# 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

### 1. 特征值界的估计

•**定理 1：** 设$A\in R^{n\times n}$，λ为 A 的任意特征值，则有$|Im(\lambda)|\le M\sqrt{\frac{n(n-1)}{2}}$，其中，$M=\max\limits_{1\le i,j\le n}\left|\frac{a_{ij}-a_{ji}}{2}\right|$

证明：设 x 为 A 的属于特征值 λ 的单位特征向量，即$\mathbf{Ax}=\lambda\mathbf{x}$，$\mathbf{x}^H\mathbf{x}=1$，则

$$\lambda=\mathbf{x}^H\mathbf{A}\mathbf{x}\ \rightarrow\ \bar{\lambda}=\left(\overline{\mathbf{x}^H\mathbf{A}\mathbf{x}}\right)=\left(\mathbf{x}^H\mathbf{A}\mathbf{x}\right)^H=\mathbf{x}^H\mathbf{A}^H\mathbf{x}$$

$$\lambda-\bar{\lambda}=2j\mathrm{Im}(\lambda)=\mathbf{x}^H\left(\mathbf{A}-\mathbf{A}^H\right)\mathbf{x}=\mathbf{x}^H\left(\mathbf{A}-\mathbf{A}^T\right)\mathbf{x}$$

将 x 写成$\mathbf{x}=[\xi_1,\xi_2,\cdots,\xi_n]^T$

$$\mathbf{x}^H\left(\mathbf{A}-\mathbf{A}^T\right)\mathbf{x}=\sum_{i=1}^n\sum_{j=1}^n\bar{\xi}_i\left(a_{ij}-a_{ji}\right)\xi_j$$

$$2|\mathrm{Im}(\lambda)|=\left|\sum_{i=1}^n\sum_{j=1}^n\bar{\xi}_i\left(a_{ij}-a_{ji}\right)\xi_j\right|$$

$$\le\sum_{i=1}^n\sum_{j=1}^n\left|\bar{\xi}_i\left(a_{ij}-a_{ji}\right)\xi_j\right|$$

$$=\sum_{i,j=1}^{n}{}'\left|\xi_i\xi_j\right|\left|a_{ij}-a_{ji}\right|\qquad\left(\sum{}'\text{ 表示不含 } i=j\right)$$

$$\le 2M\sum_{i,j=1}^{n}{}'\left|\xi_i\xi_j\right|$$

$$|\mathrm{Im}(\lambda)|^2\le M^2\left(\sum_{i,j=1}^{n}{}'\left|\xi_i\xi_j\right|\right)^2$$

$$\le M^2n(n-1)\sum_{i,j=1}^{n}{}'\left|\xi_i\xi_j\right|^2$$

$$=M^2n(n-1)\sum_{i,j=1}^{n}{}'\left|\xi_i\right|^2\left|\xi_j\right|^2$$

$$\sum_{i,j=1}^{n}{}'\left|\xi_i\right|^2\left|\xi_j\right|^2=\sum_{i,j=1}^{n}\left|\xi_i\right|^2\left|\xi_j\right|^2-\sum_{i=1}^{n}\left|\xi_i\right|^4\le\sum_{i=1}^{n}\left|\xi_i\right|^2-\sum_{i=1}^{n}\left|\xi_i\right|^4$$

$$=\sum_{i=1}^{n}\left|\xi_i\right|^2\left(1-\left|\xi_i\right|^2\right)$$

不妨写为：
$$=\left|\xi_1\right|^2\left(1-\left|\xi_1\right|^2\right)+\left|\xi_2\right|^2\left(1-\left|\xi_2\right|^2\right)+\sum_{i=3}^{n}\left|\xi_i\right|^2\left(1-\left|\xi_i\right|^2\right)$$

$$\le\left(\frac{\left|\xi_1\right|^2+\left(1-\left|\xi_1\right|^2\right)}{2}\right)^2+\left(\frac{\left|\xi_2\right|^2+\left(1-\left|\xi_2\right|^2\right)}{2}\right)^2+\sum_{i=3}^{n}\left|\xi_i\right|^2\left(1-\left|\xi_i\right|^2\right)$$

$$\le\frac{1}{2}$$

取等号的条件为$\left|\xi_1\right|^2=\left|\xi_2\right|^2=\frac{1}{2}$，但$\|\mathbf{x}\|^2=1$，所以其它$\left|\xi_i\right|^2=0$

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

•**定理 2：** 设$A\in C^{n\times n}$，λ为 A 的任意特征值，则有：

$$|\lambda|\le\|A\|_{m_\infty}=n\cdot\max_{1\le i,j\le n}|a_{ij}|$$

$$|Re(\lambda)|\le\frac{1}{2}\|A+A^H\|_{m_\infty}=n\cdot\max_{1\le i,j\le n}|a_{ij}+\overline{a_{ji}}|$$

$$|Im(\lambda)|\le\frac{1}{2}\|A-A^H\|_{m_\infty}==n\cdot\max_{1\le i,j\le n}|a_{ij}-\overline{a_{ji}}|$$

**例 5.1** 估计矩阵 $\mathbf{A}=\begin{bmatrix}1 & -0.8\\ 0.5 & 0\end{bmatrix}$的特征值的上界.

**解** 应用定理 5.2，可得 $|\lambda|\leqslant 2$，$|\mathrm{Re}(\lambda)|\leqslant 2$，$|\mathrm{Im}(\lambda)|\leqslant 1.3$. 应用定理 5.1 估计 $\mathrm{Im}(\lambda)$，有

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

实际上，A 的两个特征值是 $\lambda_{1,2}=\frac{1}{2}(1\pm \mathrm{j}\sqrt{0.6})$，从而 $|\lambda_{1,2}|=0.632\,456$，$|\mathrm{Re}(\lambda_{1,2})|=0.5$，$|\mathrm{Im}(\lambda_{1,2})|=0.387\,298$.

例 5.1 表明，在估计实矩阵的特征值的虚部上界时，定理 5.1 的结果优于定理 5.2 的结果.

## 2. 盖尔圆法

•**定义：** 设$A=(a_{ij})_{n\times n}\in C^{n\times n}$，称由不等式$|z-a_{ii}|\le R_i$在复平面上确定的区域为矩阵 A 的第 i 个盖尔圆，用$G_i$表示，其中：

$$R_i=R_i(A)=\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ij}|$$

称为盖尔圆$G_i$的半径。

•**定理 3：** 矩阵$A=(a_{ij})_{n\times n}\in C^{n\times n}$的一切特征值都在它的 n 个盖尔圆的并集之中。

**证明：设 $\mathbf{A}=(\mathbf{a}_{ij})_{n\times n}\in\mathbf{C}^{n\times n}$，$\boldsymbol{\lambda}$ 为 A 的某一个特征值，x 为相应的特征向量，将 x 写成 $\mathbf{x}=[\xi_1,\xi_2,\cdots,\xi_n]^T$，**

**设 $|\xi_{i_0}|=\max|\xi_i|$**

**由 $\mathbf{Ax}=\boldsymbol{\lambda}\mathbf{x}$，考虑 $\mathbf{i}_0$ 行**

$$\sum_{j=1}^{n}a_{i_0j}\xi_j=\lambda\xi_{i_0}$$

$$(\lambda-a_{i_0i_0})\xi_{i_0}={\sum_{j=1}^{n}}'a_{i_0j}\xi_j\quad(j\ne i_0)$$

$$|\lambda-a_{i_0i_0}|=\left|{\sum_{j=1}^{n}}'a_{i_0j}\frac{\xi_j}{\xi_{i_0}}\right|\le{\sum_{j=1}^{n}}'|a_{i_0j}|\left|\frac{\xi_j}{\xi_{i_0}}\right|\le R_{i_0}$$

**对于 A 的特征值λ，一定存在 $\mathbf{i}_0$ $(1\le i_0\le n)$，使λ落在 A 的第 $\mathbf{i}_0$ 个盖尔圆中，对于每个特征值都有相同的结论。**

•**例题：**

**例 5.5** 估计矩阵 $A=\begin{bmatrix}1 & 0.1 & 0.2 & 0.3\\ 0.5 & 3 & 0.1 & 0.2\\ 1 & 0.3 & -1 & 0.5\\ 0.2 & -0.3 & -0.1 & -4\end{bmatrix}$ 的特征值范围.

**解** $A$ 的 4 个盖尔圆为

$$|z-1|\leqslant 0.1+0.2+0.3=0.6$$
$$|z-3|\leqslant 0.5+0.1+0.2=0.8$$
$$|z+1|\leqslant 1+0.3+0.5=1.8$$
$$|z+4|\leqslant 0.2+0.3+0.1=0.6$$

画在复平面上如图 5.1 所示.

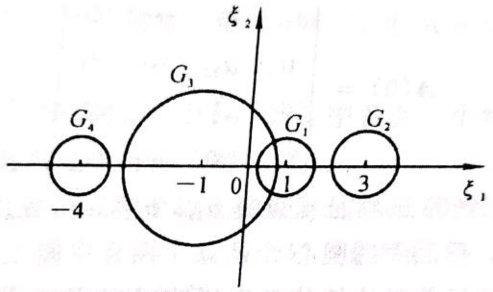
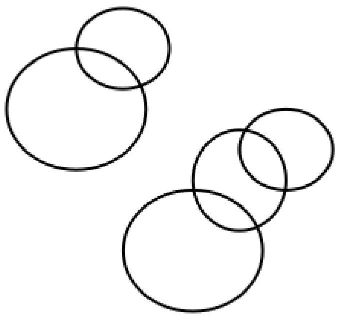



图 5.1

**于是，$A$ 的全部特征值就在这四个盖尔圆并起来的区域之中.**

•**定理 4：**由矩阵 A 的所有盖尔圆组成的连通部分中任取一个，如果它是由 k 个盖尔圆构成的，则在这个连通部分中有且仅有 A 的 k 个特征值（盖尔圆想重时重复计算，特征值相同时也重复计算）

**证明：**设 $A=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}\in C^{n\times n}$，



令 $B(u)=\begin{bmatrix}a_{11} & ua_{12} & ua_{13} & \cdots & ua_{1n}\\ ua_{21} & a_{22} & ua_{23} & \cdots & ua_{2n}\\ ua_{31} & ua_{32} & a_{33} & \cdots & ua_{3n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ ua_{n1} & ua_{n2} & ua_{n3} & \cdots & a_{nn}\end{bmatrix}$

$0\leq u\leq 1$，$B(0)=\mathrm{diag}[a_{11}\quad a_{22}\quad \cdots\quad a_{nn}]$，$B(1)=A$

**$B(u)$ 的特征多项式是 u 的多项式，其特征值是 u 的连续函数，观察 u（$0\leq u\leq 1$）变化的过程中 $B(u)$ 特征值的变化，特征值只能在盖尔圆连通的子集内变动，而不能跨出连通子集。**

**由此可见，由 K 个盖尔圆组成的连通子集恰好包含 K 个特征值。**

**应该注意到：连通的这些盖尔圆中，有些盖尔圆可能包含两个或多个特征值，而其它盖尔圆中可能无特征值。**

•**推论：**

（1）孤立盖尔圆中恰好包含一个特征值

（2）实矩阵的孤立盖尔圆恰好包含一个实特征值

（3）盖尔圆方法中盖尔圆半径可以按列求和。（因为方阵转置后特征值不变）

（4）盖尔圆半径变为 $R_i=\sum_{j=1}^{n}{}'\frac{\alpha_i}{\alpha_j}|a_{ij}|$，两个盖尔圆定理仍然成立

## 5.2 广义特征值与极小极大原理

### 1. 广义特征值问题

**•定义：**

设 A、B 为 n 阶方阵，若存在数$\lambda$，使得方程$Ax=\lambda Bx$存在非零解，则称$\lambda$为 A 相对于 B 的广义特征值，x 为 A 相对于 B 的属于广义特征值$\lambda$的特征向量。

注意：其是标准特征值问题的推广，当 B = I（单位矩阵）时，广义特征值问题退化为标准特征值问题；特征向量是非零的；广义特征值的求解：

$$(A-\lambda B)x=0 \quad \text{或者} \quad (\lambda B-A)x=0$$

$$\rightarrow \quad \text{特征方程} \ \det(A-\lambda B)=0$$

求得$\lambda$后代回原方程$Ax=\lambda Bx$可求出 x

**•等价表述：**

(1)B 正定，$B^{-1}$存在 $\rightarrow B^{-1}Ax=\lambda x$，广义特征值问题化为了标准特征值问题，但一般来说，$B^{-1}A$一般不再是厄米矩阵。

(2)B 厄米，存在 Cholesky 分解，$B=GG^H$，G 满秩

$$Ax=\lambda GG^Hx \qquad \text{令} G^Hx=y$$

则 $$G^{-1}A\left(G^H\right)^{-1}y=\lambda y$$ 也成为标准特征值问题。

$G^{-1}A\left(G^H\right)^{-1}$为厄米矩阵，广义特征值是实数，可以按大小顺序排列$\lambda_1\le\lambda_2\le\cdots\le\lambda_n$，一定存在一组正交归一的特征向量，即存在$y_1,y_2\cdots,y_n$满足

$$G^{-1}A\left(G^H\right)^{-1}y_i=\lambda y_i$$

$$y_i^Hy_j=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\ne j\end{cases}$$

还原为$x_i=\left(G^H\right)^{-1}y_i$ (i=1,2,···,n)，则

$$y_i^Hy_j=\left(x_i^HG\right)\left(G^Hx_j\right)=x_i^HBx_j=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\ne j\end{cases} \qquad \text{（带权正交）}$$

## 2. 瑞丽商

A、B 为 n 阶厄米矩阵，且 B 正定，称$\mathbf{R}(\mathbf{x})=\dfrac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{B}\mathbf{x}}(\mathbf{x}\neq 0)$为 A 相对于 B 的广义瑞利商。

$\mathbf{x}_1,\mathbf{x}_2\cdots,\mathbf{x}_n$线性无关，所以，$\forall \mathbf{x}\in \mathbf{C}^n$，存在$\mathbf{a}_1,\mathbf{a}_2\cdots,\mathbf{a}_n\in \mathbf{C}$，使得$\mathbf{x}=\sum_{i=1}^{n}\mathbf{a}_i\mathbf{x}_i$

$$\mathbf{x}^H\mathbf{B}\mathbf{x}=\left(\sum_{i=1}^{n}\mathbf{a}_i\mathbf{x}_i\right)^H\mathbf{B}\left(\sum_{j=1}^{n}\mathbf{a}_j\mathbf{x}_j\right)=\sum_{i,j=1}^{n}\bar{\mathbf{a}}_i\mathbf{a}_j\mathbf{x}_i^H\mathbf{B}\mathbf{x}_j=\sum_{i=1}^{n}|\mathbf{a}_i|^2$$

$$\mathbf{x}^H\mathbf{A}\mathbf{x}=\sum_{i,j=1}^{n}\bar{\mathbf{a}}_i\mathbf{a}_j\mathbf{x}_i^H\mathbf{A}\mathbf{x}_j=\sum_{i,j=1}^{n}\bar{\mathbf{a}}_i\mathbf{a}_j\mathbf{x}_i^H\lambda_j\mathbf{B}\mathbf{x}_j=\sum_{i=1}^{n}\lambda_i|\mathbf{a}_i|^2$$

$$\therefore\quad \mathbf{R}(\mathbf{x})=\frac{\sum_{i=1}^{n}\lambda_i|\mathbf{a}_i|^2}{\sum_{i=1}^{n}|\mathbf{a}_i|^2}$$

有：$\min\limits_{x\neq 0}\mathbf{R}(\mathbf{x})=\lambda_1\qquad \max\limits_{x\neq 0}\mathbf{R}(\mathbf{x})=\lambda_n$

证明：$\mathbf{R}(\mathbf{x})=\dfrac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{B}\mathbf{x}}=\dfrac{(\mathbf{k}\mathbf{x})^H\mathbf{A}(\mathbf{k}\mathbf{x})}{(\mathbf{k}\mathbf{x})^H\mathbf{B}(\mathbf{k}\mathbf{x})}$　　k 为非零常数

可取$\mathbf{k}=\dfrac{1}{\|\mathbf{x}\|}$，　$\|\mathbf{k}\mathbf{x}\|=1$

$\therefore\quad \mathbf{R}(\mathbf{x})=\left.\dfrac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{B}\mathbf{x}}\right|_{\|\mathbf{x}\|=1}$　　（闭区域）

当$\mathbf{x}=\mathbf{x}_1$或$\mathbf{a}_i=0(i=2,3,\cdots,n)$时，$\mathbf{R}(\mathbf{x})=\lambda_1$

$$\lambda_i\geq\lambda_1\qquad \mathbf{R}(\mathbf{x})\geq\lambda_1\frac{\sum_{i=1}^{n}|\mathbf{a}_i|^2}{\sum_{i=1}^{n}|\mathbf{a}_i|^2}=\lambda_1$$

$$\therefore\quad \min_{x\neq 0}\mathbf{R}(\mathbf{x})=\lambda_1$$

另一方面，$\lambda_i\leq\lambda_n$　　$\mathbf{R}(\mathbf{x})\leq\lambda_n\dfrac{\sum_{i=1}^{n}|\mathbf{a}_i|^2}{\sum_{i=1}^{n}|\mathbf{a}_i|^2}=\lambda_n$

$$\therefore\quad \max_{x\neq 0}\mathbf{R}(\mathbf{x})=\lambda_n$$

[证毕]

当 B＝I 时，标准特征值问题　$\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$　　（$\mathbf{A}^H=\mathbf{A}$）

$$\begin{cases}\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_n\\ \mathbf{x}_i^H\mathbf{x}_j=\delta_{ij}\end{cases}$$

则　$\min\limits_{(x\neq 0)}\dfrac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}}=\lambda_1\qquad \max\limits_{(x\neq 0)}\dfrac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}}=\lambda_n$

进一步分析可得

$$\min_{x\neq 0}\mathbf{R}(\mathbf{x})\Big|_{a_1=0}=\lambda_2\qquad\qquad \max_{x\neq 0}\mathbf{R}(\mathbf{x})\Big|_{a_n=0}=\lambda_{n-1}$$

$$\vdots\qquad\qquad\qquad\qquad\vdots$$

$$\min_{x\neq 0}\mathbf{R}(\mathbf{x})\Big|_{a_1=a_2=\cdots=a_k=0}=\lambda_{k+1}\qquad \max_{x\neq 0}\mathbf{R}(\mathbf{x})\Big|_{a_n=a_{n-1}=\cdots=a_{n-k}=0}=\lambda_{n-k-1}$$

•**定理 1：**

设 $\mathbf{L}=\operatorname{span}\{\mathbf{x}_r,\mathbf{x}_{r+1},\cdots,\mathbf{x}_s\}$ $(\lambda_r\le\lambda_{r+1}\le\cdots\le\lambda_s)$ ,则:

$$\min_{\substack{x\neq0\\x\in L}} R(x)=\lambda_r \qquad \max_{\substack{x\neq0\\x\in L}} R(x)=\lambda_s$$

•**定理 2：**

设 $\mathbf{V}_k$ 是 $\mathbf{C}^n$ 的一个 k 维子空间，则

$$\max_{V_k\in C^n}\left[\min_{\substack{x\neq0\\x\in V_k}} R(x)\right]=\lambda_{n-k+1} \qquad \min_{V_k\in C^n}\left[\max_{\substack{x\neq0\\x\in V_k}} R(x)\right]=\lambda_k$$

以上两式称为广义特征值的极小极大原理。

B = I 时，标准特征值问题同样存在上述关系；

矩阵奇异值问题：$[\sigma(A)]^2=\lambda(A^HA)$ （非零）

$$R(x)=\frac{x^H(A^HA)x}{x^Hx}=\frac{\|Ax\|_2^2}{\|x\|_2^2}$$

$$\sigma_{n-k+1}=\max_{V_k\in C^n}\left[\min_{\substack{x\neq0\\x\in V_k}}\frac{\|Ax\|_2}{\|x\|_2}\right]$$

$$\sigma_k=\min_{V_k\in C^n}\left[\max_{\substack{x\neq0\\x\in V_k}}\frac{\|Ax\|_2}{\|x\|_2}\right]$$