数值线代の题库三

by 23 大数据 miracle

1. 证明 Gershgorin 圆盘定理: 设 $A \in \mathbb{C}^{n\times n}$, $A = (a_{ij})$, 令

$$G_i(A) = \{z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|\}$$

其中 $i = 1, 2, \ldots, n$, 则有

$$\lambda(A) \subset G_1(A) \cup G_2(A) \cup \ldots \cup G_n(A)$$

Solution: 设 λ 是 A 的特征值, $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{C}^n \setminus \{0\}$ 是对应的特征向量, 满足 $Ax = \lambda x$. 取下标 r 使得

$$|x_r| = \max_{1\leq k\leq n} \{|x_k|\}$$

因 $x \neq 0$, 有 $|x_r| > 0$.

考虑 $Ax = \lambda x$ 展开后的第 r 个分量:

$$\sum_{j=1}^{n} a_{rj} x_j = \lambda x_r$$

移项得

$$(\lambda - a_{rr}) x_r = \sum_{j\neq r} a_{rj} x_j$$

取绝对值并应用三角不等式:

$$|\lambda - a_{rr}| \cdot |x_r| = \left| \sum_{j\neq r} a_{rj} x_j \right| \leq \sum_{j\neq r} |a_{rj}| \cdot |x_j|$$

又 $|x_j| \leq |x_r|$ 对所有 j 成立, 故

$$\sum_{j\neq r} |a_{rj}| \cdot |x_j| \leq \sum_{j\neq r} |a_{rj}| \cdot |x_r| = |x_r| \sum_{j\neq r} |a_{rj}|$$

代入原不等式, 得

$$|\lambda - a_{rr}| \cdot |x_r| \leq |x_r| \sum_{j\neq r} |a_{rj}|$$

两边除以 $|x_r| > 0$:

$$|\lambda - a_{rr}| \leq \sum_{j\neq r} |a_{rj}|$$

因此, $\lambda \in G_r \subset G_1(A) \cup G_2(A) \cup \ldots \cup G_n(A)$, 由于 λ 的任意性, 命题得证.

另解: 反证法. 假设存在 A 的特征值 λ^* 满足 $\lambda^* \notin \bigcup_{i=1}^n G_i(A)$, 即

$$|\lambda^* - a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad \forall i = 1, 2, \ldots, n.$$

考虑矩阵 $B = \lambda^* I - A$, 则一定有 $\det(B) = 0$, 且又由假设得对于任意的 i 都有

$$|b_{ii}| = |\lambda^* - a_{ii}| > \sum_{j \neq i} |a_{ij}| = \sum_{j \neq i} |b_{ij}|$$

故 B 严格对角占优. 而严格对角占优矩阵非奇异, 则 $\det(B) \neq 0$, 矛盾. 因此假设不成立, 原命题得证.

2. 设 A 是 $n \times m$ 矩阵, B 是 $m \times n$ 矩阵, 且 $m \geq n$. 证明:

$$\lambda(BA) = \lambda(AB) \cup \{\underbrace{0, \ldots, 0}_{m-n\text{ 个}}\}$$

Solution: 直接应用降阶公式（按两种方式化为分块三角阵）:

$$\det \begin{bmatrix} \lambda I_n & -A \\ -B & I_m \end{bmatrix} = \det(\lambda I_n - AB) = \lambda^{n-m} \det(\lambda I_m - BA)$$

由此得到特征多项式关系:

$$\det(\lambda I_m - BA) = \lambda^{m-n} \det(\lambda I_n - AB)$$

因此, 直接有特征值集合满足:

$$\lambda(BA) = \lambda(AB) \cup \{\underbrace{0, \ldots, 0}_{m-n\text{个}}\}$$

3. 设 $A = \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}$, $\alpha \neq \beta$, 求 A 的特征值 α 和 β 的条件数.

Solution: 由单特征值的条件数定义, 即:

$$\operatorname{cond}(\lambda) = \|y\|_2$$

其中 y 满足 $y^{\mathrm{T}} A = \lambda y^{\mathrm{T}}$ 且 $y^{\mathrm{T}} x = 1$, x 为 A 对应特征值 λ 的单位特征向量, 可知

1. 对于特征值 α:

$$A-\alpha I=\begin{bmatrix}0 & \gamma \\ 0 & \beta-\alpha\end{bmatrix}$$

自然有单位右特征向量:

$$x_1=\begin{bmatrix}1 \\ 0\end{bmatrix}$$

又有:

$$A^{\mathrm{T}}-\alpha I=\begin{bmatrix}0 & 0 \\ \gamma & \beta-\alpha\end{bmatrix}$$

取满足与 x_1 内积为 1 的左特征向量 y_1, 满足:

$$y_1=\begin{bmatrix}1 \\ \frac{\gamma}{\alpha-\beta}\end{bmatrix}$$

则 α 的条件数:

$$\operatorname{cond}(\alpha)=\|y_1\|_2=\frac{\sqrt{|\alpha-\beta|^2+|\gamma|^2}}{|\alpha-\beta|}$$

2. 对于特征值 β:

$$A-\beta I=\begin{bmatrix}\alpha-\beta & \gamma \\ 0 & 0\end{bmatrix}$$

自然有单位右特征向量:

$$x_2=\frac{1}{\sqrt{|\gamma|^2+|\beta-\alpha|^2}}\begin{bmatrix}\gamma \\ \beta-\alpha\end{bmatrix}$$

又有:

$$A^{\mathrm{T}}-\beta I=\begin{bmatrix}\alpha-\beta & 0 \\ \gamma & 0\end{bmatrix}$$

取满足与 x_2 内积为 1 的左特征向量 y_2, 满足:

$$y_2=\frac{1}{\beta-\alpha}\begin{bmatrix}0 \\ \sqrt{|\gamma|^2+|\beta-\alpha|^2}\end{bmatrix}$$

则 β 的条件数:

$$\operatorname{cond}(\beta)=\|y\|_2=\frac{\sqrt{|\beta-\alpha|^2+|\gamma|^2}}{|\beta-\alpha|}$$

4. 证明特征值和特征向量的条件数在酉相似条件下保持不变.

Solution: 单位特征向量的条件数定义（单特征值的条件数定义见上题）, 即:

$$\mathrm{cond}(x) = \|\Sigma^{\perp}\|_2 = \|U_2(\lambda I - A_2)^{-1}U_2^*\|_2$$

其中, $U^*AU = \begin{bmatrix} \lambda & * \\ 0 & A_2 \end{bmatrix}$, $U = \begin{bmatrix} x & U_2 \end{bmatrix}$. 设 A 的右、左特征向量分别是 x 和 y^{T}, 再令 $B = W^*AW$, W 满足 $W^*W = I$, 于是有

1. 特征值条件数的不变性:

   B 的特征值仍为 λ. 令 $x_B = W^*x$, 则 $Bx_B = \lambda x_B$ 且 $\|x_B\|_2 = 1$. 令 $y_B = W^{\mathrm{T}}y$, 则

   $$y_B^{\mathrm{T}}B = y_B^{\mathrm{T}}(W^*AW) = y^{\mathrm{T}}AW = \lambda y^{\mathrm{T}}W = \lambda y_B^{\mathrm{T}}$$

   且

   $$y_B^{\mathrm{T}}x_B = (W^{\mathrm{T}}y)^{\mathrm{T}}(W^*x) = y^{\mathrm{T}}x = 1$$

   因此,

   $$\mathrm{cond}_B(\lambda) = \|y_B\|_2 = \|W^{\mathrm{T}}y\|_2 = \|y\|_2 = \mathrm{cond}_A(\lambda)$$

2. 特征向量条件数的不变性:

   令 $U_B = W^*U = \begin{bmatrix} W^*x & W^*U_2 \end{bmatrix} = \begin{bmatrix} x_B & U_{2,B} \end{bmatrix}$, 其中 $U_{2,B} = W^*U_2$. 则

   $$U_B^*BU_B = (W^*U)^*(W^*AW)(W^*U) = U^*AU = \begin{bmatrix} \lambda & * \\ 0 & A_2 \end{bmatrix}$$

   故 $B_2 = A_2$. 则特征向量条件数为

   $$\mathrm{cond}_B(x_B) = \|U_{2,B}(\lambda I - B_2)^{-1}U_{2,B}^*\|_2 = \|W^*U_2(\lambda I - A_2)^{-1}U_2^*W\|_2$$

   由于谱范数在酉变换下不变, 有

   $$\mathrm{cond}_B(x_B) = \|U_2(\lambda I - A_2)^{-1}U_2^*\|_2 = \mathrm{cond}_A(x)$$

综上, $\mathrm{cond}(\lambda)$ 和 $\mathrm{cond}(x)$ 在酉相似变换下均保持不变.

5. 在幂法中, 取 $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $u_0 = (0, 0, 1)^{\mathrm{T}}$. 得到一个精确到 5 位有效数字的特征向量需要多少次迭代?

Solution: 注意到有分解 $A = I + J_3(0)$, 其中

$$J_3(0) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

而又

$$(J_3(0))^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

且 $(J_3(0))^k = 0, \forall k \geq 3.$

由二项式定理（又单位矩阵和任意矩阵都可交换）有

$$A^n = (I + J_3(0))^n = \binom{n}{0} I + \binom{n}{1} J_3(0) + \binom{n}{2} (J_3(0))^2$$

代入得

$$A^n = \begin{bmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$$

于是有

$$v_n = A^n u_0 = \begin{bmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{n(n-1)}{2} \\ n \\ 1 \end{bmatrix}$$

又 $\|v_n\|_\infty = \frac{n(n-1)}{2}, \forall n \geq 3$, 根据幂法迭代公式,

$$u_n = \frac{v_n}{\|v_n\|_\infty} = \begin{bmatrix} 1 \\ \frac{2}{n-1} \\ \frac{2}{n(n-1)} \end{bmatrix}$$

事实上, 由于这种情况下可以收敛到 A 的精确特征向量为 $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$,

则精确到 5 位有效数字需满足:

$$\left| \frac{2}{n-1} \right| \leq 10^{-5} \implies n - 1 \geq 2 \times 10^5 \implies n \geq 200001$$

故需 $\boxed{200001}$ 次迭代.

6. 设 $A \in \mathbb{C}^{n \times n}$ 有实特征值并满足 $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{n-1} > \lambda_n$. 现应用幂法于矩阵 $A - \mu I$. 试证: 选择 $\mu = \frac{1}{2}(\lambda_2 + \lambda_n)$ 时, 所产生的向量序列收敛到属于 λ_1 的特征向量的速度最快.

Solution: **证明:** 由于 A 特征值满足 $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{n-1} > \lambda_n$, 不论 μ 取何值, $A - \mu I$ 按模最大的特征值为 $\lambda_1 - \mu$ 或 $\lambda_n - \mu$, 因本题要求收敛到 λ_1, 所以首先选择 μ 使得

$$|\lambda_1 - \mu| > |\lambda_n - \mu|$$

由此可知

$$\mu < \frac{1}{2}(\lambda_1 + \lambda_n)$$

定义

$$f(\mu) := \frac{\max\limits_{2\le i\le n} |\lambda_i - \mu|}{|\lambda_1 - \mu|}$$

则 $f(\mu)$ 表征了最大模特征值与次最大模特征值的分离程度, $f(\mu)$ 越小, 幂法收敛速度越快, 所以有:

1. 当 $|\lambda_2 - \mu| \ge |\lambda_n - \mu|$, 即 $\mu \le \frac{1}{2}(\lambda_2 + \lambda_n)$ 时:

$$f(\mu) = \frac{|\lambda_2 - \mu|}{|\lambda_1 - \mu|} = \frac{\lambda_2 - \mu}{\lambda_1 - \mu} = 1 - \frac{\lambda_1 - \lambda_2}{\lambda_1 - \mu}$$

   函数单调递减, 在 $\mu = \frac{1}{2}(\lambda_2 + \lambda_n)$ 处取最小值

2. 当 $|\lambda_2 - \mu| < |\lambda_n - \mu|$, 即 $\mu > \frac{1}{2}(\lambda_2 + \lambda_n)$ 时:

$$f(\mu) = \frac{|\lambda_n - \mu|}{|\lambda_1 - \mu|} = \frac{\mu - \lambda_n}{\lambda_1 - \mu} = \frac{\lambda_1 - \lambda_n}{\lambda_1 - \mu} - 1$$

   函数单调递增, 此时 $f(\mu) > f(\frac{\lambda_2 + \lambda_n}{2})$ 恒成立

综上, 当 $\mu = \frac{1}{2}(\lambda_2 + \lambda_n)$ 时, $f(\mu)$ 为最小值, 此时收敛速度最快, 命题得证.

7. 设 $A \in \mathbb{C}^{n\times n}$, $x \in \mathbb{C}^n$, $X = [x, Ax, \cdots, A^{n-1}x]$. 证明: 如果 X 是非奇异的, 则 $X^{-1}AX$ 是上 Hessenberg 矩阵.

Solution: 因为 $X = [x, Ax, \cdots, A^{n-1}x]$ 非奇异, 所以 $\{x, Ax, \cdots, A^{n-1}x\}$ 线性无关.

设 $H = X^{-1}AX = [h_{ij}] \in \mathbb{C}^{n\times n}$, 则 $AX = XH$. 考虑 $AX = XH$ 的第 k 列, 由矩阵乘法得:

$$A \cdot (A^{k-1}x) = A^k x = Xh_k = \sum_{i=1}^{n} h_{ik} A^{i-1}x, \quad 1 \le k \le n.$$

当 $1 \le k \le n-1$ 时, 移项后:

$$\sum_{\substack{i=1 \\ i\ne k+1}}^{n} h_{ik} A^{i-1}x + (h_{k+1,k} - 1)A^k x = 0$$

由于 $\{x, Ax, \cdots, A^{n-1}x\}$ 线性无关, 各项系数必须为零:

$$h_{ik} = 0, 1 \le i \le n, i \ne k+1$$

$$h_{k+1,k} = 1, 1 \le k \le n-1$$

因此,

$$H=\begin{pmatrix}0 & 0 & 0 & \cdots & h_{1n}\\ 1 & 0 & 0 & \cdots & h_{2n}\\ 0 & 1 & 0 & \cdots & h_{3n}\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \cdots & 0 & 1 & h_{nn}\end{pmatrix}$$

所以, $X^{-1}AX=H$ 是上 Hessenberg 矩阵.

8. 设 H 是一个不可约的上 Hessenberg 矩阵. 证明: 存在一个对角阵 D, 使得 $D^{-1}HD$ 的次对角元均为 1, 此时 $\kappa_2(D)=\|D\|_2\|D^{-1}\|_2$ 是多少?

Solution: 设 H 的次对角元为 $a_1,\dots,a_{n-1}$. 由于 H 是不可约的上 Hessenberg 矩阵, 所以 $a_i\neq 0, \forall 1\le i\le n-1$. 取对角矩阵 $D=\operatorname{diag}(d_1,d_2,\dots,d_n)$, $d_i\neq 0, \forall 1\le i\le n$, 则

$$D^{-1}=\operatorname{diag}\left(\frac{1}{d_1},\frac{1}{d_2},\dots,\frac{1}{d_n}\right)$$

下考虑在矩阵 $D^{-1}HD$ 中的元素:
在本题中, 由于左乘对角阵 D^{-1} 相当于让矩阵 H 的对应行向量分别乘 $\dfrac{1}{d_i}$, 右乘对角阵 D 相当于让矩阵 H 的对应列向量分别乘 d_j , 于是有:

$$(D^{-1}HD)_{ij}=\frac{d_j}{d_i}h_{ij}$$

要求次对角元（即 $i=j+1$ 的元素）为 1, 即 $\forall 1\le j\le n-1$

$$\frac{d_j}{d_{j+1}}a_j=1 \quad\Longrightarrow\quad d_{j+1}=a_jd_j \quad (1\le j\le n-1)$$

由此得到一个欠定方程组:

$$Ad=0, A\in\mathbb{C}^{(n-1)\times n}, d\in\mathbb{C}^n, d=(d_1,\dots,d_n)^{\mathrm{T}}$$

$$A=\begin{bmatrix}a_1 & -1 & & & \\ & a_2 & -1 & & \\ & & \ddots & \ddots & \\ & & & a_{n-1} & -1\end{bmatrix}$$

对这个方程, 给定 $d_1\neq 0$ 后, 解得

$$d_2=a_1d_1,\quad d_3=a_1a_2d_1,\quad \cdots,\quad d_n=a_1\cdots a_{n-1}d_1$$

所以得到所求对角阵

$$D=\operatorname{diag}\left(d_1,\ a_1d_1,\ \cdots,\ \prod_{i=1}^{n-1}a_id_1\right)$$

故所求矩阵 D 一定存在, 再求条件数:

由矩阵 2-范数的性质知, 对于任一矩阵 $A \in \mathbb{C}^{n\times n}$

$$||A||_2 = \sqrt{\rho(A^{\mathrm{T}}A)}$$

所以有

$$||D||_2 = \sqrt{\rho(D^{\mathrm{T}}D)} = \sqrt{\rho(D^2)} = \max_{1\le i\le n} |d_i|$$

$$||D^{-1}||_2 = \sqrt{\rho((D^{-1})^{\mathrm{T}}D^{-1})} = \sqrt{\rho((D^{-1})^2)} = \max_{1\le j\le n} |d_j^{-1}| = \frac{1}{\min_{1\le j\le n} |d_j|}$$

所以

$$\kappa_2(D) = ||D||_2||D^{-1}||_2 = \frac{\max_{1\le i\le n}|d_i|}{\min_{1\le j\le n}|d_j|} = \frac{\max_{1\le i\le n}\prod_{k=1}^{i}|a_{k-1}|}{\min_{1\le j\le n}\prod_{l=1}^{j}|a_{l-1}|}$$

其中, $a_0 = 1$.

9. 设 H 是一个奇异不可约上 Hessenberg 矩阵. 证明: 进行一次基本的 QR 迭代后, H 的零特征值将出现.

Solution: 由奇异性和不可约性知 $\mathrm{rank}(H) = n-1$, 且其下次对角线元素全非零.

对 H 进行 QR 分解. 由于 H 是上 Hessenberg 矩阵, 可通过 $n-1$ 次 Givens 变换实现:

$$H = QR, \quad Q \in \mathbb{R}^{n\times n} \text{ 正交}, \quad R \in \mathbb{R}^{n\times n} \text{ 上三角}.$$

因 $\mathrm{rank}(H) = n-1$ 且 Givens 变换保秩, 所以 $\mathrm{rank}(R) = n-1$. 结合 H 不可约的性质:

- R 的前 $n-1$ 个对角元非零（不可约性保证在 Givens 变换过程中变化后的 4 个元素里左上角元素一定非 0）.
- R 的最后一行全零（因为是秩为 $n-1$ 的上三角矩阵）.

故 R 可写为分块形式:

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \quad R_1 \in \mathbb{R}^{(n-1)\times n} \text{ 满秩上三角}$$

QR 迭代后的矩阵为:

$$\tilde{H} = RQ = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} Q = \begin{pmatrix} R_1 Q \\ 0 \end{pmatrix}$$

上式表明: $\tilde{H}$ 的最后一行全零, 自然有零特征值出现, 命题得证.