第十七讲 广义特征值与极小极大原理

一、广义特征值问题

1.定义： 设 $\mathbf{A}$、$\mathbf{B}$ 为 $\mathbf{n}$ 阶方阵，若存在数 λ，使得方程

$$\mathbf{Ax} = \lambda \mathbf{Bx}$$

存在非零解，则称 λ 为 $\mathbf{A}$ 相对于 $\mathbf{B}$ 的广义特征值，$\mathbf{x}$ 为 $\mathbf{A}$ 相对于 $\mathbf{B}$ 的属于广义特征值 λ 的特征向量。

- 广义特征值的求解

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

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

求得 λ 后代回原方程 $\mathbf{Ax} = \lambda \mathbf{Bx}$ 可求出 $\mathbf{x}$

考虑A、B厄米且为正定矩阵的情况

2. 广义特征值问题的等价形式

- B 正定，B^{-1}存在 $\to B^{-1}Ax=\lambda x$，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。
- B 厄米，存在 Cholesky 分解，$B=GG^H$，G 为下三角矩阵

$$Ax=\lambda GG^H x$$

令$G^H x=y$，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，其广义特征值为实数，且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：

$$\mathbf{G}^{-1}\mathbf{A}\left(\mathbf{G}^{H}\right)^{-1}\mathbf{y}_{i}=\lambda\mathbf{y}_{i} \qquad \mathbf{y}_{i}^{H}\mathbf{y}_{j}=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

还原为

$$\mathbf{x}_{i}=\left(\mathbf{G}^{H}\right)^{-1}\mathbf{y}_{i}\quad \mathbf{(i=1,2,\cdots,n)}，\text{则}$$

$$\mathbf{y}_{i}^{H}\mathbf{y}_{j}=\left(\mathbf{x}_{i}^{H}\mathbf{G}\right)\left(\mathbf{G}^{H}\mathbf{x}_{j}\right)=\mathbf{x}_{i}^{H}\mathbf{B}\mathbf{x}_{j}=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

这一组 x_i 称为按 **B** 标准正交化向量系。

按 **B** 标准正交化向量系 $x_1, x_2, \cdots, x_n$ 具有如下性质：

（**1**） $x_i \neq 0$；　　（**2**） $x_1, x_2, \cdots, x_n$ 线性无关。

二、瑞利商和广义特征值的极小极大原理

1.瑞利商的定义： 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 的瑞利商。

设$\lambda_1,\lambda_2,\cdots,\lambda_n$是 A 相对于 B 的广义特征值，且$\lambda_1\le\lambda_2\le\cdots\le\lambda_n$，$x_1,x_2,\cdots,x_n$为对应的按 B 标准正交化向量系，则$x_1,x_2,\cdots,x_n$线性无关。

$\forall x\in C^n$存在$a_1,a_2,\cdots,a_n\in C$，使得 $\mathbf{x}=\sum_{i=1}^{n}\mathbf{a}_i\mathbf{x}_i$，则

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

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

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

2. $\min\limits_{x\neq 0} R(x)=\lambda_1 \qquad \max\limits_{x\neq 0} R(x)=\lambda_n$

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

可取 $k=\dfrac{1}{\|x\|}$， $\quad \|kx\|=1 \quad \therefore R(x)=\left.\dfrac{x^H Ax}{x^H Bx}\right|_{\|x\|=1}$

当 $x=x_1$ 或 $a_i=0\,(i=2,3,\cdots,n)$ 时， $R(x)=\lambda_1$

$$\lambda_i \geq \lambda_1 \qquad R(x)\geq \lambda_1\frac{\sum_{i=1}^{n}|a_i|^2}{\sum_{i=1}^{n}|a_i|^2}=\lambda_1 \qquad \therefore \min_{x\neq 0} R(x)=\lambda_1$$

另一方面，$\lambda_i \le \lambda_n$　　$R(x) \le \lambda_n \dfrac{\sum_{i=1}^{n} |a_i|^2}{\sum_{i=1}^{n} |a_i|^2} = \lambda_n$

$\therefore \quad \max_{x \ne 0} R(x) = \lambda_n$　　[证毕]

当 B＝I 时，化为标准特征值问题　$Ax = \lambda x$　（$A^H = A$）

设 $\begin{cases} \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \\ x_i^H x_j = \delta_{ij} \end{cases}$　则 $\min\limits_{(x \ne 0)} \dfrac{x^H A x}{x^H x} = \lambda_1$　$\max\limits_{(x \ne 0)} \dfrac{x^H A x}{x^H x} = \lambda_n$

进一步分析可得

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

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

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

3. 定理 1. 设 $L=\mathrm{span}\{x_r, x_{r+1}, \cdots, x_s\}$ $(\lambda_r \le \lambda_{r+1} \le \cdots \le \lambda_s)$,则

$$\min_{\substack{x\neq 0\\ x\in L}} R(x)=\lambda_r \qquad \max_{\substack{x\neq 0\\ x\in L}} R(x)=\lambda_s$$

$a_1=0$ 和 $a_n=0$ 的情况均对应于 **x** 在(**n-1**)维的子空间内变动，**x** 在 **L** 中变动是在一个(**s-r+1**)维子空间中变化。

一般的，**x** 在 C^n 的(**n-1**)维子空间 V_{n-1} 中变动时，

$$\min_{\substack{x\neq 0\\ x\in V_{n-1}}} R(x)\le\lambda_2 \qquad \max_{\substack{x\neq 0\\ x\in V_{n-1}}} R(x)\ge\lambda_{n-1}$$

即对于不同的 V_{n-1}，$R(x)$ 的最小值及最大值有可能不同，其中各个最小值中最大者为 λ_2，各个最大值中的最小者为 λ_{n-1}：

$$\max_{V_{n-1}\in C^n}\left[\min_{\substack{x\neq 0\\ x\in V_{n-1}}} R(x)\right]=\lambda_2 \qquad \min_{V_{n-1}\in C^n}\left[\max_{\substack{x\neq 0\\ x\in V_{n-1}}} R(x)\right]=\lambda_{n-1}$$

定理 2. 设 V_k 是 C^n 的一个 k 维子空间，则

$$\max_{V_k \in C^n}\left[\min_{\substack{x\neq 0\\ x\in V_k}} R(x)\right]=\lambda_{n-k+1} \qquad \min_{V_k \in C^n}\left[\max_{\substack{x\neq 0\\ x\in V_k}} R(x)\right]=\lambda_{k}$$

说明：(1)B＝I 时，标准特征值问题同样存在上述关系。

(2)矩阵奇异值问题：

$$[\sigma(A)]^2=\lambda(A^H A) \qquad （非零）$$

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

$$\sigma_{n-k+1}=\max_{V_k \in C^n}\left[\min_{\substack{x\neq 0\\ x\in V_k}} \frac{\|Ax\|_2}{\|x\|_2}\right]$$

$$\sigma_{k}=\min_{V_k \in C^n}\left[\max_{\substack{x\neq 0\\ x\in V_k}} \frac{\|Ax\|_2}{\|x\|_2}\right]$$