第4章 逐次逼近法

4.3 计算特征值的幂法

在很多工程技术和科学计算中，经常会碰到计算矩阵的特征问题。设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$, 矩阵**A**的特征问题是求数 λ 和非零向量**x**, 使得

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}\ , \qquad \textbf{（4-32）}$$

在线性代数中曾经计算过低阶矩阵的特征问题，即由***n***次代数方程

$$f(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0 \qquad \textbf{（4-33）}$$

中，求出（4−33）的根 λ_i，即为**A**的特征值，再从线性方程组

$$(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0} \qquad \textbf{（4-34）}$$

中求出 $\boldsymbol{x}^{(i)}$, 即为 λ_i的关于的特征向量 $\boldsymbol{x}^{(i)}$。λ_i和 $\boldsymbol{x}^{(i)}$称为**A**的特征对。

4.3.1 幂法

用解非线性方程和解线性方程组的方法求矩阵特征问题，一般情况计算量很大。因此，须考虑用其他方法计算A的特征对。在不少工程实际中往往需要求矩阵A绝对值（模）最大或最小的特征值。幂法就是求这种特征值与相应的特征向量的方法。

矩阵A绝对值最大的特征值称主特征值。 设A的特征值和对应的特征向量分别为

$$\lambda_1, \lambda_2, \cdots, \lambda_n \qquad \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(n)}$$

且$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$； $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(n)}$ 构成A的线性无关特征向量组。 则对任一n维向量$\boldsymbol{v}^{(0)}$均可表示成

$$\boldsymbol{v}^{(0)} = \alpha_1 \boldsymbol{x}^{(1)} + \alpha_2 \boldsymbol{x}^{(2)} + \cdots + \alpha_n \boldsymbol{x}^{(n)} \qquad (\alpha_1 \neq 0)$$

和

$$\boldsymbol{A}\boldsymbol{v}^{(0)} = \alpha_1 \boldsymbol{A}\boldsymbol{x}^{(1)} + \alpha_2 \boldsymbol{A}\boldsymbol{x}^{(2)} + \cdots + \alpha_n \boldsymbol{A}\boldsymbol{x}^{(n)}$$

$$\boldsymbol{A}\boldsymbol{v}^{(0)} = \alpha_1 \boldsymbol{A}\boldsymbol{x}^{(1)} + \alpha_2 \boldsymbol{A}\boldsymbol{x}^{(2)} + \cdots + \alpha_n \boldsymbol{A}\boldsymbol{x}^{(n)}$$

$$= \alpha_1 \lambda_1 \boldsymbol{x}^{(1)} + \alpha_2 \lambda_2 \boldsymbol{x}^{(2)} + \cdots + \alpha_n \lambda_n$$

$$\boldsymbol{A}^2\boldsymbol{v}^{(0)} = \alpha_1 \lambda_1^2 \boldsymbol{x}^{(1)} + \alpha_2 \lambda_2^2 \boldsymbol{x}^{(2)} + \cdots + \alpha_n \lambda_n^2 \boldsymbol{x}^{(n)}$$

$$\vdots$$

$$\boldsymbol{A}^k\boldsymbol{v}^{(0)} = \alpha_1 \lambda_1^k \boldsymbol{x}^{(1)} + \alpha_2 \lambda_2^k \boldsymbol{x}^{(2)} + \cdots + \alpha_n \lambda_n^k \boldsymbol{x}^{(n)}$$

$$= \lambda_1^k \cdot \left[\alpha_1 \boldsymbol{x}^{(1)} + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{x}^{(2)} + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{x}^{(n)} \right]$$

则 $$\frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\lambda_1^k} = \alpha_1 \boldsymbol{x}^{(1)} + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{x}^{(2)} + \cdots + \alpha \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{x}^{(n)}$$

因为 $\left|\frac{\lambda_i}{\lambda_1}\right| < 1,\ (i = 2, 3, \cdots, n)$，所以

$$\lim_{k\to\infty} \frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\lambda_1^k} = \alpha_1 \boldsymbol{x}^{(1)} \qquad \textbf{（4-35）}$$

(4-35)说明 $\dfrac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\lambda_1^k}$ 随着的$\boldsymbol{k}$无限增大，趋于$\boldsymbol{A}$的主特征值对应的特征向量$\boldsymbol{x}^{(1)}$ （特征向量乘一个不为0的常数 α 特征向量后，仍为原特征值对应的特征向量）。

若令 $\boldsymbol{v}^{(k)}=\boldsymbol{A}^k\boldsymbol{v}^{(0)}$，则 $\boldsymbol{v}^{(k)}\approx\lambda_1^k\alpha_1\boldsymbol{x}^{(1)}=\left(\alpha_1\lambda_1^k\right)\boldsymbol{x}^{(1)}$

故$\boldsymbol{v}^{(k)}$也可作为λ_1对应的近似特征向量。 又因为

$$\frac{\boldsymbol{v}_i^{(k)}}{\boldsymbol{v}_i^{(k-1)}}=\frac{\left(\boldsymbol{A}^k\boldsymbol{v}^{(0)}\right)_i}{\left(\boldsymbol{A}^{k-1}\boldsymbol{v}^{(0)}\right)_i}=\frac{\left[\lambda_1^k\left(\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\left(\dfrac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\dfrac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right)\right]_i}{\left[\lambda_1^{k-1}\left(\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\left(\dfrac{\lambda_2}{\lambda_1}\right)^{k-1}\boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\dfrac{\lambda_{n2}}{\lambda_1}\right)^{k-1}\boldsymbol{x}^{(n)}\right)\right]_i}$$

$$=\lambda_1 \frac{\left[\alpha_1 \boldsymbol{x}^{(1)}+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k \boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k \boldsymbol{x}^{(n)}\right]_i}{\left[\alpha_1 \boldsymbol{x}^{(1)}+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k-1} \boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\frac{\lambda_{n2}}{\lambda_1}\right)^{k-1} \boldsymbol{x}^{(n)}\right]_i}$$

其中 $\boldsymbol{v}^{(k)}=\left(v_1^{(k)}, v_2^{(k)}, \cdots, v_n^{(k)}\right)^T$。 所以

$$\lim_{k\to\infty}\frac{\boldsymbol{v}_i^{(k)}}{\boldsymbol{v}_i^{(k-1)}}=\lambda_1 \qquad \text{(4-36)}$$

（**4-36**）表明，序列 $\frac{\boldsymbol{v}_i^{(k)}}{\boldsymbol{v}_i^{(k-1)}}(k=1,2,\cdots)$ 收敛于***A***的主特征值 λ_1，其收敛速度取决于比值 $r=\left|\frac{\lambda_2}{\lambda_1}\right|$ 的大小，***r***越小，收敛速度越快，如果***r***≈1，则收敛速度就很慢，需要采用加速技术。

综上所述，可得如下定理

定理4.8 若矩阵$\boldsymbol{A}$具有$\boldsymbol{n}$个线性无关的特征向量

$$\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(n)}$$

且对应的特征值满足$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$，则取$\boldsymbol{v}^{(0)} \neq 0$, 经使用

$$\boldsymbol{v}^{(k)} = \boldsymbol{A}^k \boldsymbol{v}^{(0)} \quad k = 1, 2, \cdots \tag{4-37}$$

迭代计算可得

1） $$\boldsymbol{v}^{(k)} = \alpha_1 \lambda_1^k \boldsymbol{x}^{(1)} \quad (\alpha_1 \neq 0) \tag{4-38}$$

2） $$\frac{\boldsymbol{v}_i^{(k)}}{\boldsymbol{v}_i^{(k-1)}} \approx \lambda_1 \tag{4-39}$$

利用（4-37）、（4-38）、（4-39）计算$\boldsymbol{A}$的主特征值 λ_1，及其对应的特征向量$\boldsymbol{x}^{(1)}$的方法，称为**幂法**。其中$\left(\lambda_1, \boldsymbol{x}^{(1)}\right)$也称**极端特征对**。

在使用幂法法时还应注意到：

1）由于在取初始向量$\boldsymbol{v}^{(k)}$时，$\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(n)}$ 事先不知，可能使 $\boldsymbol{v}^{(0)}=\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}+\cdots+\alpha_n\boldsymbol{x}^{(n)}$ 中的 $\alpha_1=0$。从理论上讲，此时必须换初始向量$\boldsymbol{v}^{(0)}$，否则算不出$\boldsymbol{x}^{(1)}$，但是在实际计算时，由于舍入误差的影响，可能经几步迭代后得到

$$\boldsymbol{v}^{(t)}=\boldsymbol{A}^t\boldsymbol{v}^{(0)}=\beta_1\boldsymbol{x}^{(1)}+\beta_2\boldsymbol{x}^{(2)}+\cdots+\beta_n\boldsymbol{x}^{(n)}$$

其中 $\beta_1\neq 0$，这样继续算下去仍可算出$\boldsymbol{x}^{(1)}$的近似值，但是迭代次数就要增加，　如果发现收敛速度很慢，可以更换初始向量$\boldsymbol{v}^{(0)}$后再进行迭代计算。

2）从（4-38）中易知，当$|\lambda_1|>1\left(\text{或}|\lambda_1|<1\right)$时，数值计算将产生“溢出”，即计算$\boldsymbol{v}^{(k)}$，当$k\to\infty$时，$\boldsymbol{v}^{(k)}$中的非零分量将趋于零或无穷大 。 为了避免这种现象的产生， 需要将由（4-38）式每次算得的向量$\boldsymbol{v}$进行规范化，即在向量上除以一个常数。

为了方便起见，除数可取该向量绝对值最大的分量，并记为 **max(*v*)**。

若 $\boldsymbol{v}=(v_1,v_2,\cdots v_n)^T$ 且 $|v_k|=\max\limits_{1\le i\le n}|v_i|$

则 $$\max(\boldsymbol{v})=v_k$$

令 $$\boldsymbol{u}=\frac{\boldsymbol{v}}{\max(\boldsymbol{v})}=\left(\frac{v_1}{v_k},\frac{v_2}{v_k},\cdots,\frac{v_n}{v_k}\right)^T$$

如此规范后的向量$\boldsymbol{v}$，其绝对值最大的分量的值为1。

于是幂法可作如下表达：

（1）取初始向量$\boldsymbol{v}^{(0)} \neq \boldsymbol{0}$,且 $\alpha_1 \neq 0$, 并令 $\boldsymbol{u}^{(0)} = \boldsymbol{v}^{(0)}$

（2） $\boldsymbol{v}^{(1)} = \boldsymbol{A}\boldsymbol{v}^{(0)} = \boldsymbol{A}\boldsymbol{u}^{(0)}$

$$\boldsymbol{u}^{(1)} = \frac{\boldsymbol{v}^{(1)}}{\max\left(\boldsymbol{v}^{(1)}\right)} = \frac{\boldsymbol{A}\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}\boldsymbol{v}^{(0)}\right)}$$

（3）$\boldsymbol{v}^{(2)} = \boldsymbol{A}\boldsymbol{u}^{(1)} = \dfrac{\boldsymbol{A}^2\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}\boldsymbol{v}^{(0)}\right)} = \dfrac{\boldsymbol{A}\boldsymbol{v}^{(1)}}{\max\left(\boldsymbol{A}\boldsymbol{v}^{(0)}\right)}$

$$\boldsymbol{u}^{(2)} = \frac{\boldsymbol{v}^{(2)}}{\max\left(\boldsymbol{v}^{(2)}\right)} = \frac{\boldsymbol{A}^2\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^2\boldsymbol{v}^{(0)}\right)}$$

（4）
$$\begin{cases} \boldsymbol{v}^{(k)} = \boldsymbol{A}\boldsymbol{u}^{(k-1)} = \dfrac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^{k-1}\boldsymbol{v}^{(0)}\right)} \\ \boldsymbol{u}^{(k)} = \dfrac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)} = \dfrac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^k\boldsymbol{v}^{(0)}\right)} \\ k = 3, 4, \cdots \end{cases} \tag{4-40}$$

因为

$$\boldsymbol{u}^{(k)}=\frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^k\boldsymbol{v}^{(0)}\right)}=\frac{\lambda_1^k\left[\alpha_1\boldsymbol{x}^{(1)}+\cdots+\alpha_n\left(\dfrac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right]}{\max\left(\lambda_1^k\left[\alpha_1\boldsymbol{x}^{(1)}+\cdots+\alpha_n\left(\dfrac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right]\right)}$$

$$=\frac{\left[\alpha_1\boldsymbol{x}^{(1)}+\alpha_n\left(\dfrac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\dfrac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right]}{\max\left(\alpha_1\boldsymbol{x}^{(1)}+\alpha_n\left(\dfrac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\dfrac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right)}$$

所以

$$\lim_{k\to\infty}\boldsymbol{u}^{(k)}=\frac{\alpha_1\boldsymbol{x}^{(1)}}{\max\left(\alpha_1\boldsymbol{x}^{(1)}\right)}=\frac{\boldsymbol{x}^{(1)}}{\max\left(\boldsymbol{x}^{(1)}\right)}$$

又因为 $$\boldsymbol{v}^{(k)}=\frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^{k-1}\boldsymbol{v}^{(0)}\right)}=\frac{\lambda_1^k\left[\alpha_1\boldsymbol{x}^{(1)}+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right]}{\lambda_1^{k-1}\max\left[\alpha_1\boldsymbol{x}^{(1)}+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k-1}\boldsymbol{x}^{(n)}\right]}$$

$$=\lambda_1\frac{\left[\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right]}{\max\left[\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k-1}\boldsymbol{x}^{(2)}+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k-1}\boldsymbol{x}^{(n)}\right]}$$

所以 $$\lim_{k\to\infty}\max\left(\boldsymbol{v}^{(k)}\right)=\lambda_1$$

故当$\boldsymbol{k}$足够大时 $$\boxed{\boldsymbol{u}^{(k)}\approx\frac{\boldsymbol{x}^{(1)}}{\max\left(\boldsymbol{x}^{(1)}\right)}}\qquad（4\text{-}41）$$

即$\boldsymbol{u}^{(k)}$为主特征值对应的特征向量的近似向量。

$$\boxed{\max\left(\boldsymbol{v}^{(k)}\right)\approx\lambda_1}\qquad（4\text{-}42）$$

$\max\left(\boldsymbol{v}^{(k)}\right)$为主特征值的近似值。

例1　用幂法计算矩阵 $\boldsymbol{A}=\begin{pmatrix}1.0 & 1.0 & 0.5\\ 1.0 & 1.0 & 0.25\\ 0.5 & 0.25 & 2.0\end{pmatrix}$ 的极端特征对，

取 $\boldsymbol{v}^{(0)}=(1,1,1)^T$。 结果见下表

k	$\left(\boldsymbol{u}^{(k)}\right)^T$（规范化向量）	$\max\left(\boldsymbol{v}^{(k)}\right)$
1	$(0.9091,0.8182,1)$	2.750000
$\vdots$	$\vdots$	$\vdots$
5	$(0.7651,0.6674,1)$	2.558792
$\vdots$	$\vdots$	$\vdots$
10	$(0.7494,0.6508,1)$	2.538003
$\vdots$	$\vdots$	$\vdots$
15	$(0.7483,0.6497,1)$	2.536626
16	$(0.7483,0.6497,1)$	2.536584
17	$(0.7483,0.6497,1)$	2.536560
18	$(0.7483,0.6497,1)$	2.536546
19	$(0.7483,0.6497,1)$	2.536537
20	$\boxed{(0.7483,0.6497,1)}$	$\boxed{2.536532}$

主特征值的计算值为：

$$\lambda_1\approx 2.536532$$

λ_1 对应的特征向量为：

$$\boldsymbol{x}^{(1)}\approx(0.7482,0.6497,1)^T$$

如果矩阵**A**的特征值不满足假设

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

此时幂法的收敛性分析就变得复杂，它可能产生如下几种情况

（1）$\lambda_1 = \lambda_2 = \cdots = \lambda_r$，且 $|\lambda_1| > |\lambda_{r+1}|$

（2）$\lambda_1 = \lambda_2 = \cdots = \lambda_t$，$\lambda_{t+1} = \lambda_{t+2} = \cdots = \lambda_r = -\lambda_1$

且 $|\lambda_1| > |\lambda_{r+1}|$

（3）$\lambda_1 = \overline{\lambda_2}$，且 $|\lambda_1| > |\lambda_3|$

对于情况（1），由于

$$\boldsymbol{A}^k\boldsymbol{v}^{(0)}=\alpha_1\lambda_1^k\boldsymbol{x}^{(1)}+\cdots+\alpha_r\lambda_1^k\boldsymbol{x}^{(r)}+\alpha_{r+1}\lambda_{r+1}^k\boldsymbol{x}^{(r+1)}+\cdots+\alpha_n\lambda_n^k\boldsymbol{x}^{(n)}$$

$$=\lambda_1^k\left[\left(\alpha_1\boldsymbol{x}^{(1)}+\cdots+\alpha_r\boldsymbol{x}^{(r)}\right)+\alpha_{r+1}\left(\frac{\lambda_{r+1}}{\lambda_1}\right)^k\boldsymbol{x}^{(k+1)}+\cdots+\alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^{(n)}\right]$$

$$\boldsymbol{u}^{(k)}=\frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^k\boldsymbol{v}^{(0)}\right)}\to\frac{\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}+\cdots+\alpha_r\boldsymbol{x}^{(r)}}{\max\left(\alpha_1\boldsymbol{x}^{(1)}+\cdots+\alpha_r\boldsymbol{x}^{(r)}\right)},$$ 当$\boldsymbol{k}\to\infty$时。

$$\boldsymbol{v}^{(k)}=\frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^{k-1}\boldsymbol{v}^{(0)}\right)}\to\lambda_1\frac{\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}+\cdots+\alpha_r\boldsymbol{x}^{(r)}}{\max\left(\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}\cdots+\alpha_r\boldsymbol{x}^{(r)}\right)}$$

则当$\boldsymbol{k}\to\infty$时， $\max\left(\boldsymbol{v}^{(k)}\right)\to\lambda_1$。

故在情况(1)，若$\boldsymbol{v}^{(0)}$的展开式中，如果 $\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}+\cdots+\alpha_r\boldsymbol{x}^{(r)}\neq 0$

则矩阵$\boldsymbol{A}$的主特征值近似于$\mathbf{max}(\boldsymbol{v}^{(k)})$， 对应的近似特征向量为

$$\frac{\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}+\cdots+\alpha_r\boldsymbol{x}^{(r)}}{\max\left(\alpha_1\boldsymbol{x}^{(1)}+\alpha_2\boldsymbol{x}^{(2)}\cdots+\alpha_r\boldsymbol{x}^{(r)}\right)}$$

情况(2)和(3)的分析较(1)要复杂，读者可以参看[5]的第二章。

THE END