矩 阵 论

主讲教师：徐乐

2015年1月13日星期二

上讲回顾

- 第22讲 矩阵特征值估计
  - 特征值界的估计
  - 盖尔圆法

特征值界的估计

❖矩阵特征值估计

- 特征值计算较困难，希望找到简便的特征值界限或分布范围的估计方法

❖定理1

- 设$A \in R^{n\times n}$， λ 为A的任意特征值，则有

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

• 其中

$$M = \max_{1\le i,j\le n}\left|\frac{a_{ij}-a_{ji}}{2}\right|$$

❖定理2

- 设$A\in R^{n\times n}$， λ 为A的任意特征值，则有

$$|\lambda| \le n\rho \qquad |\mathrm{Re}(\lambda)| \le \frac{1}{2}n\tau \qquad |\mathrm{Im}(\lambda)| \le \frac{1}{2}ns$$

- 其中

$$\rho = \max_{1\le i,j\le n} |a_{ij}| \qquad \tau = \max_{1\le i,j\le n} |a_{ij} + \overline{a_{ji}}| \qquad s = \max_{1\le i,j\le n} |a_{ij} - \overline{a_{ji}}|$$

盖尔圆法

❖定义

- 设$A=(a_{i,j})_{n\times n}\in R^{n\times n}$，由方程所确定的圆称为A的第$i$个盖尔圆，$R_i$称为盖尔圆的半径

$$\left|z-a_{ii}\right|\le R_i=\sum_{\substack{j=1\\ i\ne j}}^{n}\left|a_{ij}\right|$$

❖定理3

- 矩阵A的所有特征值均落在它的所有盖尔圆的并集之中

盖尔圆法

❖定理4

- 将矩阵A的全体盖尔圆的并集按连通部分分成若干个子集
  - 一个子集由完全连通的盖尔圆组成，不同子集没有相连通的部分
  - 对每个子集，若它恰好由K个盖尔圆组成，则该子集中恰好包含A的K个特征值
- 说明
  - 盖尔圆相互重叠时重复计算，特征值相重时也重复计算

盖尔圆法

❖推论

- 孤立盖尔圆中恰好包含一个特征值
- 实矩阵的孤立盖尔圆恰好包含一个实特征值
- 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)
- 盖尔圆半径变为R_i，两个盖尔圆定理仍然成立

$$\boxed{R_i = \sum_{j=1}^{n}{}' \frac{\alpha_i}{\alpha_j} |a_{ij}|}$$

❖说明

- 相似矩阵$P^{-1}AP$与A具有相同的特征值，取

$$\boxed{P=\mathrm{diag}\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix},(\alpha_i>0)}$$

$$B=P^{-1}AP=\begin{bmatrix}\frac{1}{\alpha_1} & & & 0\\ & \frac{1}{\alpha_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\alpha_n}\end{bmatrix}\begin{bmatrix}a_{ij}\end{bmatrix}\begin{bmatrix}\alpha_1 & & & 0\\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_n\end{bmatrix}$$

$$\boxed{=\begin{bmatrix}\frac{\alpha_i}{\alpha_j}a_{ij}\end{bmatrix}}$$

- 广义特征值问题
- 瑞利商
- 极小极大原理

广义特征值问题

❖定义

- 设A、B为n阶方阵，若存在数λ，使得方程 $Ax=\lambda Bx$存在非零解
  - 则称λ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λ的特征向量
    - 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

广义特征值问题

- 广义特征值的求解

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

  - 特征方程　$\det(\mathbf{A}-\lambda\mathbf{B})=\mathbf{0}$
  - 求得 λ 后代回原方程**Ax= λ Bx**可求出**x**

- 本课程进一步考虑**A**、**B**厄米且为正定矩阵的情况

广义特征值问题

❖等价表述

- $\mathbf{B}$正定，$\mathbf{B}^{-1}$存在

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

  - 广义特征值问题化为了标准特征值问题
  - 但一般来说，一般$\mathbf{B}^{-1}\mathbf{A}$不再是厄米矩阵

- $\mathbf{B}$厄米，存在Cholesky分解，$\mathbf{B}=\mathbf{G}\mathbf{G}^{H}$，$\mathbf{G}$满秩

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{G}\mathbf{G}^{H}\mathbf{x} \qquad \mathbf{G}^{H}\mathbf{x} = \mathbf{y}$$

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

广义特征值问题

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

- $G^{-1}A\left(G^H\right)^{-1}$ 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满足

$$G^{-1}A\left(G^H\right)^{-1}y_i=\lambda y_i \qquad y_i^H y_j=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\ne j\end{cases}$$

- 还原为 $x_i=\left(G^H\right)^{-1}y_i$

$$y_i^H y_j=\left(x_i^H G\right)\left(G^H x_j\right)=x_i^H B x_j=\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\ne j\end{cases}$$

定义

- **A、B为n阶厄米矩阵，且B正定，称A相对于B的瑞利商为**

$$R(x)=\frac{x^H A x}{x^H B x}\left(x\neq 0\right)$$

  - 说明：若$x_1,x_2,\dots,x_n$为线性无关向量组，则

$\forall x\in C^n$　　$\exists a_1,a_2\cdots,a_n\in C$　　$x=\sum_{i=1}^{n}a_i x_i$

$$x^H B x=\left(\sum_{i=1}^{n}a_i x_i\right)^H B\left(\sum_{j=1}^{n}a_j x_j\right)=\sum_{i,j=1}^{n}\overline{a_i}a_j x_i^H B x_j=\sum_{i=1}^{n}\left|a_i\right|^2$$

$$x^H A x=\sum_{i,j=1}^{n}\overline{a_i}a_j x_i^H A x_j=\sum_{i,j=1}^{n}\overline{a_i}a_j x_i^H \lambda_j B x_j=\sum_{i=1}^{n}\lambda_i\left|a_i\right|^2$$

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

瑞利商

❖瑞利商性质

$$\boxed{\min_{x\neq 0} R(x)=\lambda_1} \qquad \boxed{\max_{x\neq 0} R(x)=\lambda_n}$$

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

• 取 $k=\dfrac{1}{\|x\|}$ $\qquad \|kx\|=1 \qquad R(x)=\left.\dfrac{x^H Ax}{x^H Bx}\right|_{\|x\|=1}$

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

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

• 另一方面

$\lambda_i \leq \lambda_n \qquad R(x)\leq \lambda_n \dfrac{\sum_{i=1}^{n}|a_i|^2}{\sum_{i=1}^{n}|a_i|^2}=\lambda_n \qquad \boxed{\max_{x\neq 0} R(x)=\lambda_n}$

瑞利商

❖ 当B＝I时，标准特征值问题 $Ax=\lambda x \quad A^H=A$

$$\begin{cases}\lambda_1\le\lambda_2\le\cdots\le\lambda_n\\ x_i^H x_j=\delta_{ij}\end{cases}$$

$$\min_{(x\neq 0)}\frac{x^H A x}{x^H x}=\lambda_1$$

$$\boxed{\min_{x\neq 0} R(x)\Big|_{a_1=0}=\lambda_2}$$

$$\vdots$$

$$\min_{x\neq 0} R(x)\Big|_{a_1=a_2=\cdots=a_k=0}=\lambda_{k+1}$$

$$\max_{(x\neq 0)}\frac{x^H A x}{x^H x}=\lambda_n$$

$$\boxed{\max_{x\neq 0} R(x)\Big|_{a_n=0}=\lambda_{n-1}}$$

$$\vdots$$

$$\max_{x\neq 0} R(x)\Big|_{a_n=a_{n-1}=\cdots=a_{n-k}=0}=\lambda_{n-k-1}$$

定理1

- 设 $L=\text{span}\{x_r,x_{r+1},\cdots,x_s\}\quad(\lambda_r\le\lambda_{r+1}\le\cdots\le\lambda_s)$
- 则 $\min\limits_{\substack{x\neq0\\x\in L}}R(x)=\lambda_r\qquad\max\limits_{\substack{x\neq0\\x\in L}}R(x)=\lambda_s$
  - 这一结果不便于应用，希望对上述结果进行改造，改造成不依赖于x_i的一种表达方式
  - $a_1=0,a_n=0$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

$$\min_{\substack{x\neq0\\x\in V_{n-1}}}R(x)\le\lambda_2\qquad\max_{\substack{x\neq0\\x\in V_{n-1}}}R(x)\ge\lambda_{n-1}$$

极小极大原理

- 对于不同的V_{n-1}，R(x)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

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

    - 以上两式称为广义特征值的极小极大原理

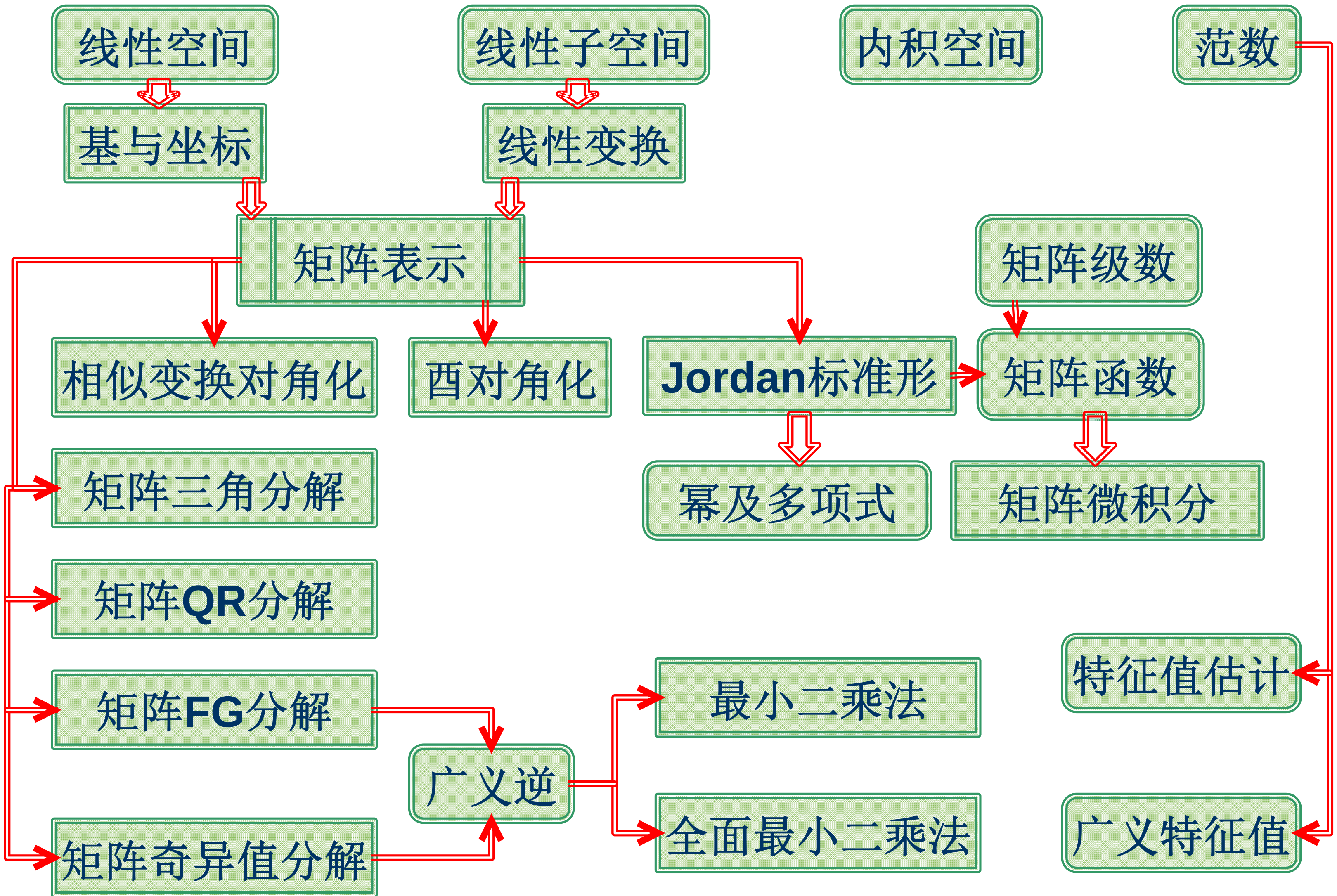
❖讨论

- 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 $\left[\sigma(A)\right]^2=\lambda\left(A^H A\right)$ 非零

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

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

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


线性空间
基与坐标
线性子空间
线性变换
内积空间
范数
矩阵表示
矩阵级数
相似变换对角化
酉对角化
Jordan标准形
矩阵函数
矩阵三角分解
幂及多项式
矩阵微积分
矩阵QR分解
特征值估计
矩阵FG分解
最小二乘法
广义逆
全面最小二乘法
矩阵奇异值分解
广义特征值

祝大家考试取得好成绩

态度端正、复习认真

做题仔细、步骤详细