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 $A\in\mathbf{R}^{n\times n}$，矩阵***A***的特征问题是求数 λ 和非零向量***x***，使得

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}\ ,\tag{4-32}$$

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

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

中，求出（**4-33**）的根 λ_i，即为***A***的特征值，再从线性方程组

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

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$\boldsymbol{A}$的特征对。在不少工程实际中往往需要求矩阵$\boldsymbol{A}$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$\boldsymbol{A}$绝对值最大的特征值称主特征值。设$\boldsymbol{A}$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$\boldsymbol{A}$的线性无关特征向量组。则对任一$\boldsymbol{n}$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

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

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

(4-35)说明 $\frac{\boldsymbol{A}^k\boldsymbol{v}^{(0)}}{\lambda_1^k}$ 随着的k无限增大，趋于A的主特征值对应的特征向量$x^{(1)}$ （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

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

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

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

$$\lim_{k\to\infty} \frac{v_i^{(k)}}{v_i^{(k-1)}} = \lambda_1 \qquad \textbf{（4-36）}$$

（4-36）表明, 序列 $\frac{v_i^{(k)}}{v_i^{(k-1)}} \ (k = 1, 2, \cdots)$ 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）$$\frac{v_i^{(k)}}{v_i^{(k-1)}}\approx\lambda_1 \tag{4-39}$$

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$\boldsymbol{v}^{(0)}$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。

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

2）从 $\boldsymbol{v}^{(k)} = \alpha_1 \lambda_1^k \boldsymbol{x}^{(1)}$ 中易知, 当$|\lambda_1|>1$(或$|\lambda_1|<1$) 时, 数值计算将产生“溢出”，即计算$\boldsymbol{v}^{(k)}$，当$\boldsymbol{k}\to\infty$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 $$\boldsymbol{u} = \frac{\boldsymbol{v}}{\max(\boldsymbol{v})} = \left(\frac{v_1}{v_k}, \frac{v_2}{v_k}, \cdots \frac{v_{k-1}}{v_k}, 1, \cdots, \frac{v_n}{v_k}\right)^T$$

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

因为

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

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

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

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

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

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

即$u^{(k)}$为主特征值对应的特征向量的近似向量。

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

$\max\left(v^{(k)}\right)$为主特征值的近似值。

计算$\boldsymbol{A}$的主特征值 λ_1，及其对应的特征向量$\boldsymbol{x}^{(1)}$的**幂法**。

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

$$\begin{cases} \boldsymbol{v}^{(k)} = \boldsymbol{A}\boldsymbol{u}^{(k-1)} = \dfrac{\boldsymbol{A}^k \boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^{k-1}\boldsymbol{v}^{(0)}\right)} \\ \boldsymbol{u}^{(k)} = \dfrac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)} = \dfrac{\boldsymbol{A}^k \boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^k \boldsymbol{v}^{(0)}\right)} \\ k = 1, 2, \cdots \end{cases}$$

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

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

$$\boxed{\max\left(\boldsymbol{v}^{(k)}\right) \approx \lambda_1}$$

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

例1 用幂法计算矩阵

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

$$\boldsymbol{A}=\begin{pmatrix}1.0 & 1.0 & 0.5\\ 1.0 & 1.0 & 0.25\\ 0.5 & 0.25 & 2.0\end{pmatrix}$$

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

的极端特征对。 取$\boldsymbol{v}^{(0)}=(1,1,1)^T$。 令 $\boldsymbol{u}^{(0)}=\boldsymbol{v}^{(0)}$

$$\boldsymbol{v}^{(1)}=\boldsymbol{A}\boldsymbol{v}^{(0)}=\begin{pmatrix}1.0 & 1.0 & 0.5\\ 1.0 & 1.0 & 0.25\\ 0.5 & 0.25 & 2.0\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}2.5\\2.25\\2.75\end{pmatrix}=\boldsymbol{A}\boldsymbol{u}^{(0)}$$

$$\max\left(\boldsymbol{v}^{(1)}\right)=v_3^{(1)}=2.7500\,,$$

$$\boldsymbol{u}^{(1)}=\frac{\boldsymbol{v}^{(1)}}{\max\left(\boldsymbol{v}^{(1)}\right)}=\frac{\boldsymbol{A}\boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}\boldsymbol{v}^{(0)}\right)}=\left(\frac{2.5}{2.75},\frac{2.25}{2.75},1\right)^T$$

$$=\left(0.9091\,,0.8182\,,1\right)^T$$

$$\boldsymbol{v}^{(2)} = \boldsymbol{A}\boldsymbol{u}^{(1)} = \begin{pmatrix} 1.0 & 1.0 & 0.5 \\ 1.0 & 1.0 & 0.25 \\ 0.5 & 0.25 & 2.0 \end{pmatrix} \begin{pmatrix} 0.9091 \\ 0.8182 \\ 1 \end{pmatrix}$$

$$= (2.2192\,, 1.9692\,, 2.6591)^T$$

$$\max\left(\boldsymbol{v}^{(2)}\right) = v_3^{(2)} = 2.6591\,,$$

$$\boldsymbol{u}^{(2)} = \frac{\boldsymbol{v}^{(2)}}{\max(\boldsymbol{v}^{(2)})} = \left(\frac{2.2192}{2.6591}, \frac{1.9692}{2.6591}, 1\right)^T$$

$$= (0.8346\,, 0.7406\,, 1)^T$$

$$\boldsymbol{v}^{(3)} = \boldsymbol{A}\boldsymbol{u}^{(2)} = \begin{pmatrix} 1.0 & 1.0 & 0.5 \\ 1.0 & 1.0 & 0.25 \\ 0.5 & 0.25 & 2.0 \end{pmatrix} \begin{pmatrix} 0.8346 \\ 0.7406 \\ 1 \end{pmatrix}$$

$$= (2.0752\,, 1.8252\,, 2.6025)^T, \qquad \max\left(\boldsymbol{v}^{(3)}\right) = v_3^{(3)} = 2.6025\,,$$

$$\max\left(\boldsymbol{v}^{(3)}\right) = v_3^{(3)} = 2.6025\,,$$

$$\boldsymbol{u}^{(3)} = \frac{\boldsymbol{v}^{(3)}}{\max\left(\boldsymbol{v}^{(3)}\right)} = \left(\frac{2.0752}{2.6025}\,, \frac{1.8252}{2.6025}\,, 1.0000\right)^T$$

$$= \left(0.7974\,, 0.7013\,, 1.0000\right)^T$$

$$\boldsymbol{v}^{(4)} = \boldsymbol{A}\boldsymbol{u}^{(3)} = \begin{pmatrix} 1.0 & 1.0 & 0.5 \\ 1.0 & 1.0 & 0.25 \\ 0.5 & 0.25 & 2.0 \end{pmatrix}\begin{pmatrix} 0.7974 \\ 0.7013 \\ 1.0000 \end{pmatrix}$$

$$= \left(1.9987\,, 1.7487\,, 2.5740\right)^T$$

$$\max\left(\boldsymbol{v}^{(4)}\right) = v_3^{(4)} = 2.5740\,,$$

$$\boldsymbol{u}^{(4)} = \frac{\boldsymbol{v}^{(4)}}{\max\left(\boldsymbol{v}^{(4)}\right)} = \left(\frac{1.9987}{2.5740}\,, \frac{1.7487}{2.5740}\,, 1.0000\right)^T$$

$$= \left(0.7765\,, 0.6794\,, 1.0000\right)^T$$

$$\boldsymbol{u}^{(4)} = (0.7765\,,\,0.6794\,,1.0000)^T$$

$$\boldsymbol{v}^{(5)} = \boldsymbol{A}\boldsymbol{u}^{(4)} = \begin{pmatrix} 1.0 & 1.0 & 0.5 \\ 1.0 & 1.0 & 0.25 \\ 0.5 & 0.25 & 2.0 \end{pmatrix}\begin{pmatrix} 0.7765 \\ 0.6794 \\ 1.0000 \end{pmatrix}$$

$$= (1.9559\,,1.7059\,,\,2.5582)^T$$

$$\max\left(\boldsymbol{v}^{(5)}\right) = v_3^{(5)} = 2.5582\,,$$

$$\boldsymbol{u}^{(5)} = \frac{\boldsymbol{v}^{(5)}}{\max(\boldsymbol{v}^{(5)})} = \left(\frac{1.9559}{2.5582}\,,\,\frac{1.7059}{2.5582}\,,1.0000\right)^T$$

$$= (0.7646\,,\,0.6668\,,1.0000)^T$$

结果见下表

k	$\left(u^{(k)}\right)^T$（规范化向量）	$\max\left(v^{(k)}\right)$
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	$\boxed{(0.7483, 0.6497, 1)}$	$\boxed{2.536532}$

$$\boldsymbol{A} = \begin{pmatrix} 1.0 & 1.0 & 0.5 \\ 1.0 & 1.0 & 0.25 \\ 0.5 & 0.25 & 2.0 \end{pmatrix}$$

主特征值的计算值为:

$$\lambda_1 \approx 2.536532$$

λ_1 对应的特征向量为:

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

验证:

$$\boldsymbol{A}\boldsymbol{x}^{(1)} = \begin{pmatrix} 1.8979 \\ 1.6479 \\ 2.5365 \end{pmatrix}$$

$$\lambda_1 \boldsymbol{x}^{(1)} = \begin{pmatrix} 1.8976 \\ 1.6480 \\ 2.5365 \end{pmatrix}$$

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 $|\lambda_1| > |\lambda_3|$

只就(1)讨论。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**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。

4.3.2 反幂法

如果对表达式

$$Ax = \lambda x$$

的两边乘A^{-1}后，得到

$$A^{-1}Ax = Ix = \lambda A^{-1}x$$

故

$$A^{-1}x = \frac{1}{\lambda}x \qquad \text{(4-43)}$$

(4-43)式说明A的绝对值最小的特征值的倒数为A^{-1}的主特征值。

因此,我们可以对A^{-1}用幂法求的绝对值最小的特征值,简称最小特征值.

计算$\boldsymbol{A}$的主特征值 λ_1，及其对应的特征向量$\boldsymbol{x}^{(1)}$的**幂法**。

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

$$\begin{cases} \boldsymbol{v}^{(k)} = \boldsymbol{A}\boldsymbol{u}^{(k-1)} = \dfrac{\boldsymbol{A}^k \boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^{k-1}\boldsymbol{v}^{(0)}\right)} \\ \boldsymbol{u}^{(k)} = \dfrac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)} = \dfrac{\boldsymbol{A}^k \boldsymbol{v}^{(0)}}{\max\left(\boldsymbol{A}^k \boldsymbol{v}^{(0)}\right)} \\ k = 1, 2, \cdots \end{cases} \tag{4-40}$$

故当k足够大时 $\boxed{\boldsymbol{u}^{(k)} \approx \dfrac{\boldsymbol{x}^{(1)}}{\max\left(\boldsymbol{x}^{(1)}\right)}}$

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

$$\boxed{\max\left(\boldsymbol{v}^{(k)}\right) \approx \lambda_1}$$

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

任取初始向量$\boldsymbol{v}^{(0)}=\boldsymbol{u}^{(0)}$，使用幂法（**4-40**），得到迭代格式

$$\begin{cases} \boldsymbol{v}^{(k)} = \boldsymbol{A}^{-1}\boldsymbol{u}^{(k-1)} \\ \boldsymbol{u}^{(k)} = \dfrac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)} \qquad k=1,2,\cdots \end{cases} \tag{4-44}$$

（**4-44**）称为反幂法，其收敛速度取决于

$$r = \frac{\left|\dfrac{1}{\lambda_{n-1}}\right|}{\left|\dfrac{1}{\lambda_n}\right|} = \left|\frac{\lambda_n}{\lambda_{n-1}}\right|$$

在反幂法（**4-44**）中需要求$\boldsymbol{A}$的逆矩阵$\boldsymbol{A}^{-1}$，为了避免求$\boldsymbol{A}^{-1}$，可以通过解线性方程组

$$\boldsymbol{A}\boldsymbol{v}^{(k)}=\boldsymbol{u}^{(k-1)}$$

的方法求出$\boldsymbol{v}^{(k)}$ 。 由于需要反复求解以为$\boldsymbol{A}$系数矩阵的方程组，所以可先对进行三角分解$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$后，求解三角方程组，即求解

$$\begin{cases}\boldsymbol{L}\boldsymbol{y}^{(k)}=\boldsymbol{v}^{(k-1)}\\ \boldsymbol{U}\boldsymbol{v}^{(k)}=\boldsymbol{y}^{(k)}\end{cases}$$

其中$\boldsymbol{L}$, $\boldsymbol{U}$分别为下、上三角阵。

取初始向量$\boldsymbol{u}^{(0)}\neq\boldsymbol{0}$，由迭代格式

$$\begin{cases}\boldsymbol{A}\boldsymbol{v}^{(k)}=\boldsymbol{L}\boldsymbol{U}\boldsymbol{v}^{(k)}=\boldsymbol{u}^{(k-1)}\\ \boldsymbol{u}^{(k)}=\dfrac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)}\quad k=1,2,\cdots\end{cases}\qquad(4\text{-}45)$$

计算出的绝对值最小的特征值 λ_n 和对应的特征向量$\boldsymbol{x}^{(n)}$, $\left(\lambda_n,\boldsymbol{x}^{(n)}\right)$ 也称极端特征对。

当非奇异矩阵***A***有***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

$$|\lambda_1| \geq |\lambda_2| \geq \cdots |\lambda_{n-1}| > |\lambda_n| > 0$$

则对于非零初始向量$\boldsymbol{v}^{(0)}=\boldsymbol{u}^{(0)}\neq\boldsymbol{0}$ $(\alpha_n \neq 0)$，由反幂法**(4-45)**求出

$\boldsymbol{v}^{(k)}, \boldsymbol{u}^{(k)}$，当$k$足够大时，

$$\boldsymbol{u}^{(k)} \approx \frac{\boldsymbol{v}^{(n)}}{\max\left(\boldsymbol{v}^{(n)}\right)}, \quad \lambda_n \approx \frac{1}{\max\left(\boldsymbol{v}^{(k)}\right)}$$

其中 $\dfrac{1}{\max\left(\boldsymbol{v}^{(k)}\right)}$和 $\boldsymbol{v}^{(k)}$ 分别为***A***的最小特征值和对应的特征向量

的近似(计算)值 。

由受到反幂法启发，如果知道矩阵**A**的某个特征值 λ_i 的近似值*p*, 或由某种方法求出特征值 λ_i 的近似值*p*, 需要提高精度,

此时若有 $\left|\lambda_i - p\right| < \left|\lambda_j - p\right| \qquad i \neq j$ （**4-46**）

由于 $$(\boldsymbol{A} - p\boldsymbol{I})\boldsymbol{x}^{(i)} = \boldsymbol{A}\boldsymbol{x}^{(i)} - p\boldsymbol{x}^{(i)} = (\lambda_i - p)\boldsymbol{x}^{(i)}$$

故若 $(\boldsymbol{A} - p\boldsymbol{I})^{-1}$ 存在，则

$$(\boldsymbol{A} - p\boldsymbol{I})^{-1}\boldsymbol{x}^{(i)} = \frac{1}{\lambda_i - p}\boldsymbol{x}^{(i)} \qquad \textbf{(4-47)}$$

从（**4-46**）和（**4-47**）知, $\dfrac{1}{\lambda_i - p}$ 是 $(\boldsymbol{A} - p\boldsymbol{I})^{-1}$ 的主特征值。

于是使用反幂法 $$\begin{cases} (\boldsymbol{A} - p\boldsymbol{I})\boldsymbol{v}^{(k)} = \boldsymbol{u}^{(k-1)} \\ \boldsymbol{u}^{(k)} = \dfrac{\boldsymbol{v}^{(k)}}{\max(\boldsymbol{v}^{(k)})} \qquad k = 1, 2, \cdots \end{cases}$$

求出$\boldsymbol{v}^{(k)}$, $\boldsymbol{u}^{(k)}$，当k足够大时,

$$\boldsymbol{u}^{(k)} \approx \frac{\boldsymbol{x}^{(i)}}{\max\left(\boldsymbol{x}^{(i)}\right)}, \qquad \max\left(\boldsymbol{v}^{(k)}\right) \approx \frac{1}{\lambda_i - P}$$

即
$$\lambda_i \approx p + \frac{1}{\max\left(\boldsymbol{v}^{(k)}\right)} \qquad \text{（4-50）}$$

（**4-49**）和（**4-50**）分别为***A***的第个***i*** 特征向量和对应的特征值的近似(计算)值， 此法为原点位移的反幂法。

与反幂法一样, 为了节省工作量, 应先对(***A−pI***)进行三角分解, 即

$$\boldsymbol{A} - p\boldsymbol{I} = \boldsymbol{L}\boldsymbol{U}$$

求$\boldsymbol{v}^{(k)}$求时, 相当于解两个三角方程组

$$\begin{cases} \boldsymbol{L}\boldsymbol{y}^{(k)} = \boldsymbol{v}^{(k-1)} \\ \boldsymbol{U}\boldsymbol{v}^{(k)} = \boldsymbol{y}^{(k)} \end{cases}$$

因此反幂法可以写成

$$\begin{cases} \boldsymbol{L}\boldsymbol{y}^{(k)} = \boldsymbol{u}^{(k-1)} \\ \boldsymbol{U}\boldsymbol{v}^{(k)} = \boldsymbol{y}^{(k)} \\ \boldsymbol{u}^{(k)} = \dfrac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)} \quad k = 1, 2, \cdots \end{cases} \qquad \text{（4-51）}$$

如果***LU***分解的计算精度不够时，可以采用列主元三角分解，即

$$\boldsymbol{PA} = \boldsymbol{LU} \quad \left(\boldsymbol{P}(\boldsymbol{A} - p\boldsymbol{I}) = \boldsymbol{LU}\right)$$

此时，应将（**4-51**）的第一式改为

$$\boldsymbol{L}\boldsymbol{y}^{(k)} = \boldsymbol{P}\boldsymbol{u}^{(k-1)}$$

其中***P***为排列阵。

练习　设 $\boldsymbol{A}=\begin{pmatrix}-1 & 2 & 1\\ 2 & -4 & 1\\ 1 & 1 & -6\end{pmatrix}$

有一近似特征值 $\tilde{\lambda}_j=-6.42$，用反幂法求对应的特征向量, 并改进近似特征值的精度。

解　对矩阵 $\boldsymbol{A}-\tilde{\lambda}_j\boldsymbol{I}=\boldsymbol{A}+6.42\boldsymbol{I}$ 作三角分解, 得

$$\boldsymbol{A}+6.42\boldsymbol{I}=\boldsymbol{LU}=\begin{pmatrix}1 & 0 & 0\\ 0.369003 & 1 & 0\\ 0.184503 & 0.375148 & 1\end{pmatrix}\begin{pmatrix}5.42 & 3 & 1\\ 0 & 1.681993 & 0.630993\\ 0 & 0 & -1.21884\times10^{-3}\end{pmatrix}$$

利用公式（**4-51**）

$$\boldsymbol{Ly}^{(k)}=\boldsymbol{u}^{(k-1)},\ \boldsymbol{Uv}^{(k)}=\boldsymbol{y}^{(k)},\ \boldsymbol{u}^{(k)}=\frac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)},\quad k=1,2,\cdots$$

取 $\boldsymbol{u}^{(0)}=(1,1,1)^T$，　则　$\boldsymbol{y}^{(1)}=(1,0.631000,0.578779)^T$,

$$\boldsymbol{v}^{(1)}=(37.764,308.38912,-820.446848)^T,\ \max\left(\boldsymbol{v}^{(1)}\right)=-820.446848$$

$$\boldsymbol{L}\boldsymbol{y}^{(k)}=\boldsymbol{u}^{(k-1)},\ \boldsymbol{U}\boldsymbol{v}^{(k)}=\boldsymbol{y}^{(k)},\ \boldsymbol{u}^{(k)}=\frac{\boldsymbol{v}^{(k)}}{\max\left(\boldsymbol{v}^{(k)}\right)},\quad k=1,2,\cdots$$

$$\boldsymbol{u}^{(1)}=\frac{\boldsymbol{v}^{(1)}}{\max\left(\boldsymbol{v}^{(1)}\right)}=\left(-0.046029,-0.375871,1\right)^T,$$

$$\boldsymbol{y}^{(2)}=\left(-0.046029,-0.358886,1.143128\right)^T,$$

$$\boldsymbol{v}^{(2)}=\left(43.279655,351.627002,-937.875765\right)^T,$$

$$\max\left(\boldsymbol{v}^{(2)}\right)=-937.875765$$

$$\boldsymbol{u}^{(2)}=\frac{\boldsymbol{v}^{(2)}}{\max\left(\boldsymbol{v}^{(2)}\right)}=\left(-0.046147,-0.374918,1\right)^T,$$

由于$\boldsymbol{u}^{(1)}$与$\boldsymbol{u}^{(2)}$对应的分量近似相等, $\left\|\boldsymbol{u}^{(2)}-\boldsymbol{u}^{(1)}\right\|_\infty=0.000953\le\frac{1}{2}\times10^{-2}$

故$\boldsymbol{A}$的特征值为：$\lambda_j=-6.42+\frac{1}{\max\left(\boldsymbol{v}^{(2)}\right)}\approx-6.42107$

对应的特征向量为：$\boldsymbol{u}^{(3)}=\left(-0.046147,-0.374918,1\right)^T$。

THE END