第四章　矩阵的特征值与特征向量的计算

武艳云

重庆邮电大学

2023年7月

目录

◇主特征值：按模最大的特征值.

◇关于特征值的一些结果：

1,设A的特征值为$\lambda_1, \lambda_2, \cdots, \lambda_n$,则

(1)$\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$;

(2)$\lambda_1 \lambda_2 \cdots \lambda_n = |A|$.

2,设A的特征值为λ,且$Ax = \lambda x$,$x \neq 0$, 则

(1)$\lambda - p$ 为$A - pI$的特征值;

(2)λ^2为A^2的特征值;

(3)设A为非奇异矩阵，那么$\lambda \neq 0$,且$\frac{1}{\lambda}$为A^{-1}的特征值。

4.1.1 乘幂法

乘幂法：计算一个矩阵的主特征值及其相应的特征向量的一种迭代法。

设A的特征值按模大小排列顺序为$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$，并且对应的特征向量$x_1, x_2, \cdots, x_n$线性无关,设$v_0$是任意的一个$n$维非零向量，则$v_0$可以唯一地表示为：

$$v_0 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

令$v_k = Av_{k-1}$,则

$$v_k = Av_{k-1} = A^2 v_{k-2} = \cdots = A^k v_0 = \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \cdots + \alpha_n \lambda_n^k x_n$$

$$v_{k+1} = \alpha_1\lambda_1^{k+1}x_1 + \alpha_2\lambda_2^{k+1}x_2 + \cdots + \alpha_n\lambda_n^{k+1}x_n =$$
$$\lambda_1^{k+1}\left(\alpha_1 x_1 + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k+1}x_2 + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1}x_n\right)$$

当k充分大时，就有

$$v_{k+1} = \lambda_1^{k+1}\left(\alpha_1 x_1 + \sum_{j=2}^{n}\alpha_j\left(\frac{\lambda_j}{\lambda_1}\right)^{k+1}x_j\right) \approx \lambda_1^{k+1}\alpha_1 x_1$$

相应地有

$$v_k \approx \lambda_1^k \alpha_1 x_1 \tag{1}$$

所以

$$\lambda_1 \approx \frac{(v_{k+1})_i}{(v_k)_i}$$

事实上，可以证明 $\frac{(v_{k+1})_i}{(v_k)_i} = \lambda_1 + O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^k\right).$

.

由(1)可知，当$|\lambda_1| > 1$时，$\lambda_1^k \to \infty$;当$|\lambda_1| < 1$，$\lambda_1^k \to 0$,在计算机上会出现"上溢"或"下溢"。所以在每步计算时将z_k单位化。求A的按模最大的特征值λ_1的乘幂法如下：

算法1(乘幂法):

(1)任取初始向量v_0,$m_0 = \max\{v_0\}$,允许误差ϵ,最大迭代步数N;

(2)for $k = 1, 2, \cdots, N$;

(i)计算$u_{k-1} = \frac{v_{k-1}}{m_{k-1}}$,

(ii)计算$v_k = Au_{k-1}$,$m_k = \max\{v_k\}$;

(iii)若$|m_k - m_{k-1}| \le \epsilon$，则取$\lambda_1 = m_k$,停止计算.

例4.1 利用乘幂法计算一下矩阵的特征值和相应的特征向量。

$$A = \begin{pmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{pmatrix}$$

解 取$v_0 = (0,0,1)^T$,则$m_0 = 1$,$u_0 = (0,0,1)^T$, $v_1 = Au_0 = (2,4,1)^T$,且$m_1 = \max\{v_1\} = 4$,$u_1 = \frac{v_1}{m_1} = (0.5,1,0.25)^T$,利用上述算法，计算结果如下表

k		u_k^T		a
0	0	0	1	1
1	0.5	1.0	0.25	4
2	0.5	1.0	0.8611	9
3	0.5	1.0	0.7306	11.44
4	0.5	1.0	0.753	10.92
5	0.5	1.0	0.749	11.01
6	0.5	1.0	0.750	10.99
7	0.5	1.0	0.750	11.00
8	0.5	1.0	0.750	11.00

从表中可以看出，矩阵A的绝对值最大的特征值为$\lambda_1 = 11$,相应的特征向量$x_1 = [0.5, 1.0, 0.75]^T$

若$|\lambda_1| = |\lambda_2| = \cdots = |\lambda_m| > |\lambda_{m+1}| \geq \cdots \geq |\lambda_n|$时：

(1)λ_1是m重根，即$\lambda_1 = \lambda_2 = \cdots = \lambda_m$，

$$\begin{aligned} v_{k+1} =& \alpha_1\lambda_1^{k+1}x_1 + \alpha_2\lambda_2^{k+1}x_2 + \cdots + \alpha_n\lambda_n^{k+1}x_n \\ =& \lambda_1^{k+1}\Big[\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m \\ & + \alpha_{m+1}\left(\frac{\lambda_{m+1}}{\lambda_1}\right)^{k+1} x_{m+1} + \cdots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1} x_n\Big] \end{aligned} \tag{2}$$

当$k \to \infty$,$v_{k+1} \approx \lambda_1^{k+1}(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m)$，因为$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m$也是矩阵$A$的对应于$\lambda_1$的特征值，因此有

$$\lambda_1 = \lambda_2 = \cdots = \lambda_m = \frac{(v_{k+1})_i}{(v_k)_i}.$$

$(2)\lambda_1 = -\lambda_2, |\lambda_1| > |\lambda_3|$

$$\begin{aligned} v_{k+1} =& \alpha_1 \lambda_1^{k+1} x_1 + \alpha_2 \lambda_2^{k+1} x_2 + \cdots + \alpha_n \lambda_n^{k+1} x_n \\ =& \lambda_1^{k+1} \left[\alpha_1 x_1 + (-1)^{k+1} \alpha_2 x_2 + \alpha_3 \left(\frac{\lambda_3}{\lambda_1} \right)^{k+1} x_3 + \cdots + \right. \\ & \left. \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^{k+1} x_n \right] \end{aligned} \tag{3}$$

v_{k+1}是个摆动数列，当$k \to \infty$,有

$$v_{2k-1} \approx \lambda_1^{2k-1} (\alpha_1 x_1 - \alpha_2 x_2),$$

$$v_{2k} \approx \lambda_1^{2k} (\alpha_1 x_1 + \alpha_2 x_2)$$

因此有

$$\lambda_1^2 \approx \frac{(v_{k+2})_i}{(v_k)_i},$$

即 $$\lambda_{1,2} \approx \pm\sqrt{\frac{(v_{k+2})_i}{(v_k)_i}}.$$

又由 $$v_{k+1} \approx \lambda_1^{k+1}\left[\alpha_1 x_1 + (-1)^{k+1}\alpha_2 x_2\right],$$ 可以导出

$$\begin{cases} v_{k+1} + \lambda_1 v_k \approx 2\lambda_1^{k+1}\alpha_1 x_1, \\ v_{k+1} - \lambda_1 v_k \approx 2\lambda_1^{k+1}(-1)^{k+1}\alpha_2 x_2. \end{cases}$$

故在这种情况下，仍可以按乘幂法产生向量序列，得到主特征值和对应的特征向量。

4.1.2 反幂法

反幂法：求矩阵按模最小的特征值和特征向量。

设有n阶非奇异矩阵A，其特征值与相应的特征向量分别是λ_i和$x_i(i=1,2,\cdots,n)$,则$A^{-1}x_i=\lambda^{-1}x_i$，即$\lambda_i^{-1}$是$A^{-1}$的特征值，其特征向量仍然是$x_i$.
如果A的特征值为如此情况：

$$|\lambda_1| \geq |\lambda_2| \geq \cdots |\lambda_{n-1}| > |\lambda_n|$$

则λ_n^{-1}一定是A^{-1}的主特征值。

对A^{-1}采用采用乘幂法即可求得矩阵A^{-1}的主特征值$\mu_n=\frac{1}{\lambda_n}$，得到矩阵$A$的按模最小的特征值$\lambda_n$。

反幂法的具体计算步骤为：

1,任取初始向量$v_0 = u_0 \neq 0$;

2,计算$v_k = A^{-1}u_{k-1}(k = 1, 2, \cdots)$;

3,$m_k = \max\{v_k\}$;$v_k = \frac{v_k}{m_k}$;

4,如果k从某时以后有$\frac{(v_k)_j}{(v_{k-1})_j} \approx c$(常数)$(j = 1, 2, \cdots, n)$,则取$\lambda_n \approx \frac{1}{c}$;而$u_k$就是与$\lambda_n$对应的特征向量。

例4.2 用反幂法求矩阵$A=\begin{pmatrix} 3 & 2 \\ 4 & 5 \end{pmatrix}$的按模最小的特征值和特征向量，精确至6位有效数字。

解 取$v_0=u_0=(1,1)^T$,由$v_1=A^{-1}u_0$,即$Av_1=u_0$, 得

$$v_1=(0.428571,-0.142857)^T, m_1=\max v_1=0.428571$$

归一化得$u_1=(1.000000,-0.333333)^T$. 反复进行，计算结果如下表：

则A的按模最小的特征值为$\frac{1}{m_7}\approx 1.00000$,相应的特征向量为

$$(1,00000,-0.999993)^T$$

乘幂法的收敛速度主要去取决于$r=\frac{|\lambda_2|}{|\lambda_1|}$,当它接近1时，收敛速度很慢，此时需要加速。

4.2.1 艾特金加速法

如果序列a_k线性收敛于a,即$\lim_{k\to\infty}\frac{a_{k+1}-a}{a_k-a}=c\neq 0$,则当$k$充分大时，有$\frac{a_{k+2}-a}{a_{k+1}-a}\approx\frac{a_{k+1}-a}{a_k-a}$，由此可解出

$$a\approx\frac{a_{k+2a_k-a_{k+1}^2}}{a_{k+2}-2a_{k+1}+a_k}=a_k-\frac{(a_{k+1-a_k})^2}{a_{k+2}-2a_{k+1}+a_k}.$$

记上式右端为$a_k^{'}$,此时序列$a_k^{'}$比a_k更快地收敛于a。

利用艾特金加速法修改乘幂法得到的加速算法:

(1)输入矩阵$A=(a_{ij})$,初始向量v_0,允许的误差ϵ, 最大迭代次数N;

(2)置$k=1$,$a_0=0, a_1=0, \lambda_0=1$;

(3)$a=\max v_k$;

(4)计算$u_k=\frac{v_k}{a}$;$v_{k+1}=Au_k$;置$\max v_{k+1}=a_2$;

(5)计算$\lambda=a_0-\frac{(a_1-a_0)^2}{a_2-2a_1+a_0}$;

(6)若$|\lambda-\lambda_0|<\epsilon$,输出$\lambda, x$.

(7)若$k<N$, 置$k+1\Rightarrow k$,$\lambda\Rightarrow\lambda_0$, $a_1\Rightarrow a_0$,$a_2\Rightarrow a_1$,转到(3); 否则输出失败信息, 停机。

例4.3 利用乘幂法的加速法求矩阵$A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$的按模最大的特征值与对应的特征向量。取$v_0 = (0,0,1)^T$.

解 取$v_0 = (0,0,1)^T$,$a_0 = 0, a_1 = 0, \lambda_0 = 1$, 则$v_1 = Av_0 = (0,-1,2)^T$,$\max\{v_1\} =$ 2,从而$a_2 = 2$,此时$\lambda = a_0 - \frac{(a_1-a_0)^2}{a_2-2a_1+a_0} = 0$,$|\lambda - \lambda_0| = 1$, 不满足$|\lambda - \lambda_0| < \epsilon$,转至(7).

取$v_1 = (0,-1,2)^T$,$u_1 = v_1/2 = (0,-0.5,1)^T$, $v_2 = Au_1 = (0.5,-2,2.5)^T$, $a_3 = \max\{v_3\} = 2.5$, $\lambda = a_1 - \frac{(a_2-a_1)^2}{a_3-2a_2+a_1} = \frac{8}{3}$, 不满足条件，转至(7).

取$v_2 = (0.5,-2,2.5)^T$, $v_3 = Au_2 = (1.2,-2.6,2.8)^T$, $a_4 = \max\{v_3\}$,此时$\lambda = a_2 - \frac{(a_3-a_2)^2}{a_4-2a_3+a_2} = 3.25$, 计算结果见书上第54页。

4.2.1 原点平移法

适当地选择p，使$\lambda_1 - p$为矩阵$A - pI$的按模最大的特征值，且使得$\frac{\lambda_2-p}{\lambda_1-p}$较$|\frac{\lambda_2}{\lambda_1}|$小得多，从而显著提高收敛速度。

例4.4 设$A \in R^{4\times 4}$有特征值$\lambda_j = 15-j, j = 1,2,3,4$,则比值$r = |\frac{\lambda_2}{\lambda_1}| \approx 0.9$.而对矩阵$B = A - 12I$, 则$B$的特征值分别为$\mu_1 = 2, \mu_2 = 1, \mu_3 = 0, \mu_4 = -1$, 应用乘幂法计算$B$的主特征值的收敛速度的比值为$|\frac{\mu_2}{\mu_1}| = 0.5 < 0.9$.

例4.5 取$p=2.9$, 用原点平移法求$A=\begin{pmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2.8 \end{pmatrix}$的主特征值和相应的特征向量，要求误差不超过$10^{-4}$.

解 由原点平移法，令

$$B=A-pI=\begin{pmatrix} -6.9 & 14 & 0 \\ -5 & 10.1 & 0 \\ -1 & 0 & -0.1 \end{pmatrix},$$

取$v_0=(1,1,1)^T$,则$a=1.u_0=(1,1,1)^T$.计算结果见书第55页。

谢谢！