i(A)R_j(A)\right\}$$

证 设 $x=(\xi_1,\cdots,\xi_n)^T$ 是A的属于特征值λ的特征向量，即有 $Ax=\lambda x$. 选择 $r,t(r\neq t)$ 使满足

$$|\xi_r|\geq|\xi_t|\geq|\xi_j|(j\neq r)$$

如果 $\xi_t=0$,则$\xi_j=0(j\neq r)$,此时$\xi_r\neq 0$.于是有

$$\lambda\xi_r=\sum_{j=1}^{n}a_{rj}\xi_j=a_{rr}\xi_r$$

即$\lambda=a_{rr}$. 所以 $|\lambda-a_{rr}||\lambda-a_{tt}|=0\leq R_r(A)R_t(A)$

如果 $\xi_t\neq 0$,则有 $(\lambda-a_{ii})\xi_i=\sum_{j\neq i}a_{ij}\xi_j\ \ (i=1,\cdots,n)$

可得

$$|\lambda - a_{rr}||\xi_r| \le \sum_{j\neq r}|a_{rj}||\xi_j| \le |\xi_t| R_r(A)$$

$$|\lambda - a_{tt}||\xi_t| \le \sum_{j\neq t}|a_{tj}||\xi_j| \le |\xi_r| R_t(A)$$

于是 $$|\lambda \quad a_{rr}||\lambda \quad a_{tt}| \quad R_r(A)R_t(A)$$

总之，无论 ξ_t 是否为零，必存在**r** 和**t** $(r \neq t)$

使得 $\lambda \in \Omega_{rt}(A).$ 证毕

推论 设 $A = (a_{ij}) \in C^{n\times n} (n \ge 2)$, 如果对于所有 ***i≠j*** ，

恒有 $|a_{ii}||a_{jj}| \quad R_i(A)R_j(A)$, 则 $\mathbf{\det A} \quad 0.$

例5.11 讨论矩阵 $A = \begin{pmatrix} 2 & 1.1 & 1 \\ 0.8 & 3 & 2 \\ 1.5 & 1.1 & 3 \end{pmatrix}$ 的奇异性.

解:

$$R_1(A) = 2.1, R_2(A) = 2.8, R_3(A) = 2.6$$

$$|a_{11}||a_{22}| = 6 > 2.1 \times 2.8 = R_1(A)R_2(A)$$

$$|a_{11}||a_{33}| = 6 > 2.1 \times 2.6 = R_1(A)R_3(A)$$

$$|a_{22}||a_{33}| = 9 > 2.8 \times 2.6 = R_2(A)R_3(A)$$

所以 $\det A \neq 0$，即A 非奇异.

5.2 广义特征值问题

定义5.5 设A 是n 阶实对称矩阵，B是n 阶实对称正定矩阵，如果存在$\lambda\in C$与 $O\neq x\in C^n$ 使满足$Ax=\lambda Bx$，则 λ 为矩阵A 相对于矩阵B的特征值；而与λ 相对应的非零解 x 称之为属于λ 的特征向量.

一. 广义特征值问题的等价形式

第一种： $B^{-1}Ax=\lambda x$

第二种：将B进行 **Cholesky** 分解，得到

$B=GG^T$, 于是可表为 $Ax=\lambda GG^Tx$.

令 $y=G^Tx$，则 $x=(G^T)^{-1}y$，且有

$$Sy=\lambda y$$

其中 $S=G^{-1}A(G^{-1})^{\mathrm{T}}$.

二. 特征向量的共轭性

因为 S 是实对称矩阵，所以它的特征值 $\lambda_1,\cdots,\lambda_n$ 均为实数，且有完备特征向量系 $y_1,\cdots,y_n$, 满足

$$y_i^T y_j=\begin{cases}0, & i\neq j\\ 1, & i=j\end{cases}$$

因为 $x_i = \left(G^{-1}\right)^T y_i$，所以

$$x_i^T B x_j = x_i^T G G^T x_j = \left(G^T x_i\right)^T \left(G^T x_j\right) = y_i^T y_j$$

即

$$x_i^T B x_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

定义5.6 满足 $\boldsymbol{x}_i^T \boldsymbol{B} \boldsymbol{x}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$ 的向量系 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 称为按B标准正交向量系；第一式称为B正交(共轭)条件.

性质1 $x_i \neq O(i = 1, \cdots, n)$

性质2 $x_1, \cdots x_n$ 线性无关.

5.3 对称矩阵特征值的极性

5.3.1 实对称矩阵的Rayleigh商的极性

定义5.7 设A是n阶实对称矩阵，$x \in R^n$. 称

$$R(\boldsymbol{x}) = \frac{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} \quad \boldsymbol{x} \neq \boldsymbol{O}$$

为矩阵A的**Rayleigh**商.

性质1 $R(x)$ 是 x 的连续函数.

性质2 $R(x)$ 是 x 的零次齐次函数.

性质3 $x \in L(x_0)(x_0 \neq O)$时，$R(x)$是一常数.

性质4 $R(x)$的最大值和最小值存在，且能够在单位球面 $S = \{x \mid x \in R^n, \|x\|_2 = 1\}$ 上达到.

注：$R(x)$在S 上的最大值与最小值，就是在 $x \neq O$ 的整个区域上的最大值与最小值.

设实对称矩阵A 的特征值按其大小顺序排列为

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

对应的标准正交特征向量系为 $p_1, p_2, \cdots, p_n$

定理5.16 设A为实对称矩阵，则

$$\min_{x \neq O} R(\boldsymbol{x}) = \lambda_1 \quad , \max_{x \neq O} R(\boldsymbol{x}) = \lambda_n$$

证 任取 $O \neq x \in R^n$， 则

$$x = c_1 p_1 + c_2 p_2 + \cdots + c_n p_n \quad (c_1^2 + c_2^2 + \cdots + c_n^2 \neq 0)$$

于是有 $$\boldsymbol{Ax} = c_1\lambda_1\boldsymbol{p}_1 + c_2\lambda_2\boldsymbol{p}_2 + \cdots + c_n\lambda_n\boldsymbol{p}_n$$

$$\boldsymbol{x}^T\boldsymbol{Ax} = c_1^2\lambda_1 + c_2^2\lambda_2 + \cdots + c_n^2\lambda_n$$

$$\boldsymbol{x}^T\boldsymbol{x} = c_1^2 + c_2^2 + \cdots + c_n^2$$

令 $k_i = \dfrac{c_i^2}{c_1^2 + c_2^2 + \cdots + c_n^2}$ $(i = 1, 2, \cdots, n)$，则有

$k_1 + k_2 + \cdots + k_n = 1$ 且 $R(\boldsymbol{x}) = k_1\lambda_1 + k_2\lambda_2 + \cdots + k_n\lambda_n$

由此可得 $$\lambda_1 \le R(\boldsymbol{x}) \le \lambda_n$$

又 $$R(\boldsymbol{p}_1) = \lambda_1,\ R(\boldsymbol{p}_n) = \lambda_n,$$

所以 $$\min_{\boldsymbol{x}\ne\boldsymbol{O}} R(\boldsymbol{x}) = \lambda_1 \quad , \max_{\boldsymbol{x}\ne\boldsymbol{O}} R(\boldsymbol{x}) = \lambda_n$$

证毕

推论1 在 $\|x\|_2 = 1$ 上，$\boldsymbol{p}_1$ 与 $\boldsymbol{p}_n$ 分别是 $R(\boldsymbol{x})$ 的一个极小点和极大点.

推论2 如果 $\lambda_1 = \cdots = \lambda_k (1 \le k \le n)$，则在 $\|x\|_2 = 1$ 上，$R(x)$ 的所有极小点为

$$l_1 p_1 \quad l_2 p_2 \qquad + l_k p_k$$

其中 $l_i \in R(i = 1, \cdots, k)$，且满足 $l_1^2 + l_2^2 + \cdots + l_k^2 = 1$

定理5.17 设 $x \in L(p_r, p_{r+1}, \cdots, p_s)$, $1 \le r \le s \le n$, 则有 $\min\limits_{x \ne O} R(x) = \lambda_r \quad , \max\limits_{x \ne O} R(x) = \lambda_s$

定理5.18 设实对称矩阵A，则A的第k个特征值

$$\lambda_k = \min_{V_k} \max \left\{ \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \,\middle|\, \boldsymbol{x} \in V_k, \|\boldsymbol{x}\|_2 = 1 \right\}$$

其中 V_k 是 R^n 的任意一个 k 维子空间，$1 \le k \le n$

证 构造 R^n 的子空间 $W_k = L(\boldsymbol{p}_k, \boldsymbol{p}_{k+1}, \cdots, \boldsymbol{p}_n)$

则 $\dim W_k = n - k + 1$ 由于 $V_k + W_k \subset R^n$， 所以

$$n \ge \dim(V_k + W_k) = \dim V_k + \dim W_k - \dim(V_k \cap W_k)$$

$$= n + 1 - \dim(V_k \cap W_k)$$

即有 $\dim(V_k \cap W_k) \ge 1$，于是存在 $x_0 \in V_k \cap W_k$，满足 $\|\boldsymbol{x}_0\|_2 = 1$， 且有

$$\boldsymbol{x}_0 = c_k \boldsymbol{p}_k + \cdots + c_n \boldsymbol{p}_n \quad (c_k^2 + \cdots + c_n^2 = 1)$$

故 $\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0 = c_k^2\lambda_k + \cdots + c_n^2\lambda_n \geq \lambda_k$,即

$$\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \middle| \boldsymbol{x} \in V_k, \|\boldsymbol{x}\|_2 = 1\right\} \geq \lambda_k$$

所以 $$\min_{V_k}\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \middle| \boldsymbol{x} \in V_k, \|\boldsymbol{x}\|_2 = 1\right\} \geq \lambda_k$$

令 $V_k^0 = L(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_k)$,取$\boldsymbol{x} \in V_k^0$满足$\|\boldsymbol{x}\|_2 = 1$, 则有

$$x = l_1 p_1 + \cdots + l_k p_k \quad (l_1^2 + \cdots + l_k^2 = 1)$$

于是 $$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = l_1^2\lambda_1 + \cdots + l_k^2\lambda_k \leq \lambda_k$$

所以 $$\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \middle| \boldsymbol{x} \in V_k^0, \|\boldsymbol{x}\|_2 = 1\right\} \leq \lambda_k$$

因而 $$\min_{V_k}\max\left\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \middle| \boldsymbol{x} \in V_k, \|\boldsymbol{x}\|_2 = 1\right\} = \lambda_k$$

证毕

***定理 5.19** 设实对称矩阵$\boldsymbol{A}$ 和$\boldsymbol{A+Q}$的特征值分别为 $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$和$\mu_1 \le \mu_2 \le \cdots \le \mu_n$，则有

$$|\lambda_i - \mu_i| \le \|\boldsymbol{Q}\|_2 \qquad (i = 1, \cdots, n)$$

证 令 $c = \|\boldsymbol{Q}\|_2$, 则 $\boldsymbol{Q}+c\boldsymbol{I}$ 是半正定的,

因为 $\boldsymbol{A}+\boldsymbol{Q}+c\boldsymbol{I}$ 的特征值为

$$\mu_1 + c \le \mu_2 + c \le \cdots \le \mu_n + c$$

所以 $\mu_i + c = \min\limits_{V_i} \max \left\{ \boldsymbol{x}^{\mathrm{T}} (\boldsymbol{A} + \boldsymbol{Q} + c\boldsymbol{I}) \boldsymbol{x} \,\middle|\, \boldsymbol{x} \in V_i, \|\boldsymbol{x}\|_2 = 1 \right\}$

$$\ge \min_{V_i} \max \left\{ \boldsymbol{x}^{\mathrm{T}} (\boldsymbol{A}) \boldsymbol{x} \,\middle|\, \boldsymbol{x} \in V_i, \|\boldsymbol{x}\|_2 = 1 \right\} = \lambda_i$$

因此, $\lambda_i - \mu_i \le c$. 类似地，因为 $\boldsymbol{Q}-c\boldsymbol{I}$ 半负定，所以

$\boldsymbol{A}+\boldsymbol{Q}-c\boldsymbol{I}$ 的特征值为 $\mu_1 - c \le \mu_2 - c \le \cdots \le \mu_n - c$

于是可得

$$\mu_i - c = \min_{V_i} \max \left\{ \boldsymbol{x}^{\mathrm{T}} (\boldsymbol{A} + \boldsymbol{Q} - c\boldsymbol{I}) \boldsymbol{x} \,\middle|\, \boldsymbol{x} \in V_i, \|\boldsymbol{x}\|_2 = 1 \right\}$$

$$\le \min_{V_i} \max \left\{ \boldsymbol{x}^{\mathrm{T}} (\boldsymbol{A}) \boldsymbol{x} \,\middle|\, \boldsymbol{x} \in V_i, \|\boldsymbol{x}\|_2 = 1 \right\} = \lambda_i$$

因此 $\lambda_i - \mu_i \ge -c$. 因而 $|\lambda_i - \mu_i| \le c$. 即

$$|\lambda_i - \mu_i| \le \|Q\|_2.$$

证毕

***定理5.20** 设实对称矩阵$\boldsymbol{A}$，$\boldsymbol{A}+\boldsymbol{Q}$ 和$\boldsymbol{Q}$ 的特征值分别是 $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n, \mu_1 \le \mu_2 \le \cdots \le \mu_n$

和 $\gamma_1 \le \gamma_2 \le \cdots \le \gamma_n$，并定义向量

$$\boldsymbol{u} = (\lambda_1, \lambda_2, \cdots, \lambda_n)^{\mathrm{T}}, \boldsymbol{v} = (\mu_1, \mu_2, \cdots, \mu_n)^{\mathrm{T}}, \boldsymbol{w} = (\gamma_1, \gamma_2, \cdots, \gamma_n)^{\mathrm{T}},$$

则 $\|\boldsymbol{u} - \boldsymbol{v}\|_2 \le \|\boldsymbol{w}\|_2$

二. 广义特征值的极小极大原理

定义5.8 设 A，B 为n 阶实对称矩阵，且B 正定，$x\in R^n$. 称

$$R(x)=\frac{x^T Ax}{x^T Bx},\quad x\neq O$$

为矩阵A 相对于矩阵B 的广义**Rayleigh商.**

定理5.22 非零向量 x_0是 $R(x)$ 的驻点 $\Leftrightarrow$ x_0是 $Ax=\lambda Bx$ 的属于特征值 λ 的特征向量.

证 因为

$$(x^T Bx)R(x)=x^T Ax$$

所以

$$(2Bx)R(x)+(x^T Bx)\frac{dR(x)}{dx}=2Ax$$

$$\frac{dR(x)}{dx}=\frac{2}{x^T Bx}[Ax-R(x)Bx]$$

必要性. 因为 x_0 是 $R(x)$ 的驻点，所以

$$\left.\frac{dR(x)}{dx}\right|_{x=x_0}=0\text{，即 }Ax_0=R(x_0)Bx_0$$

充分性. 设 x_0 满足 $Ax_0=\lambda Bx_0$，则有 $\lambda=R(x_0)$，

所以 x_0 是 $R(x)$ 的驻点. 证毕

推论 若 x_0 是 $Ax=\lambda Bx$ 的特征向量，则 $R(x_0)$ 是与之对应的特征值.

设广义特征值为 $\lambda_1\le\lambda_2\le\cdots\le\lambda_n$，与之对应的按 B 标准正交特征向量系为 $p_1,p_2,\cdots,p_n$.

定理5.23 设 V_k 为 R^n 中的任意一个 k 维子空间，则广义特征问题的第 k 个特征值和第 $n-k+1$ 个特征值具有下列的极小极大性质

$$\lambda_k=\min_{V_k}\left[\max_{O\neq x\in V_k} R(\boldsymbol{x})\right],\quad \lambda_{n-k+1}=\max_{V_k}\left[\min_{O\neq x\in V_k} R(\boldsymbol{x})\right]$$

证 与**定理5.18**相同．只证第二个等式．

因为$-Ax=\lambda Bx$的第k个特征值为 $-\lambda_{n-k+1}$, 应用第一式可得

$$-\lambda_{n-k+1}=\min_{V_k}\left[\max_{O\neq x\in V_k}\frac{\boldsymbol{x}^{\mathrm{T}}(-A)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}}\right]=\min_{V_k}\left[(-1)\min_{O\neq x\in V_k} R(\boldsymbol{x})\right]$$

$$=(-1)\max_{V_k}\left[\min_{O\neq x\in V_k} R(\boldsymbol{x})\right]$$

证毕

推论1 设 V_k 为 R^n中的任意一个 k 维子空间，则对实对称矩阵A 的第 k 个特征值和第$n-k+1$个特征值具有以下性质

$$\lambda_k = \min_{V_k}\left[\max_{O\neq x\in V_k} R(x)\right],\quad \lambda_{n-k+1} = \max_{V_k}\left[\min_{O\neq x\in V_k} R(x)\right]$$

第一式被称为特征值的**极小极大原理**

第二式被称为特征值的**极大极小原理**

推论2 设 V_k是 R^n 的任意一个 k 维子空间，则定理5.23或推论1 的结论可写成如下形式

$$\lambda_k = \max_{V_{n-k+1}}\left[\min_{O\neq x\in V_{n-k+1}} R(\boldsymbol{x})\right],\quad \lambda_{n-k+1} = \min_{V_{n-k+1}}\left[\max_{O\neq x\in V_{n-k+1}} R(\boldsymbol{x})\right]$$

三. 矩阵奇异值的极小极大性质

定理5.24 设 $A \in R_r^{m \times n}$ 的奇异值排列为

$$0 = \sigma_1 = \cdots = \sigma_{n-r} < \sigma_{n-r+1} \leq \cdots \leq \sigma_n$$

则A 的第 k 个奇异值和第 $n-k+1$ 个奇异值具有下列的极性质

$$\sigma_k = \min_{V_k}\left[\max_{O \neq x \in V_k} \frac{\|Ax\|_2}{\|x\|_2}\right], \quad \sigma_{n-k+1} = \max_{V_k}\left[\min_{O \neq x \in V_k} \frac{\|Ax\|_2}{\|x\|_2}\right]$$

其中 V_k 是 R^n 的任一 k 维子空间.

证 设 A^TA 的特征值排列为

$$0 = \lambda_1 = \cdots = \lambda_{n-r} < \lambda_{n-r+1} \leq \cdots \leq \lambda_n$$

由于 $\sigma_i = \sqrt{\lambda_i}(i = 1, \cdots, n)$

所以

$$\sigma_k = \sqrt{\lambda_k} = \left[\min_{V_k}\left(\max_{O \neq x \in V_k} \frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\right)\right]^{\frac{1}{2}}$$

$$= \left[\min_{V_k}\left(\max_{O \neq x \in V_k} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2^2}{\|\boldsymbol{x}\|_2^2}\right)\right]^{\frac{1}{2}} = \min_{V_k}\left(\max_{O \neq x \in V_k} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}\right)$$

证毕

定理5.25 设$A \in R_r^{m\times n}$的奇异值排列同上，$(A+Q) \in R_{r'}^{m\times n}$ 的奇异值排列为 $0=\tau_1=\cdots=\tau_{n-r'}<\tau_{n-r'+1}\le\cdots\le\tau_n$

则有 $$|\sigma_i-\tau_i| \le \|Q\|_2 \quad (i=1,\cdots,n)$$

证 设 $A^{\mathrm{T}}A$ 的相应的标准正交特征向量系为 $x_1,x_2,\cdots,x_n$ ，$V_i^0=L(x_1,x_2,\cdots,x_i)(i=1,\cdots,n)$

则有

$$\tau_i \le \max_{O\ne x\in V_i^0} \frac{\|(A+Q)x\|_2}{\|x\|_2}$$

$$\le \max_{O\ne x\in V_i^0}\left(\frac{\|Ax\|_2}{\|x\|_2}+\frac{\|Qx\|_2}{\|x\|_2}\right)$$

$$\le \max_{O\ne x\in V_i^0}\frac{\|Ax\|_2}{\|x\|_2}+\max_{x\ne O}\frac{\|Qx\|_2}{\|x\|_2}=\sigma_i+\|Q\|_2$$

设 $(A+Q)^T(A+Q)$ 的特征值排列为

$$0 = \mu_1 = \cdots = \mu_{n-r'} < \mu_{n-r'+1} \le \cdots \le \mu_n$$

相应的标准正交特征向量系为 $y_1, y_2, \cdots, y_n$

$$V_i^1 = L(y_1, y_2, \cdots, y_i)(i = 1, \cdots, n)$$

$$\sigma_i \quad \max_{O \quad x \quad V_i^1} \frac{\|Ax\|_2}{\|x\|_2}$$

$$\max_{O \quad x \quad V_i^1} \frac{\|(A+Q)x\|_2}{\|x\|_2} \quad \max_{x \quad O} \frac{\|Qx\|_2}{\|x\|_2} \quad \tau_i \quad \|Q\|_2$$

所以 $|\sigma_i - \tau_i| ? \|Q\|_2 \quad (i \quad 1, \mathrm{L}\ , n)$ 证毕

定理5.26 设 $A \in R_r^{m \times n}$ 和 $(A+Q) \in R_r^{m \times n}$ 的奇异值排列同上，$Q \in R_r^{m \times n}$ 的奇异值排列为

$$0 = \delta_1 = \cdots = \delta_{n-r''} < \delta_{n-r''+1} \leq \cdots \leq \delta_n$$

定义向量

$$u = (\sigma_1, \cdots, \sigma_n), v = (\tau_1, \cdots, \tau_n), w = (\delta_1, \cdots, \delta_n)$$

则有
$$\|u - v\|_2 \leq \|w\|_2$$

5.4 矩阵的直积及应用

一. 直积的概念

定义5.9 设 $\boldsymbol{A} \quad (a_{ij}) \quad C^{m \quad n}$, $\boldsymbol{B} \quad (b_{ij}) \quad C^{p \quad q}$, 则称如下的分块矩阵

$$\boldsymbol{A} \quad \boldsymbol{B} \quad \begin{matrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & & a_{2n}\boldsymbol{B} \\ & & & \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & & a_{mn}\boldsymbol{B} \end{matrix} \quad C^{mp \quad nq}$$

为$\boldsymbol{A}$ 与$\boldsymbol{B}$ 的直积（***Kronecker***积）

例5.12 设 $A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}, B=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}$

求 $A\otimes B, B\otimes A$.

解

$$A\otimes B=\begin{pmatrix}B & 2B\\ 3B & 4B\end{pmatrix}=\begin{pmatrix}0 & 0 & 1 & 0 & 0 & 2\\ 0 & 1 & 0 & 0 & 2 & 0\\ 1 & 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 3 & 0 & 0 & 4\\ 0 & 3 & 0 & 0 & 4 & 0\\ 3 & 0 & 0 & 4 & 0 & 0\end{pmatrix}$$

$$B \quad A \quad \begin{matrix} & & A \\ & A & \\ A & & \end{matrix} \quad \begin{matrix} 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 4 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \end{matrix}$$

性质1 $k(A \otimes B) = (kA) \otimes B = A \otimes (kB)$

性质2 设 A_1与A_2 为同阶矩阵，则

$$(A_1 \quad A_2) \quad B \quad A_1 \quad B \quad A_2 \quad B$$

$$B \quad (A_1 \quad A_2) \quad B \quad A_1 \quad B \quad A_2$$

性质3 $(A \quad B) \quad C \quad A \quad (B \quad C)$

证　由于 $(\boldsymbol{A} \otimes \boldsymbol{B})_{ij} = a_{ij}\boldsymbol{B} = \begin{pmatrix} a_{ij}b_{11} & \cdots & a_{ij}b_{1q} \\ \vdots & & \vdots \\ a_{ij}b_{p1} & \cdots & a_{ij}b_{pq} \end{pmatrix}$

所以 $(\boldsymbol{A} \otimes \boldsymbol{B})_{ij} \otimes \boldsymbol{C} = \begin{pmatrix} a_{ij}b_{11}\boldsymbol{C} & \cdots & a_{ij}b_{1q}\boldsymbol{C} \\ \vdots & & \vdots \\ a_{ij}b_{p1}\boldsymbol{C} & \cdots & a_{ij}b_{pq}\boldsymbol{C} \end{pmatrix} = a_{ij}(\boldsymbol{B} \otimes \boldsymbol{C})$

$$\begin{array}{lll} & (\boldsymbol{A} \quad \boldsymbol{B})_{11} \quad \boldsymbol{C} & (\boldsymbol{A} \quad \boldsymbol{B})_{1n} \quad \boldsymbol{C} \\ (A \quad B) \quad \boldsymbol{C} & & \\ & (\boldsymbol{A} \quad \boldsymbol{B})_{m1} \quad \boldsymbol{C} & (\boldsymbol{A} \quad \boldsymbol{B})_{mn} \quad \boldsymbol{C} \end{array}$$

$$\begin{array}{lll} a_{11}(\boldsymbol{B} \quad \boldsymbol{C}) & a_{1n}(\boldsymbol{B} \quad \boldsymbol{C}) & \\ & & \boldsymbol{A} \quad (\boldsymbol{B} \quad \boldsymbol{C}) \\ a_{m1}(\boldsymbol{B} \quad \boldsymbol{C}) & a_{mn}(\boldsymbol{B} \quad \boldsymbol{C}) & \end{array}$$

性质4 设 $A_1=\left(a_{ij}^{(1)}\right)_{m_1\times n_1}, A_2=\left(a_{ij}^{(2)}\right)_{m_2\times n_2}$

$\boldsymbol{B}_1 \quad (b_{ij}^{(1)})_{p_1 \quad q_1}, \boldsymbol{B}_2 \quad (b_{ij}^{(2)})_{p_2 \quad q_2}, n_1=m_2, q_1=p_2$，则

$$(\boldsymbol{A}_1 \quad \boldsymbol{B}_1)(\boldsymbol{A}_2 \quad \boldsymbol{B}_2) \quad (\boldsymbol{A}_1\boldsymbol{A}_2) \quad (\boldsymbol{B}_1\boldsymbol{B}_2)$$

证 由于

$$[\text{左端}]_{ij}=(a_{i1}^{(1)}\boldsymbol{B}_1,\cdots,a_{in_1}^{(1)}\boldsymbol{B}_1)\begin{pmatrix} a_{1j}^{(2)}\boldsymbol{B}_2 \\ \vdots \\ a_{m_2j}^{(2)}\boldsymbol{B}_2 \end{pmatrix}$$

$$=\sum_{k=1}^{n_1}a_{ik}^{(1)}\boldsymbol{B}_1a_{kj}^{(2)}\boldsymbol{B}_2=\left(\sum_{k=1}^{n_1}a_{ik}^{(1)}a_{kj}^{(2)}\right)(\boldsymbol{B}_1\boldsymbol{B}_2)$$

$$=(\boldsymbol{A}_1\boldsymbol{A}_2)_{ij}(\boldsymbol{B}_1\boldsymbol{B}_2)=[\text{右端}]_{ij}$$

性质5 设 $\boldsymbol{A}_{m\times m}$与$\boldsymbol{B}_{n\times n}$ 都可逆，则$(\boldsymbol{A} \quad \boldsymbol{B})^{\ 1} \quad \boldsymbol{A}^{\ 1} \quad \boldsymbol{B}^{\ 1}$

性质6 设 $A_{m \quad m}$ 与 $B_{n \quad n}$ 都是上三角（下三角）矩阵，则 $A \quad B$ 也是上三角（下三角）矩阵.

性质7 $A \quad B^{H} \quad A^{H} \quad B^{H}$

性质8 设 $A_{m\times m}$ 与 $B_{n\times n}$ 都是正交(酉)矩阵，则 $A\otimes B$ 也是正交（酉）矩阵.

性质9 $rank(A \quad B) \quad (rankA)(rankB)$

证 设 $rankA \quad r_1$，$rankB \quad r_2$.

因为存在可逆矩阵 P_1 与 Q_1，使

$$A = P_1 A_1 Q_1，\ A_1 = \begin{pmatrix} I_{r_1} & O \\ O & O \end{pmatrix}$$

存在可逆矩阵 P_2 与 Q_2，使

$$B = P_2 B_1 Q_2, \quad B_1 = \begin{pmatrix} I_{r_2} & O \\ O & O \end{pmatrix}$$

因为 $$A \otimes B = (P_1 A_1 Q_1) \otimes (P_2 B_1 Q_2)$$

$$= (P_1 \otimes P_2)(A_1 \otimes B_1)(Q_1 \otimes Q_2)$$

所以 $rank(A \quad B) \quad rank(A_1 \quad B_1)$

因为

$$B_1 \qquad O$$

$$A_1 \quad B_1$$

$$B_1 \quad O$$

$$O \qquad O \quad O$$

所以 $rank(A_1 \quad B_1) \quad r_1 r_2$

即 $rank(A \quad B) \quad (rank A)(rank B)$

性质10 设 $A\in C^{m\times m}$, $B\in C^{n\times n}$, 则有 A Ä $B \sim B \quad A$.

证 因为 $A \quad B \quad (A \quad I_n)(I_m \quad B)$, 对矩阵

$$A \quad I_n \quad \begin{array}{cc|c|cc} a_{11} & & & a_{1m} & \\ & a_{11} & & & a_{1m} \\ \hline & & & & \\ \hline a_{m1} & & & a_{mm} & \\ & a_{m1} & & & a_{mm} \end{array}$$

调换它的行，且同时调换它的同序号的列，可使之变为

a_{11} a_{1m}

a_{m1} a_{mm}

$I_n \quad A$

a_{11} a_{1m}

a_{m1} a_{mm}

即存在可逆矩阵$\boldsymbol{P}$，使得

$$\boldsymbol{P}^{\;1}(\boldsymbol{A} \quad \boldsymbol{I}_n)\boldsymbol{P} \quad \boldsymbol{I}_n \quad \boldsymbol{A}$$

同理，有 $$\boldsymbol{P}^{\;1}(\boldsymbol{I}_m \quad \boldsymbol{B})\boldsymbol{P} \quad \boldsymbol{B} \quad \boldsymbol{I}_m$$

于是 $\quad A \quad B \quad (A \quad I_n)(I_m \quad B)$

$$P(I_n \quad A)P^{\ 1}P(B \quad I_m)P$$

$$P(B \quad A)P^{\ 1}$$

所以 $A \quad B$ 与 $B \quad A$ 相似. 证毕

设二元多项式 $f(x,y)=\sum_{i,j=0}^{l} c_{ij}x^i y^j$

及矩阵 $A_{m\times m}$ 与 $B_{n\times n}$，定义mn阶矩阵

$$f(A,B)=\sum_{i,j=0}^{l} c_{ij}A^i \otimes B^j$$

其中 $\quad A^0=I_m, B^0=I_n.$

定理5.27 设 $A_{m' \ m}$的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，$B_{n \ n}$ 的特征值为 $\mu_1,\mu_2,\cdots,\mu_n$,则 $f(A,B)$的全体特征值为 $f(\lambda_i,\mu_j)(i \quad 1,2, \quad ,m; j \quad 1,2, \quad ,n)$.

证 存在可逆矩阵 $P_{m \ m}$与$Q_{n \ n}$ 使

$$P^{\ 1}AP \qquad \begin{matrix} \lambda_1 & \\ & \\ & \lambda_m \end{matrix} \qquad T_1,$$

$$Q^{-1}BQ \qquad \begin{matrix} \mu_1 & \\ & \\ & \mu_n \end{matrix} \qquad T_2$$

因为 $(P \quad Q)^{1} A \quad B(P \quad Q)$

$$(P^{1}AP) \quad (Q^{1}BQ) \quad T_1 \quad T_2$$

所以 $(P \quad Q)^{1} f(A,B)(P \quad Q) \quad f(T_1,T_2)$

因为

$$\lambda_1^i T_2^j$$

$$T_1^i \quad T_2^j \qquad ,$$

$$\lambda_m^i T_2^j$$

$$\lambda_k^i \mu_1^j$$

$$\lambda_k^i T_2^j$$

$$\lambda_k^i \mu_n^j$$

所以 $f(A,B)$ 的特征值为 $f(\lambda_i,\mu_j)(1\le i\le m,1\le j\le n)$

推论1 设 $A_{m\times m}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，$B_{n\times n}$ 的特征值为 $\mu_1,\mu_2,\cdots,\mu_n$，则 $A\otimes B$ 的全体特征值为

$$\lambda_i\mu_j(i=1,2,\cdots,m;\ j=1,2,\cdots,n)$$

推论2 设 $\boldsymbol{A}\in\mathbf{C}^{m\times m},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，则有

$$\det(A\otimes B)=(\det A)^n(\det B)^m$$

推论3 设 $\boldsymbol{A}\in\mathbf{C}^{m\times m},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，则有

$$tr(A\otimes B)=(trA)(trB)$$

二. 线性矩阵方程的可解性

线性矩阵方程 $$\sum_{i=1}^{l} \boldsymbol{A}_i \boldsymbol{X} \boldsymbol{B}_i = \boldsymbol{F}$$

其中 $\boldsymbol{A}_i \in C^{m\times p}, \boldsymbol{B}_i \in C^{q\times n}, \boldsymbol{F} \in C^{m\times n}$ 为已知矩阵, $\boldsymbol{X}\in \boldsymbol{C}^{p\times q}$ 为未知矩阵.

定义 设矩阵 $\boldsymbol{X}\in \boldsymbol{C}^{p\times q}$ 的第$\boldsymbol{i}$行为 $\boldsymbol{x}_i^T (i = 1,\cdots,p)$,则向量$(x_1^{\mathrm{T}}, x_2^{\mathrm{T}}, \cdots, x_p^{\mathrm{T}})^{\mathrm{T}}$称为矩阵 $\boldsymbol{X}$ 按行拉直记作 $\overline{vec}(X)$

用直积表示 $\overline{vec}(AXB)$.

因为

$$AXB = \begin{pmatrix} \vdots \\ (a_{i1}x_1^T + \cdots a_{ip}x_p^T)B \\ \vdots \end{pmatrix}$$

从而

$$\overline{vec}(AXB) = \begin{pmatrix} \vdots \\ B^T(a_{i1}x_1 + \cdots + a_{ip}x_p) \\ \vdots \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}B^T & \cdots & a_{1p}B^T \\ \vdots & & \vdots \\ a_{m1}B^T & \cdots & a_{mp}B^T \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

$$\overline{vec}(\boldsymbol{AXB}) \quad (\boldsymbol{A} \quad \boldsymbol{B}^T)\overline{vec}(\boldsymbol{X})$$

定理5.28 方程 $\sum_{i=1}^{l} \boldsymbol{A}_i \boldsymbol{X} \boldsymbol{B}_i = \boldsymbol{F}$ 有解 $\Leftrightarrow$

$$\overline{vec}(\boldsymbol{F}) \in R(\sum_{i=1}^{l} \boldsymbol{A}_i \quad \boldsymbol{B}_i^T)$$

定理5.29 设 $\boldsymbol{A}_{m \times m}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_m$, $\boldsymbol{B}_{\boldsymbol{n}\times\boldsymbol{n}}$的特征值为 $\mu_1, \mu_2, \cdots, \mu_n$, 则方程 $\boldsymbol{AX} + \boldsymbol{XB} = \boldsymbol{F}$ 有唯一解 $\Leftrightarrow \lambda_i + \mu_j \neq 0 (i = 1, \cdots, m; j = 1, \cdots, n)$

推论1 设 $\boldsymbol{A}_{m\times m}$ 的特征值为 $\lambda_1,\lambda_2,\mathrm{L}\ ,\lambda_m$, $\boldsymbol{B}_{n\times n}$的特征值为$\mu_1,\mu_2,\cdots,\mu_n$, 则齐次方程 $\boldsymbol{AX} \quad \boldsymbol{XB} \quad \boldsymbol{O}$

有非零解⇔存在$\boldsymbol{i_0}$与$\boldsymbol{j_0}$满足$\lambda_{i_0}+\mu_{j_0}=0$.

推论2 设 $\boldsymbol{A}$是$\boldsymbol{m}$阶方阵，则齐次方程组$\boldsymbol{AX-XA=0}$

一定有非零解.

例5.13 设 $A\in C^{m\times m}, B\in C^{n\times n}, F\in C^{m\times n}$， 如果$A$与$B$无公共特征值，则 $\begin{pmatrix} A & F \\ O & B \end{pmatrix}$ 相似于 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$.

证 设 $P=\begin{pmatrix} I_m & X \\ O & I_n \end{pmatrix}, X\in C^{m\times n}$,则

$$P^{-1}=\begin{pmatrix} I_m & -X \\ O & I_n \end{pmatrix}, P^{-1}\begin{pmatrix} A & F \\ O & B \end{pmatrix}P=\begin{pmatrix} A & F+XB-AX \\ O & B \end{pmatrix}$$

因为$\lambda_i(A)-\mu_j(B)\neq O$，所以 $AX-BX=F$ 有唯一解 $X\in C^{m\times n}$, 于是有

$$P=\begin{pmatrix} I_m & X \\ O & I_n \end{pmatrix}, P^{-1}\begin{pmatrix} A & F \\ O & B \end{pmatrix}P=\begin{pmatrix} A & O \\ O & B \end{pmatrix}$$

证毕

定理5.30 设 $A_{m\times m}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，$B_{n\times n}$ 的特征值为 $\mu_1,\mu_2,\cdots,\mu_n$，则

（1）方程 $\sum_{k=0}^{l} A^k X B^k = F$ 有唯一解 $\Leftrightarrow$

$$1+\lambda_i\mu_j+\cdots+\left(\lambda_i\mu_j\right)^l \neq 0 (i=1,\cdots,m;\ j=1,\cdots,n);$$

（2）齐次方程 $\sum_{i=1}^{l} A^k X B^k = O$ 有非零解 $\Leftrightarrow$

存在 i_0 与 j_0 使 $1+\lambda_{i_0}\mu_{j_0}+\cdots+\left(\lambda_{i_0}\mu_{j_0}\right)^l = 0.$

引理3 设 $A\in C^{m\times m}$, $B\in C^{n\times n}$, $F\in C^{m\times n}$，如果A与B的特征值的实部都小于零，则积分 $\int_0^{+\infty} e^{At}Fe^{Bt}dt$ 存在.

证 设A的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，则存在可逆矩阵 $P\in C^{m\times m}$，使得

$$P^{-1}AP=\begin{pmatrix} \lambda_1 & \delta_1 & & & \\ & \ddots & \ddots & & \\ & & & & \\ & & \lambda_{m-1} & \delta_{m-1} & \\ & & & \lambda_m \end{pmatrix}$$

其中 $\delta_i=0$ 或1．从而

$$e^{At} = P\begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_m t} \end{pmatrix} T_A P^{-1}$$

其中 T_A 表示单位上三角矩阵，它的非零元形式为

$$at^k \left(0 \le k \le m, a \in R\right)$$

设B 的特征值为 $\mu_1, \mu_2, \cdots, \mu_n$, 与上类似，有

$$e^{Bt} = Q\begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_m t} \end{pmatrix} T_B Q^{-1}$$

其中矩阵 $\mathbf{Q} \in \mathbf{C}^{\mathbf{n} \times \mathbf{n}}$ 可逆，T_B也表示单位上三角矩阵，

它的非零元形式为 $bt^k (0 \le k \le n, b \in R)$.

由于

$$e^{At}Fe^{Bt}=P\begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_m t} \end{pmatrix}T_A P^{-1}FQ\begin{pmatrix} e^{\mu_1 t} & & \\ & \ddots & \\ & & e^{\mu_n t} \end{pmatrix}T_B Q^{-1}$$

且积分$\int_0^{+\infty} t^k e^{(\lambda_i+\mu_j)t}dt(k \geq 0)$都存在，所以积分

$\int_0^{+\infty} e^{At}Fe^{Bt}dt$ 存在.

定理5.31 设 $A\in C^{m\times m}$， $B\in C^{n\times n}$，$F\in C^{m\times n}$ 的特征值之和不等于零，如果积分 $\int_0^{+\infty} e^{At}Fe^{Bt}dt$ 存在，则方程 $AX+XB=F$ 的唯一解是 $X=-\int_0^{+\infty} e^{At}Fe^{Bt}dt.$

证 令 $Y(t)=e^{At}Fe^{Bt}$，则有

$$\frac{dY(t)}{dt}=AY(t)+Y(t)B,\quad Y(t)\big|_{t=0}=F$$

由积分 $\int_0^{+\infty} Y(t)dt$ 的存在知，$\lim_{t\to\infty} Y(t)=O.$

所以 $Y(t)\big|_0 \quad A \quad {}_0 \quad Y(t)dt \quad {}_0 \quad Y(t)dt\ B$

即 $AX+XB=F$ 证毕

推论1 设 $A \in C^{m\times m}, B \in C^{n\times n}$ 的特征值满足

$$\mathrm{Re}(\lambda_i) < 0 (i = 1,\cdots,m),\ \mathrm{Re}(\mu_j) < 0 (j = 1,\cdots,n)$$

则方程$AX+XB=F$ 的唯一解为

$$X \qquad {}_{0} \quad e^{At}Fe^{Bt}dt.$$

推论2 设 $A\in C^{m\times m}$ 的特征值满足 $\mathrm{Re}(\lambda_i)<0$ ，则方程 $A^TX+XA=-F$ 的唯一解为 $X=\int_0^{+\infty}e^{A^Tt}Fe^{At}dt.$ 如果 $F\in C^{m\times m}$ 是正定矩阵，则解矩阵 X 也是正定的.

证 只需证后一结论. 设 $O\neq x\in C^n$, 由于矩阵 e^{At} 可逆，所以 $e^{At}x\neq O$ ，于是可得

$$(e^{At}x)^T F(e^{At}x)>0$$

从而有

$$x^T Xx=\int_0^{+\infty}(e^{At}x)^T F(e^{At}x)dt>0$$

故 X 是正定矩阵. 证毕

第五章　　总结

一. 特征值的估计

1. 盖尔圆定理

2. 应用（根的隔离）

二. 对称矩阵特征值的极性

1. 实对称矩阵Rayleigh商的极性

2. 广义特征值的极小极大原理

三. 矩阵的直积及其应用