第九章 特征值问题的数值解法

§1 幂法与反幂法

§2 对称Jacobi迭代法

§3 QR迭代法

§1　幂法与反幂法

□回顾

✓ 设A是n阶方阵，若数λ和非零向量x满足$Ax=\lambda x$
则称λ的A的特征值，x为A的对应于λ特征向量。

✓ λ是A的特征值$\Leftrightarrow(\lambda E-A)x=0$，其中$x\neq 0$

$\Leftrightarrow$方程$(\lambda E-A)x=0$有非零解

✓ 特征多项式：$f(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda-a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & \lambda-a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$

$$=\lambda^n-c_1\lambda^{n-1}+\cdots+(-1)^{n-1}c_{n-1}\lambda+(-1)^n c_n$$

特征方程：$f(\lambda)=|\lambda E-A|=0$

✓ 矩阵相似对角化：∃可逆矩阵P,使得$P^{-1}AP=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$

其中(λ_i,α_i)为A的特征对，$P=[\alpha_1,\alpha_2,\cdots,\alpha_n]$.

✓ A可相似对角化 $\Leftrightarrow$ A有n个线性无关的特征向量

（可构成R^n的基底）。

✓ 实对称矩阵一定相似对角化：∃正交矩阵Q,使得

$$Q^TAQ=Q^{-1}AQ=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$$

幂法

假设$A \in R^{n\times n}$有n个线性无关的特征向量，A的n个特征值按模大小排列

$$|\lambda_1| > |\lambda_2| \ge \cdots \ge |\lambda_n|$$

对应的特征向量为$x_1, x_2, \cdots, x_n$（R^n的一个基底）

任取$v_0 \in R^n (v_0 \ne 0)$，并设v_0在基底$x_1, x_2, \cdots, x_n$下的坐标为

$$v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots \alpha_n x_n \quad (\alpha_1 \ne 0)$$

考察：

$$\begin{aligned} v_k &= A^k v_0 \\ &= \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \cdots + \alpha_n \lambda_n^k x_n \\ &= \lambda_1^k \left(\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k x_n \right) \\ &\triangleq \lambda_1^k \left(\alpha_1 x_1 + \varepsilon_k \right) \end{aligned}$$

由于 $\left|\frac{\lambda_i}{\lambda_1}\right| < 1 (i = 2, \cdots, n)$

故 $\varepsilon_k \to 0 (k \to \infty)$

$$\therefore v_k \xrightarrow{k\to+\infty} x_1$$

k充分大$x_1 \approx v_k$

$$v_k = A^k v_0 = \ldots = \lambda_1^k \left(\alpha_1 x_1 + \varepsilon_k \right)$$

又因为 $Ax_1 = \lambda_1 x_1$

两边左乘x_1^T $x_1^T A x_1 = x_1^T \left(\lambda_1 x_1 \right) = \lambda_1 x_1^T x_1$

所以 $$\lambda_1 = \frac{x_1^T A x_1}{x_1^T x_1} \approx \frac{v_k^T A v_k}{v_k^T v_k}$$

问题

✓如果$|\lambda_1| > 1$,则$v_k = A^k v_0 \to \infty (\lambda_1^k \to \infty)$,产生计算溢出;

✓如果$|\lambda_1| < 1$,则$v_k = A^k v_0 \to 0 (\lambda_1^k \to 0), v_k = 0$:非特征向量;

具体算法

1 取$v_0 \in R^n$，满足$\|v_0\|_2 = 1$(要求$\alpha_1 \neq 0$)

$$\begin{cases} u_k = Av_{k-1} \\ v_k = \dfrac{u_k}{\|u_k\|_2} \\ m_k = v_k^T A v_k \end{cases}, k = 1, 2, \cdots$$

则有$\lim\limits_{k\to\infty} v_k = \pm \dfrac{x_1}{\|x_1\|_2}$和$\lim\limits_{k\to\infty} m_k = \lambda_1$.

证明 $\because v_k = \dfrac{u_k}{\|u_k\|_2} = \dfrac{Av_{k-1}}{\|u_k\|_2} = \dfrac{A^2 v_{k-2}}{\|u_k\|_2 \cdot \|u_{k-1}\|_2} = \cdots = \dfrac{A^k v_0}{\|A^k v_0\|_2}$

而 $A^k v_0 = \lambda_1^k(\alpha_1 x_1 + \varepsilon_k)$，$\varepsilon_k \to 0$

$$\therefore v_k = \frac{\lambda_1^k(\alpha_1 x_1 + \varepsilon_k)}{|\lambda_1^k| \cdot \|\alpha_1 x_1 + \varepsilon_k\|_2} \to \pm \frac{x_1}{\|x_1\|_2}, \text{ 即} v_k \to x_1.$$

$$m_k = v_k^T A v_k = \frac{\lambda_1^k(\alpha_1 x_1 + \varepsilon_k)^T}{|\lambda_1|^k \cdot \|\alpha_1 x_1 + \varepsilon_k\|_2} \cdot \frac{\lambda_1^{k+1}(\alpha_1 x_1 + \varepsilon_{k+1})}{|\lambda_1|^k \cdot \|\alpha_1 x_1 + \varepsilon_k\|_2} = \frac{\lambda_1^{2k+1}}{\lambda_1^{2k}} \cdot \frac{\alpha_1^2 x_1^T x_1 + \xi_k}{\alpha_1^2 x_1^T x_1 + \eta_k} \to \lambda_1$$

2 迭代结束标准：$\|Av_k - m_k v_k\|_2 < \varepsilon$

3 收敛速度：$\varepsilon_k = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$

$\left|\frac{\lambda_2}{\lambda_1}\right|$为收敛因子，如果$\left|\frac{\lambda_2}{\lambda_1}\right| \approx 1$，则收敛很慢；

4 其他格式

$$\begin{cases} 任取 v_0 \in R^n (\alpha_1 \neq 0) \\ u_k = Av_{k-1} \\ m_k = \max(u_k) \\ v_k = \dfrac{u_k}{m_k} \end{cases}$$

则 $v_k \to x_1$，$m_k \to \pm\lambda_1 (|m_k| \to |\lambda_1|)$。

反幂法

1 假设可逆矩阵A的值$|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_{n-1}| > |\lambda_n|$，则$A^{-1}$的特征值

$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \ge \cdots$，将幂法用于$A^{-1}$，则可求得$A$的按模最小的特征值。

2 位移反幂法　假设$\tilde{\lambda}_m$是λ_m的近似值，求λ_m及其对应的特征向量

一般地，$0 < |\tilde{\lambda}_m - \lambda_m| << |\tilde{\lambda}_m - \lambda_i|, (i \ne m)$,

考虑矩阵$A - \tilde{\lambda}_m I$的特征值$\lambda_i - \tilde{\lambda}_m$按模最小的特征值$\lambda_m - \tilde{\lambda}_m$

对$A - \tilde{\lambda}_m I$应用反幂法：　取$v_0 (\alpha_1 \ne 0)$,

$$\begin{cases} u_k = (A - \tilde{\lambda}_m I)^{-1} v_{k-1} \Leftrightarrow (A - \tilde{\lambda}_m I) u_k = v_{k-1} \\ v_k = \dfrac{u_k}{\|u_k\|_2} \\ m_k = v_k^T A v_k \quad \to \lambda_m \end{cases} , k = 1, 2, \cdots$$

则当k充分大时，v_k是$(A - \tilde{\lambda}_m I)^{-1}$的近似特征向量，也是$A$的近似特征向量。

§2 对称Jacobi迭代法

❑ 问题

设A是n阶实对称矩阵, 求A的所有特征值和对应的特征向量。

❑ 基本结论

n阶实对称矩阵A一定可以正交相似对角化, 即∃Q($Q^TQ=I$),使得

$$Q^T A Q = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

其中λ_1，λ_2,⋯，λ_n是A的n个特征值。

❑ 计算公式

✓ 2阶是对称矩阵的相似对角化

$$A=\begin{bmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{bmatrix} (a_{pq}=a_{qp}\neq 0)$$

利用Givens变换矩阵

$$Q=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \equiv \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

使得$Q^T AQ$变成对角矩阵，经计算

$$Q^T AQ=\begin{bmatrix} c^2 a_{pp}-2csa_{pq}+s^2 a_{qq} & c^2 a_{pq}+cs(a_{pp}-a_{qq})-s^2 a_{pq} \\ c^2 a_{pq}+cs(a_{pp}-a_{qq})-s^2 a_{pq} & c^2 a_{qq}+2csa_{pq}+s^2 a_{pp} \end{bmatrix}$$

令 $$c^2 a_{pq}+cs(a_{pp}-a_{qq})-s^2 a_{pq}=0$$

$$c^2 a_{pq} + cs(a_{pp} - a_{qq}) - s^2 a_{pq} = 0$$

即 $$\left(\frac{s}{c}\right)^2 + \frac{a_{qq} - a_{pp}}{a_{pq}} \frac{s}{c} - 1 = 0$$

令 $$t = \frac{s}{t} (= \tan\theta), \quad d = \frac{a_{qq} - a_{pp}}{2a_{pq}}$$

可得 $$t^2 + 2dt - 1 = 0$$

由二次多项式求根公式，解得$\left(|\theta| \le \pi/4, |t| \le 1\right)$

$$t = \frac{\mathrm{sign}(d)}{|d| + \sqrt{1 + d^2}}$$

所以 $$c = \frac{1}{\sqrt{1 + t^2}}, \quad s = c \cdot t$$

✓ n阶是对称矩阵

$$A=\begin{bmatrix}\ddots & & & & \iddots\\ & a_{pp} & \cdots & a_{pq} & \\ & \vdots & \ddots & \vdots & \\ & a_{qp} & \cdots & a_{qq} & \\ \iddots & & & & \ddots\end{bmatrix}$$

如果$a_{pq}=a_{qp}\neq 0(p<q)$，令

$$Q=G(p,q;\theta)=\begin{array}{cc}\begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I\end{bmatrix} & \begin{array}{c} \\ p \\ \\ q \\ \\ \end{array} \\ \begin{array}{ccccc} \quad & p & \quad & q & \quad \end{array} & \end{array}$$

则$Q^T AQ$将a_{pq}和a_{qp}化为0,且只改变A的第p,q行和列元素，其他不变。

✓一个例子

用古典Jacobi迭代法求$A=\begin{pmatrix}4&2&2\\2&5&1\\2&1&6\end{pmatrix}$的全部特征值。

解 记$A^0=A$，取$p=1,q=2,a_{pq}^0=a_{12}^0=2$，则

$$s=\frac{a_{11}^{(0)}-a_{22}^{(0)}}{2a_{12}^{(0)}}=-0.25,\quad t=-0.780776$$

$$\cos\theta=(1+t^2)^{-\frac{1}{2}}=0.788206,\sin\theta=t\cos\theta=-0.615412$$

故 $$Q=G(p,q;\theta)=\begin{pmatrix}\cos\theta&-\sin\theta&0\\\sin\theta&\cos\theta&0\\0&0&1\end{pmatrix}=\begin{pmatrix}0.788206&0.615412&0\\-0.615412&0.788206&0\\0&0&1\end{pmatrix}$$

所以 $$A^{(1)} = Q_1^T A^{(0)} Q_1 = \begin{pmatrix} 2.438448 & 0 & 0.961 \\ 0 & 6.561552 & 2.020190 \\ 0.961 & 2.020190 & 6 \end{pmatrix}$$

再取$p = 2, q = 3, a_{pq}^1 = a_{23}^1 = 2.020190$，则类似可得

$$A^{(2)} = \begin{pmatrix} 2.438448 & 0.631026 & 0.724794 \\ 0.631026 & 8.320386 & 0 \\ 0.724794 & 0 & 4.241166 \end{pmatrix}$$

$$A^{(3)} = \begin{pmatrix} 2.183185 & 0.595192 & 0 \\ 0.595192 & 8.320386 & 0.209614 \\ 0 & 0.209614 & 4.496424 \end{pmatrix}$$

$$A^{(4)} = \begin{pmatrix} 2.125995 & 0 & -0.020048 \\ 0 & 8.377576 & 0.208653 \\ -0.020048 & 0.208653 & 4.496424 \end{pmatrix}$$

$$\mathbf{A}^{(5)}=\begin{pmatrix} 2.125995 & -0.001073 & -0.020019 \\ -0.001073 & 8.388761 & 0 \\ -0.020019 & 0 & 4.485239 \end{pmatrix}$$

$$\mathbf{A}^{(6)}=\begin{pmatrix} 2.125825 & -0.001072 & 0 \\ 0 & 8.388761 & 0.000009 \\ -0.001072 & 0.000009 & 4.485401 \end{pmatrix}$$

$$\mathbf{A}^{(7)}=\begin{pmatrix} 2.125825 & 0 & 0 \\ 0 & 8.388761 & 0.000009 \\ 0 & 0.000009 & 4.485401 \end{pmatrix}$$

所以A的特征值可取为：

$$\lambda_1 \approx 2.125825,\quad \lambda_2 \approx 8.388761,\quad \lambda_3 \approx 4.485401$$

记 $\mathrm{off}(A) \equiv \sqrt{\sum_{i \neq j} a_{ij}^2}$

✓ 基本结论

定理1 对任意的 $p \neq q$，设 $B=Q^T AQ$，则 $\mathrm{off}^2(B) = \mathrm{off}^2(A) - 2a_{pq}^2$.

定理2 设 $A^{(k)}$ 是古典 *Jacobi* 迭代法迭代 k 次的矩阵，则

$$\mathrm{off}(A^{(k)})^2 \leq \left(1 - \frac{1}{N}\right)^k \cdot \mathrm{off}(A)^2$$

✓ Matlab命令： $[\mathrm{Q},\mathrm{d}] = \mathrm{eig}(A)$

Jacobi方法具有方法简单紧凑, 精度高, 收敛较快等优点, 是计算对称矩阵全部特征值和相应特征向量的有效方法, 但计算量较大, 一般适用于阶数不高的矩阵.