4. 对称矩阵特征值的计算

4.1 Rayleigh 商迭代

$A \in \mathbb{C}^{n\times n}$, $x \in \mathbb{C}^n$, $x \neq 0$

$$R(x) = \frac{(Ax, x)}{(x, x)}$$

称为关于 A 和 x 的 Rayleigh 商

Rayleigh 商迭代算法

$$\boxed{\begin{aligned} & v^{(0)} \in \mathbb{R}^n, \quad \|v^{(0)}\|_2 = 1 \\ & k = 0, 1, 2 \cdots \\ & \quad \mu_k = R(v^{(k)}) \\ & \quad (A - \mu_k I) y^{(k+1)} = v^{(k)} \\ & \quad v^{(k+1)} = y^{(k+1)} / \|y^{(k+1)}\|_2 \end{aligned}}$$

4.2. Jacobi 方法

$A = [a_{ij}] \in \mathbb{R}^{n\times n}$, $A^T = A$, 记

$$\mathrm{off}(A) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|^2$$

$J = J(p, q, \theta)$ 为旋转角度为 θ 的 Givens 变换矩阵

$B = JAJ^{-1}$, $c = \cos\theta$, $s = \sin\theta$

$$J = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad J^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

B的元素为

$$b_{pp} = a_{pp}c^2 + a_{qq}s^2 + a_{pq}\sin 2\theta$$
$$b_{qq} = a_{pp}s^2 + a_{qq}c^2 - a_{pq}\sin 2\theta$$
$$b_{pq} = b_{qp} = \frac{1}{2}(a_{qq} - a_{pp})\sin 2\theta + a_{pq}\cos 2\theta$$

$$b_{ip} = b_{pi} = a_{ip}c + a_{iq}s \quad (i \neq p, q)$$
$$b_{iq} = b_{qi} = -a_{ip}s + a_{iq}c \quad (i \neq p, q)$$
$$b_{ij} = a_{ij} \quad (i \neq p, q,\ j \neq p, q)$$

容易验证

$$b_{pp}^2 + b_{qq}^2 + 2b_{pq}^2 = a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2$$

则

$$\mathrm{off}(B) = \|B\|_F^2 - \sum_{i=1}^{n} b_{ii}^2 = \|A\|_F^2 - \sum_{\substack{i=1 \\ i\neq p,q}}^{n} a_{ii}^2 - (b_{pp}^2 + b_{qq}^2)$$

$$= \mathrm{off}(A) - 2a_{pq}^2 + 2b_{pq}^2$$

选择 θ 使得 $b_{pq} = 0$ 此时 $\mathrm{off}(B)$ 达到极小

θ 满足

$$\cot 2\theta = \frac{a_{pp} - a_{qq}}{2a_{pq}} = \tau$$

记 $t = \tan\theta$, t 满足

$$t^2 + 2t\tau - 1 = 0 \qquad (6.4.1)$$

取绝对值较小的根（对应 $|\theta|\le\frac{\pi}{4}$）

$$t=\frac{\mathrm{sgn}(\tau)}{|\tau|+\sqrt{1+\tau^2}} \qquad (6.4.2)$$

$$\Rightarrow \quad c=\frac{1}{\sqrt{1+t^2}}, \quad s=tc \qquad (6.4.3)$$

在经典 Jacobi 方法中 取

$$|a_{pq}|=\max_{1\le i<j\le n}|a_{ij}| \qquad (6.4.4)$$

其迭代格式为

$$A_k=J_k^T A_{k-1} J_k$$

J_k 为 (6.4.1) − (6.4.4) 确定的 Givens 旋转矩阵

引理 6.4.1 设 A, E 均为 $\mathbb{R}^{n\times n}$ 中的对称矩阵，

$A, E, A+E$ 的特征值分别为

$$\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_n$$

$$\nu_1\ge\nu_2\ge\cdots\ge\nu_n$$

$$\mu_1\ge\mu_2\ge\cdots\ge\mu_n$$

则有

$$\lambda_i+\nu_n\le\mu_i\le\lambda_i+\nu_1, \quad i=1,2,\cdots,n$$

证明： 记 $x_1\cdots x_n$ 为 A 的特征向量

$W_i=\mathrm{span}\{x_1,\cdots x_i\}$ 则有

$$\mu_i \geq \min_{\substack{x \in W_i \\ x \neq 0}} \frac{((A+E)x, x)}{(x, x)} \geq \min_{\substack{x \in W_i \\ x \neq 0}} \frac{(Ax, x)}{(x, x)} + \min_{\substack{x \in W_i \\ x \neq 0}} \frac{(Ex, x)}{(x, x)}$$

$$\geq \lambda_i + \min_{\substack{x \in W_i \\ x \neq 0}} \frac{(Ex, x)}{(x, x)} \geq \lambda_i + \nu_n$$

同理可证 $\mu_i \leq \lambda_i + \nu_1$

定理 6.4.1 存在 A 的特征值的一个排列 $\lambda_1 \cdots \lambda_n$

使得

$$\lim_{k \to \infty} A_k = \operatorname{diag}(\lambda_1 \cdots \lambda_n)$$

证明: $\operatorname{off}(A_{k+1}) = \operatorname{off}(A_k) - 2\left(a_{pq}^{(k)}\right)^2$

注意到

$$|a_{pq}^{(k)}| = \max_{1 \leq i < j \leq n} |a_{ij}^{(k)}|$$

$$\Rightarrow \quad 2\left(a_{pq}^{(k)}\right)^2 \geq \frac{2}{n(n-1)} \operatorname{off}(A_k)$$

$$\Rightarrow \quad \operatorname{off}(A_{k+1}) \leq \left(1 - \frac{1}{N}\right) \operatorname{off}(A_k), \quad N = \frac{n(n-1)}{2}$$

$$\Rightarrow \quad \lim_{k \to \infty} \operatorname{off}(A_k) = 0$$

下面证明存在特征值的一个排列 $\lambda_1 \cdots \lambda_n$

使得 $\lim\limits_{k \to \infty} a_{ii}^{(k)} = \lambda_i, \quad i = 1, \cdots, n$

令 $E(A) = \left(\operatorname{off}(A)\right)^{1/2}$

记 A 的互不相同的特征值之间的最小距离为 δ. 即

$$\delta = \min\{|\mu-\lambda| : \lambda, \mu \in \sigma(A), \lambda \neq \mu\}$$

$\forall \varepsilon < \frac{\delta}{4}$ $\exists k_0$ 使得 $\forall k \geq k_0$ 有

$$E(A_k) < \varepsilon < \frac{\delta}{4}$$

由引理 6.4.1，存在一个排列 $\lambda_1 \cdots \lambda_n$ 使得

$$|\lambda_i - a_{ii}^{(k_0)}| \leq E(A_k) < \varepsilon < \frac{\delta}{4}, \quad i=1,\cdots,n$$

下面证 $\forall k > k_0$ 有

$$|\lambda_i - a_{ii}^{(k)}| < \varepsilon$$

由归纳法只需证

$$|\lambda_i - a_{ii}^{(k_0+1)}| < \varepsilon$$

$$\begin{aligned}
a_{pp}^{(k_0+1)} &= a_{pp}^{(k_0)} + c^2\left(-2t\, a_{pq}^{(k_0)} + t^2\left(a_{qq}^{(k_0)} - a_{pp}^{(k_0)}\right)\right) \\
&= a_{pp}^{(k_0)} + c^2\left(-2t\, a_{pq}^{(k_0)} + t(1-t^2)\, a_{pq}^{(k_0)}\right) \\
&= a_{pp}^{(k_0)} - t\, a_{pq}^{(k_0)}
\end{aligned}$$

同理

$$a_{qq}^{(k_0+1)} = a_{qq}^{(k_0)} + t\, a_{pq}^{(k_0)}$$

$\forall \lambda_j \neq \lambda_p$ 有

$$\begin{aligned}
|a_{pp}^{(k_0+1)} - \lambda_j| &\geq |\lambda_p - \lambda_j| - |a_{pp}^{(k_0)} - \lambda_p| - |t\, a_{pq}^{(k_0)}| \\
&\underset{|t|\leq 1}{\geq} \delta - 2\varepsilon \geq 2\varepsilon
\end{aligned}$$

由此可得 $|a_{pp}^{(k_0+1)} - \lambda_p| < \varepsilon$

同理可证 $|a_{qq}^{(k_0+1)} - \lambda_q| < \varepsilon$

4.3 二分法

$A \in \mathbb{R}^{n\times n}$ 对称，则存在正交矩阵 Q 使得

$$QAQ^T = T$$

其中 $T \in \mathbb{R}^{n\times n}$ 为对称三对角矩阵.

设

$$T = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_n \\ & & & \beta_n & \alpha_n \end{bmatrix}$$

下面考虑 T 的特征值的计算.

不失一般性 设 $\beta_i \neq 0$，$i=2,\cdots,n$

记 $P_i(\lambda)$ 为 $T-\lambda I$ 的 i 阶顺序主子式，则有

$$P_0(\lambda) = 1, \quad P_1(\lambda) = \alpha_1 - \lambda$$

$$P_i(\lambda) = (\alpha_i - \lambda) P_{i-1}(\lambda) - \beta_i^2 P_{i-2}(\lambda) \qquad (6.4.1)$$

$$i=2,\cdots,n$$

T 实对称，所以 $P_i(\lambda)$ 的根全为实数.

定理 6.4.1.

(1) 存在 $M>0$ 使得当 $\lambda > M$ 时，

$P_i(-\lambda)>0$，　$\mathrm{sgn}(P_i(\lambda)) = (-1)^i$.

(2) $P_{i-1}(\lambda)$，$P_i(\lambda)$ 无公共根

(3) 若 $P_i(\mu)=0$，则

$$P_{i-1}(\mu)\,P_{i+1}(\mu) < 0$$

(4) $P_i(\lambda)$ 全是单根，并且 $P_i(\lambda)$ 的根严格分隔 $P_{i+1}(\lambda)$ 的根.

证明:

(1). $P_i(\lambda)$ 的首项为 $(-1)^i\lambda^i$ 由此可得 (1)

(2) 反证法. 假设存在 i 使得 $P_{i-1}(\lambda)$，$P_i(\lambda)$ 有公共根 μ. 即 $P_{i-1}(\mu) = P_i(\mu)=0$，则由 (6.4.1) 得

$$0 = P_i(\mu) = (\alpha_i-\mu)P_{i-1}(\mu) - \beta_i^2 P_{i-2}(\mu) = -\beta_i^2 P_{i-2}(\mu)$$

由于 $\beta_i \neq 0$，故 $P_{i-2}(\mu)=0$

依此类推可得 $P_0(\mu)=0$ 与 $P_0(\mu)=1$ 矛盾

(3). 设 $p_i(\mu)=0$ 则由 (6.4.1) 及 (2)

$$p_{i+1}(\mu)\,p_{i-1}(\mu) = -\beta_{i+1}^2\,p_{i-1}(\mu)^2 < 0$$

(4) 数学归纳法. 当 $i=1$ 时, $p_1(\lambda)=\alpha_1-\lambda$

α_1 为 $p_1(\lambda)$ 的根.

$p_2(\alpha_1) = -\beta_2^2 < 0$

另一方面

$$p_2(\lambda)\to+\infty, \quad \lambda\to\pm\infty$$

所以在 $(-\infty,\alpha_1)$ 与 $(\alpha_2,+\infty)$ 之间

$p_2(\lambda)$ 各有一个根.

设 $i=k$ 时结论成立. 设 p_{k-1}, p_k

的根分别为

$$\nu_1<\nu_2<\cdots<\nu_{k-1}, \quad \mu_1<\mu_2<\cdots<\mu_k$$

由归纳法假设知

$$\mu_1<\nu_1<\mu_2<\nu_2<\cdots<\nu_{k-1}<\mu_k$$

由 (1) 及 $p_{k-1}(\nu_j)=0$ 可得

$$(-1)^{j-1}p_{k-1}(\mu_j)>0$$

由递推公式 (6.4.1) 得

$$p_{k+1}(\mu_j) = -\beta_{k+1}^2\,p_{k-1}(\mu_j), \quad j=1,\cdots,k$$

所以

$$(-1)^j p_{k+1}(\mu_j)>0, \quad j=1,\cdots,k$$

注意到

$$\lim_{\mu\to+\infty} P_{k+1}(-\mu) = +\infty$$

$$\lim_{\mu\to+\infty} (-1)^{k+1} P_{k+1}(\mu) = +\infty$$

所以 $P_{k+1}(\lambda)$ 在区间

$(-\infty, \mu_1)$ (μ_1, μ_2) $\cdots$ (μ_{k-1}, μ_k), $(\mu_k, +\infty)$

各有一个根. $\square$

设 $\mu \in \mathbb{R}$, 令 $S_k(\mu)$ 为 $P_0(\mu), \cdots, P_k(\mu)$ 的变号次数. 规定

若 $P_i(\mu) = 0$, 则 $P_i(\mu)$ 与 $P_{i-1}(\mu)$ 同号

定理 6.4.2. $S_k(\mu)$ 为 $P_k(\lambda)$ 在 $(-\infty, \mu)$ 内根的个数.

证明: 数学归纳法.

$k=1$ 时结论显然成立.

假设 $k=\ell$ 时成立. 设 $P_\ell(\lambda)$, $P_{\ell+1}(\lambda)$ 的根分别为

$$\mu_1 < \mu_2 < \cdots < \mu_\ell, \quad \lambda_1 < \lambda_2 < \cdots < \lambda_{\ell+1}$$

则由定理 6.4.1 知

$$\lambda_1 < \mu_1 < \lambda_2 < \mu_2 < \cdots < \mu_\ell < \lambda_{\ell+1}$$

设 $S_l(\mu) = m$，由归纳假设有

$$\mu_m < \mu \le \mu_{m+1}$$

下面分情况讨论：

(1). $\lambda_m < \mu_m < \mu \le \lambda_{m+1}$

此时 $P_l(\mu)$ 与 $P_{l+1}(\mu)$ 有相同的符号

即 $S_{l+1}(\mu) = S_l(\mu) = m$

与 $P_{l+1}(\lambda)$ 在 $(-\infty, \mu)$ 根的个数相同.

(2) $\lambda_{m+1} < \mu < \mu_{m+1}$

此时 $P_l(\mu)$ 与 $P_{l+1}(\mu)$ 异号. 因此

$$S_{l+1}(\mu) = S_l(\mu) + 1 = m+1$$

(3) $\mu = \mu_{m+1}$

此时 $P_l(\mu) = 0$ 则有 $P_l(\mu)$ 与 $P_{l-1}(\mu)$ 同号.

由定理 6.4.1 得 $P_{l-1}(\mu)$ 与 $P_{l+1}(\mu)$ 异号

因此 $P_l(\mu)$ 与 $P_{l+1}(\mu)$ 异号. 即

$$S_{l+1}(\mu) = S_l(\mu) + 1 = m+1 \qquad \square$$

推论 6.4.1： T 是不可约对称三对角矩阵. 则 $S_n(\mu)$ 是 T 在 $(-\infty, \mu)$ 内特征值的个数.

设 T 的特征值为

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n$$

则

$$|\lambda_i| \le \rho(T) \le \|T\|_\infty$$

取 $l_0 = -\|T\|_\infty$, $u_0 = \|T\|_\infty$

则 $\lambda_m \in [l_0, \mu_0]$

令 $r_1 = (l_0 + \mu_0)/2$

若 $s_n(r_1) \geq m$ 则 $\lambda_m \in [l_0, r_1]$

否则 $\lambda_m \in [r_1, u_0]$

依此类推，可将区间不断二分.

二分法的主要计算量为计算 $s_n(\mu)$

为了避免高阶多项式溢出. 令

$$q_i(\lambda) = \frac{p_i(\lambda)}{p_{i-1}(\lambda)}, \quad i=1, \cdots, n.$$

则有

$$q_1(\lambda) = p_1(\lambda) = \alpha_1 - \lambda$$

$$q_i(\lambda) = \alpha_i - \lambda - \frac{\beta_i^2}{q_{i-1}(\lambda)}, \quad i=2, \cdots, n$$

$s_n(\mu)$ 为 $q_1(\mu) \cdots q_n(\mu)$ 中负数的个数.

$S_n(\mu)$ 计算方法：

$$
\begin{array}{l}
x = [\alpha_1, \alpha_2, \cdots, \alpha_n] \\
y = [0, \beta_2, \cdots, \beta_n] \\
s = 0, \quad q = x(1) - \mu \\
k = 1 : n \\
\quad \text{if } q < 0 \\
\quad\quad\quad s = s + 1 \\
\quad \text{end} \\
\quad \text{if } k < n \\
\quad\quad \text{if } q = 0 \\
\quad\quad\quad q = |y(k+1)|\varepsilon \\
\quad\quad \text{end} \\
\quad\quad q = x(k+1) - \mu - y(k+1)^2 / q \\
\quad \text{end} \\
\text{end}
\end{array}
$$