7.3 Jacobi方法

我们知道，若矩阵 $A \in R^{n\times n}$ 为对称矩阵，则存在一正交阵 P，使

$$PAP^T = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) = D。$$

D 的对角元 $\lambda_i (i = 1,2,\ldots, n)$ 就是 A 的特征值，P^T 的列向量就是对应于特征值的特征向量。于是求实对称矩阵A的特征值问题就转为寻找正交矩阵P，使$PAP^T = D$为对角阵，而这个问题的主要困难在于如何构造P。

*Jacobi*方法是用来计算实对称矩阵的全部特征值及对应的特征向量的一种变换方法，其基本思想是对矩阵作一系列正交相似变换，使其非对角线元素收敛到零。所用的变换是*Jacobi*旋转变换。下面先讨论*Jacobi*旋转变换及其性质。

在R^n中的$\{x_i, x_j\}$平面内的平面旋转变换为

$$\begin{cases} y_i = x_i \cos\theta + x_j \sin\theta, \\ y_j = -x_i \sin\theta + x_j \cos\theta \\ y_k = x_k, k \neq i, j, \end{cases}$$

或写成 $y = Jx$，其中，$x = (x_1, x_2, \ldots, x_n)^T$，$y = (y_1, y_2, \ldots y_n)^T$，

$$J = J(i, j) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & \cdots & s & & \\ & & \vdots & \ddots & \vdots & & \\ & & -s & \cdots & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}。$$

称J为平面旋转矩阵，它只有在$(i,i),(i,j),(j,i)$和(j,j)位置上的元素与单位矩阵不一样，分别为 $c=\cos\theta, s=\sin\theta, -\sin\theta$和$\cos\theta$。

显然，矩阵J是正交矩阵，JA只改变A的第i行与第j行的元素，AJ^T只改变A的第i列与第j列的元素，JAJ^T只改变A的第i行，第j行，第i列，第j列的元素。

设$A=(a_{ij})\in R^{n\times n}$为对称矩阵，$J(i,j)$为一平面旋转矩阵，则$B=JAJ^T=(b_{ij})$的元素的计算公式为

$$
\begin{aligned}
b_{ii} &= a_{ii}\cos^2\theta + a_{jj}\sin^2\theta + 2a_{ij}\sin\theta\cos\theta,\\
b_{jj} &= a_{ii}\sin^2\theta + a_{jj}\cos^2\theta - 2a_{ij}\sin\theta\cos\theta,\\
b_{ij} &= b_{ji} = \frac{1}{2}(a_{jj}-a_{ii})\sin 2\theta + a_{ij}\cos 2\theta,\\
b_{ik} &= b_{ki} = a_{ik}\cos\theta + a_{jk}\sin\theta, k\neq i,j,\\
b_{jk} &= b_{kj} = a_{jk}\cos\theta - a_{ik}\sin\theta, k\neq i,j,\\
b_{lm} &= b_{ml} = a_{lm}, l\neq i,j; m\neq i,j。
\end{aligned}
$$

而且，不难验证

$$b_{ii}^2+b_{jj}^2+2b_{ij}^2=a_{ii}^2+a_{jj}^2+2a_{ij}^2。\qquad \textbf{(7.3.1)}$$

定理7.7 设$A\in R^{n\times n}$为对称矩阵，若$B=PAP^T$，P为正交阵，

则有$\|B\|_F=\|A\|_F$。

证

设$\lambda_i(i=1,2,\cdots\cdots n)$为$A$的特征值，则

$$\|A\|_F^2=\sum_{i=1}^n\sum_{j=1}^n a_{ij}^2=tr(A^TA)=tr(A^2)=\sum_{i=1}^n\lambda_i^2。$$

另一方面，矩阵B的特征值也为$\lambda_i(i=1,2,\cdots\cdots n)$,

$$\|B\|_F^2=tr(B^2)=\sum_{i=1}^n\lambda_i^2$$

因此,$\|B\|_F^2=\|A\|_F^2$。定理得证。

设A的非对角线元素 $a_{ij}\neq 0$,我们可选择平面旋转矩 阵$J(i,j)$，使$B=JAJ^T$的非对角线元素$b_{ij}=b_{ji}=0$。为此，由矩阵B元素的计算公式可知，可选择θ，使

$$\tan 2\theta = \frac{2a_{ij}}{a_{ii} - a_{jj}}, |\theta| \le \frac{\pi}{4} \qquad (7.3.2)$$

如果用D（A）表示A的对角线元素的平方和，用S（A）表示A的非对角线元素的平方和，则对 $B = JAJ^T$，由（7.3.1）、（7.3.2）和定理7.7可知

$$D(B) = D(A) + 2a_{ij}^2$$

$$S(B) = S(A) - 2a_{ij}^2$$

这说明B的对角线元素的平方和比A的对角线元素的平方和增加了$2a_{ij}^2$,而B的非对角线元素的平方和减少了$2a_{ij}^2$.这就是Jacobi方法求矩阵特征值和特征向量的依据。下面说明Jacobi方法的计算过程。

先在$A = A_0 = (a_{ij}^{(0)})$中选择非对角元中绝对值最大的$a_{ij}^{(0)}$。可设$a_{ij}^{(0)} \neq 0$,否则A已经对角化了。由（7.3.2）选择平面旋转矩阵J_1,

使$J_1A_0J_1^T=A_1$的元素$a_{ij}^{(1)}=0$。计算出A_1，再类似的选择J_2，计算

$A_2=J_2A_1J_2^T$，继续这个过程，连续对A旋行一系列平面旋转变换

消除非对角线绝对值最大的元素，直到将A的非对角线元素全化为充分小为止。

定理7.8 设$A\in R^{n\times n}$为对称矩阵，A施行上述一系列平面旋转变换

$$A_m=J_mA_{m-1}J_m^T\ ,m=1,2,\cdots\cdots,$$

则有$\lim\limits_{n\to\infty}S(A_m)=0$。

证 设$\left|a_{ij}^{(m)}\right|=\max\limits_{l\neq k}\left|a_{lk}^{(m)}\right|$,由于

$$S(A_{m+1})=S(A_m)-2(a_{ij}^{(m)})^2,$$

$$S(A_m)=\sum_{l\neq k}(a_{lk}^{(m)})^2\leq n(n-1)(a_{ij}^{(m)})^2,$$

则有

$$S(A_{m+1}) \le S(A_m)\left(1-\frac{2}{n(n-1)}\right)。$$

反复利用上式，即得

$$S(A_{m+1}) \le S(A_0)\left(1-\frac{2}{n(n-1)}\right)^{n+1}, n>2。$$

故$\lim\limits_{n\to\infty} S(A_m)=0$，定理得证。

我们指出，可以证明A_m的对角线元素一定有极限。

设m充分大时，有

$$A_m = J_m \cdots J_2 J_1 A J_1^T J_2^T \cdots J_m^T \approx D,$$

D为对角阵，则A_m的对角线元素就是A的近似特征值，$Q_m = J_1^T J_2^T \cdots J_m^T$的列向量就是对应的近似特征向量。可用$S(A_m) < \varepsilon$控制迭代终止，其中$\varepsilon$是要求的精度。

例7.3 用Jacobi方法计算矩阵

$$A=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

的特征值的特征向量。

解 先取 $(i,j)=(1,2)$,按（7.3.2）有$\cot 2\theta=0, s=c=1/\sqrt{2}$，所以

$$J_1=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_1=J_1AJ_1^T=\begin{pmatrix} 1 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 3 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 \end{pmatrix}。$$

再取（i, j）$=(1,3)$有 $\cot 2\theta=\frac{1}{\sqrt{2}}$， $c=0.88808, s=0.45970,$

$$J_2 = \begin{pmatrix} 0.88808 & 0 & 0.45970 \\ 0 & 1 & 0 \\ -0.45970 & 0 & 0.88808 \end{pmatrix},$$

$$A_2 = J_2 A_1 J_2^T = \begin{pmatrix} 0.63398 & -0.32505 & 0 \\ -0.32505 & 3 & -0.62797 \\ 0 & -0.62797 & 2.36603 \end{pmatrix}。$$

这里我们看到经变换后所得的非对角线元素的最大绝对值逐次变小。继续做下去，可以得到

$$A_9 = \begin{pmatrix} 0.58578 & 0.00000 & 0.00000 \\ 0.00000 & 2.00000 & 0.00000 \\ 0.00000 & 0.00000 & 3.41421 \end{pmatrix},$$

$$Q_9 = J_1^T J_2^T \cdots J_9^T = \begin{pmatrix} 0.50000 & 0.70710 & 0.50000 \\ 0.70710 & 0.00000 & -0.70710 \\ 0.50000 & -0.70710 & 0.50000 \end{pmatrix},$$

矩阵A的近似特征值和特征向量均已求出。

用Jacobi方法求得的结果精度一般都比较高，特别是求得的特征向量正交性很好。所以Jacobi方法是求实对称矩阵全部特征值和特征向量的一个较好的方法。它的弱点是计算量大，对原矩阵是稀疏矩阵，旋转变换后不能保持其稀疏的性质。