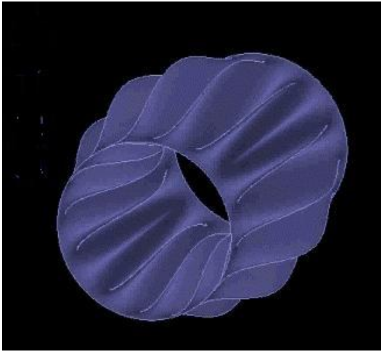
要求

1. 熟练掌握求解按模（绝对值）最大的特征值乘幂法
2. 熟练掌握求解的特征值反幂法



Definition

A 为$n \times n$ 阶矩阵，$\mathbf{x}$ 为非零向量，若存在数λ, 使$A\mathbf{x} = \lambda\mathbf{x}$ 成立，则称λ 为A 的特征值，$\mathbf{x}$ 为对应于λ 的特征向量.

Definition

若λ 满足$P_A(\lambda) = \det(\lambda I - A) = 0$, 则称$P_A(\lambda)$ 为矩阵A的特征多项式, $\det(\lambda I - A) = 0$ 称为特征方程.

Definition

若λ_i 为特征方程的k重根，称k为特征值λ_i的代数重数；齐次线性方程组$(\lambda_i I - A)\mathbf{x} = 0$的基础解系所含向量的个数（即$\lambda_i$的特征子空间的维数）称为特征值$\lambda_i$的几何重数.

矩阵特征值的基本性质

特征多项式

Property

设n 阶矩阵$A=(a_{ij})_{n\times n}$ 的全部特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$, 则有

$$\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn},$$

$$\lambda_1\lambda_2\cdots\lambda_n=\det(A).$$

Property

设λ为矩阵A的一个特征值, 则对任何多项式$P(x)=a_0+a_1x+\cdots+a_mx^m$, $P(\lambda)$为矩阵$P(A)=a_0I+a_1A+\cdots+a_mA^m$的一个特征值.

矩阵特征值的基本性质

Property

设λ为n 阶可逆矩阵A的一个特征值，则$\lambda \neq 0$，且$\frac{1}{\lambda}$为逆矩阵A^{-1}的一个特征值，$\frac{\det(A)}{\lambda}$为A的伴随矩阵A^*的一个特征值.

Property

设$\lambda_1, \lambda_2, \cdots, \lambda_m$为矩阵$A$的互不相同的特征值，$x_i$ 为A 的属于特征值λ_i的特征向量$(i = 1, 2, \cdots, m)$，则$x_1, x_2, \cdots, x_m$线性无关.

Property

矩阵A的任意特征值的几何重数不大于其代数重数.

乘幂法

乘幂法用于求模（绝对值）最大的特征值. 设A的特征值按模大小排列为$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$, 且对应的特征向量$\xi_1, \xi_2, \cdots, \xi_n$线性无关. 此时，任一非零向量$\mathbf{z}_0$均可用$\xi_1, \xi_2, \cdots, \xi_n$线性表示为

$$\mathbf{z}_0 = c_1\xi_1 + c_2\xi_2 + \cdots + c_n\xi_n,$$

作向量序列$\mathbf{z}_k = A^k\mathbf{z}_0$, 则

$$\mathbf{z}_k = \lambda_1^k\left(c_1\xi_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\xi_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\xi_n\right),$$

由此可见，若$\lambda_1 \neq 0$, 则由于$k \to \infty$时, $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0 (j = 2, \cdots, n)$, 当$k$充分大时必**有$\mathbf{z}_k \approx \lambda_1^k c_1\xi_1$, 即$\mathbf{z}_k$可近似看成$\lambda_1$对应的特征向量，而且$\mathbf{z}_k$与$\mathbf{z}_{k-1}$的非零分量之比趋近于$\lambda_1$.**

乘幂法

Algorithm

(乘幂法)

1. 任取$\mathbf{z}_0$, 如令$\mathbf{z}_0 := (1,1,\cdots,1)^T$, $\epsilon_1 > 0, \epsilon_2 > 0$
2. For $k = 1,2,\cdots,$ do
   - 令$\mathbf{y}_k := A\mathbf{z}_{k-1}$
   - 求$\mathbf{y}_k$绝对值最大的分量$m_k := \max \mathbf{y}_k$
   - 令$\mathbf{z}_k := \mathbf{y}_k / m_k$
   - 若$|m_k - m_{k-1}| \leq \epsilon_1$ 或$\|\mathbf{z}_k - \mathbf{z}_{k-1}\| \leq \epsilon_2$, 则取$\lambda_1 \approx m_k$, $\xi_1 \approx \mathbf{z}_k$, 停止计算

Theorem

乘幂法定义的序列m_k和向量序列$\mathbf{z}_k$满足

$$\lim_{k\to\infty} m_k = \lambda_1, \qquad \boxed{\lim_{k\to\infty} \mathbf{z}_k = \frac{\xi_1}{\max(\xi_1)}.}$$

特征向量

例8.1 设$A = \begin{bmatrix} -1 & 2 & 1 \\ 2 & -4 & 1 \\ 1 & 1 & -6 \end{bmatrix}$，求$A$按模最大的特征值$\lambda_1$和特征向量$\xi_1$.

解　取$\mathbf{z}_0 = (1,1,1)^T$. $\mathbf{y}_1 := A\mathbf{z}_0 = (2,-1,-4)^T$,
$m_1 := \max \mathbf{y}_1 = -4$, $\mathbf{z}_1 := \mathbf{y}_1/m_1 = (-0.5, 0.25, 1)^T$,
$\mathbf{y}_2 := A\mathbf{z}_1 = (2,-1,-6.25)^T$, $m_2 := \max \mathbf{y}_2 = -6.25$,
$\mathbf{z}_2 := \mathbf{y}_2/m_2 = (-0.32, 0.16, 1)^T$, $\cdots$
$\vdots$
$m_{90} = -6.421066614 = m_{91}$,
$\mathbf{z}_{91} = (-0.04614548303, 0.3749211313, 1)^T$
故得$\lambda_1 \approx -6.421\,066\,614$,
$\xi_1 = (-0.046\,145\,483\,03, -0.374\,921\,131\,3, 1)^T$.

乘幂法线性收敛. 渐近常数$\dfrac{\lambda_2}{\lambda_1}$称为乘幂法的收敛率, 其值越小, 收敛越快.

证明:设k充分大时, $A^k\mathbf{z}_0$绝对值最大的分量是其第i个分量, 则

$$\begin{aligned}
m_k-\lambda_1 &= \max \mathbf{y}_k-\lambda_1=\frac{\max(A^k\mathbf{z}_0)}{\max(A^{k-1}\mathbf{z}_0)}-\lambda_1\\
&= \frac{[c_1\lambda_1^k\xi_1+c_2\lambda_2^k\xi_2+\cdots+c_n\lambda_n^k\xi_n]_i}{[c_1\lambda_1^{k-1}\xi_1+c_2\lambda_2^{k-1}\xi_2+\cdots+c_n\lambda_n^{k-1}\xi_n]_i}-\lambda_1\\
&= \frac{[c_2\lambda_2^{k-1}(\lambda_2-\lambda_1)\xi_2+\cdots+c_n\lambda_n^{k-1}(\lambda_n-\lambda_1)\xi_n]_i}{[c_1\lambda_1^{k-1}\xi_1+c_2\lambda_2^{k-1}\xi_2+\cdots+c_n\lambda_n^{k-1}\xi_n]_i}\\
&= \left(\frac{\lambda_2}{\lambda_1}\right)^{k-1}M_k,\quad M_k\to M(M\text{为常数}).
\end{aligned}$$

$$\frac{|m_{k+1}-\lambda_1|}{|m_k-\lambda_1|}=\frac{|M_{k+1}(\lambda_2/\lambda_1)^k|}{|M_k(\lambda_2/\lambda_1)^{k-1}|}\to\left|\frac{\lambda_2}{\lambda_1}\right|.$$

可用Aitken方法进行加速收敛.

$$m_k^* = m_k - \frac{(m_k - m_{k-1})^2}{m_k - 2m_{k-1} + m_{k-2}}.$$

由例8.1所得数据，可得下表.

k	m_k^*	$(\mathbf{z}_k^*)_1$	$(\mathbf{z}_k^*)_2$
10	-6.423 206 251	-0.044 799 036 75	-0.377 496 088 5
20	-6.421 074 976	-0.046 140 213 23	-0.374 931 210 7
40	-6.421 066 612	-0.046 145 482 73	-0.374 921 132 1

原点位移法

乘幂法收敛快慢取决于比值$|\lambda_2/\lambda_1|$, 可选取常数p, 用$A-pI$ 代替A 作乘幂法. 适当选取p, 可使$A-pI$ 的特征值λ_i-p 满足

$$|\frac{\lambda_2-p}{\lambda_1-p}|<|\frac{\lambda_2}{\lambda_1}|$$

这时的乘幂法收敛速度快, $m_k+p\to\lambda_1$, 而$\mathbf{z}_k$ 仍收敛于A的特征向量ξ_1. 这种加速收敛的方法称为**原点位移法**.

例8.2 取$p=-2$,, $\mathbf{z}_0=(1,1,1)^T$. 按原点位移法计算例8.1可得

解 取$\mathbf{z}_0=(1,1,1)^T$. $\mathbf{y}_1:=(A+2I)\mathbf{z}_0=(4,1,-2)^T$,
$m_1:=\max\mathbf{y}_1=4$, $\mathbf{z}_1:=\mathbf{y}_1/m_1=(1,0.25,-0.5)^T$,
$\mathbf{y}_2:=(A+2I)\mathbf{z}_1=(1,1,3.25)^T$, $m_1:=\max\mathbf{y}_2=\frac{13}{4}$,
$\mathbf{z}_1:=\mathbf{y}_1/m_1=(\frac{4}{13},\frac{4}{13},1)^T$, $\cdots$

反幂法

反幂法用于求模（绝对值）最小的特征值及对应的特征向量

设A可逆，则对A^{-1} 作的乘幂法称为反幂法. 由于A^{-1} 的特征值是A 的特征值的倒数, A^{-1} 的特征向量是A 对应的特征向量，所以当A的特征值满足$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n|$ 时，按反幂法必有

$$m_k \to \frac{1}{\lambda_n}, \qquad \mathbf{z}_k \to \frac{\xi_n}{\max(\xi_n)}.$$

且收敛率为$|\lambda_n/\lambda_{n-1}|$.

反幂法用于已知矩阵近似特征值时，求矩阵的特征向量并提高特征值的精度.

已知A的特征值λ_m的近似值$\tilde{\lambda}_m$时，一般有

$$|\lambda_i - \tilde{\lambda}_m| \gg |\lambda_m - \tilde{\lambda}_m| \approx 0, \quad i \neq m$$

故对$A - \tilde{\lambda}_m I$应用反幂法时

$$m_k \to \frac{1}{\lambda_m - \tilde{\lambda}_m}, \qquad \mathbf{z}_k \to \xi_m^0.$$

而收敛率为$\dfrac{|\lambda_m - \tilde{\lambda}_m|}{|\lambda_j - \tilde{\lambda}_m|} \approx 0$，所以迭代收敛很快，往往只迭代两三步就可达到很高的精度.

反幂法

已知A的特征值近似值$\tilde{\lambda}_m$时，求矩阵的特征向量并提高特征值λ_m的精度，反幂法步骤如下：

Algorithm

(反幂法)

1. 三角分解$A-\tilde{\lambda}_m I = LR$
2. *For* $k=1,2,\cdots,$ *do*
   - 当$k=1$时，令$\mathbf{u} := (1,1,\cdots,1)^T$
     当$k\neq 1$时，解$Lu = \mathbf{z}_{k-1}$ **半次迭代法**
   - 解$R\mathbf{y}_k = u$
   - 求绝对值最大的分量$m_k := \max \mathbf{y}_k$
   - 令$\mathbf{z}_k := \mathbf{y}_k / m_k$
   - 若$|m_k - m_{k-1}| \leq \epsilon_1$ 或$\|\mathbf{z}_k - \mathbf{z}_{k-1}\| \leq \epsilon_2$，则停止计算;取$\xi_m^0 \approx \mathbf{z}_k, \lambda_m \approx \tilde{\lambda}_m + 1/m_k$.

反幂法

例8.3 用反幂法求例8.1矩阵近似于-6.42的特征值和特征向量

$$A+6.42I=\begin{pmatrix}1 & & \\ 0.369 & 1 & \\ 0.185 & 0.375 & 1\end{pmatrix}\begin{pmatrix}5.42 & 2 & 1\\ & 1.682 & 0.631\\ & & -0.0122\end{pmatrix}$$

$$Ry_1=e=(1,1,1)^T\Rightarrow y_1=(37.762,308.370,-820.410)$$

$$m_1=-820.410, LU=z_1\Rightarrow \cancel{u=}(-0.046,-0.376,1)^T$$

k	m_k	$(\mathbf{z}_k)_1$	$(\mathbf{z}_k)_2$	$(\mathbf{z}_k)_3$
1	-820.410 368 7	-0.046 028 298 46	-0.375 872 762 3	1
2	-937.834 480 8	-0.046 145 715 84	-0.374 920 570 4	1
3	-937.545 857 7	-0.046 145 482 84	-0.374 921 131 7	1

故得$\lambda_1\approx\tilde{\lambda}_1+1/m_3\approx-6.421\,066\,614$,
$\xi_1=(-0.046\,145\,482\,84,-0.374\,921\,131\,7,1)^T$.

小　结

- 乘幂法常用于求按模最大的特征值
- 带位移的反幂法用于已知特征值的近似值时，求其精确值及对应的特征向量；
- QR方法用于求一般矩阵全部特征值和特征向量
- Arnoldi方法用于求稀疏矩阵特征值
- Jacobi方法用于求实对称矩阵的全部特征值和特征向量
- Givens方法用于实对称三对角矩阵特征值的计算
- Lanczos算法用于求大规模对称稀疏矩阵的最大最小特征值
- 广义Schur分解用于求解广义特征值问题