矩阵特征值与特征向量的计算

主讲：王伟

Emal: wangw@tongji.edu.cn

同济大学数学科学学院

内容提要

内容提要

内容提要

内容提要

矩阵特征值问题:

给定n阶方阵$A \in R^{n\times n}$, 寻找常数$\lambda \in C$和非零向量$x \in C^n$, 使得

$$Ax = \lambda x,$$

其中C和C^n分别表示复数域和n维复向量空间。

定义1.1

对于n阶方阵A, 称

$$p(\lambda) = \det(A - \lambda I)$$

为矩阵A的特征多项式。

定义1.1

对于n阶方阵A，称

$$p(\lambda) = \det(A - \lambda I)$$

为矩阵A的特征多项式。

定义1.2

$p(\lambda)$是关于λ的n次多项式，且p的零点即为矩阵A的全部特征值。

定义1.1

对于n阶方阵A，称
$$p(\lambda)=\det(A-\lambda I)$$
为矩阵A的特征多项式。

定义1.2

$p(\lambda)$是关于λ的n次多项式，且p的零点即为矩阵A的全部特征值。

定义1.3

设λ_i为矩阵A的特征值，由于$\det(A-\lambda_i I)=0$，故方程$(A-\lambda_i I)x=0$必有非零解$x^{(i)}$，称$x^{(i)}$为矩阵A对应于特征值λ_i的特征向量。

§1.2 前言——特征值问题的线性代数解法

- 特征多项式与特征方程

$$\det(A-\lambda I)=0$$

- 特征多项式与特征方程

$$\det(A-\lambda I)=0$$

- 线性代数中介绍方法：

第一步 求解特征方程，得到矩阵A的所有特征值；

第二步 求解齐次方程$(A-\lambda_i I)x=0$，得到λ_i所对应的特征向量。

§1.2 前言——特征值问题的线性代数解法

- 特征多项式与特征方程

$$\det(A-\lambda I)=0$$

- 线性代数中介绍方法：

第一步 求解特征方程，得到矩阵A的所有特征值；

第二步 求解齐次方程$(A-\lambda_i I)x=0$，得到λ_i所对应的特征向量。

- 当矩阵阶数较大时，特征多项式的零点没有简单的解析表达式，只能通过近似计算得到。而高次方程近似求根的稳定性差，求得的近似解会有较大误差!

例1.1

计算矩阵$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$的全部特征值。

例1.1

计算矩阵$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}$的全部特征值。

解:矩阵A的特征多项式为

$$\begin{aligned} p(\lambda) &= \det(A-\lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 0 & 2 \\ 0 & 1-\lambda & -1 \\ -1 & 1 & 1-\lambda \end{bmatrix}\right) \\ &= (1-\lambda)(\lambda^2-2\lambda+4). \end{aligned}$$

矩阵A的特征值为$p(\lambda)=0$的解，即$\lambda_1=1$, $\lambda_{2,3}=1\pm\sqrt{3}i$。

定理1.1

若$\lambda_i(i=1,2,\cdots,n)$是矩阵$A$的特征值, 则有

(1) $\sum_{i=1}^{n}\lambda_i=\mathrm{tr}(A)$，其中$\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$表示矩阵$A$的迹；

(2) $\prod_{i=1}^{n}\lambda_i=\det(A)$。

定理1.1

若$\lambda_i(i=1,2,\cdots,n)$是矩阵$A$的特征值，则有

(1) $\sum_{i=1}^{n}\lambda_i=\mathrm{tr}(A)$，其中$\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$表示矩阵$A$的迹；

(2) $\prod_{i=1}^{n}\lambda_i=\det(A)$。

定义1.4

设A, B都是n阶矩阵，若存在可逆矩阵P，使得$P^{-1}AP=B$，则称矩阵A与矩阵B相似。

定理1.2

若矩阵A与矩阵B相似，则A与B具有相同的特征值。

定理1.3 (Gerschgorin圆盘定理)

设$A=(a_{ij})_{n\times n}$, 则A的每一个特征值必属于下述某个圆盘之中

$$|\lambda - a_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |a_{ij}|, \qquad i=1,2,\cdots,n,$$

其中上式表示复平面上以a_{ii}为中心, 以$\sum\limits_{j=1, j\neq i}^{n} |a_{ij}|$为半径的圆盘。

定理1.3 (Gerschgorin圆盘定理)

设$A=(a_{ij})_{n\times n}$, 则A的每一个特征值必属于下述某个圆盘之中

$$|\lambda - a_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |a_{ij}|, \qquad i = 1, 2, \cdots, n,$$

其中上式表示复平面上以a_{ii}为中心, 以$\sum_{j=1, j\neq i}^{n} |a_{ij}|$为半径的圆盘。

证:

定理1.3 (Gerschgorin圆盘定理)

设$A=(a_{ij})_{n\times n}$, 则A的每一个特征值必属于下述某个圆盘之中

$$|\lambda-a_{ii}|\leqslant\sum_{j=1,j\neq i}^{n}|a_{ij}|,\qquad i=1,2,\cdots,n,$$

其中上式表示复平面上以a_{ii}为中心, 以$\displaystyle\sum_{j=1,j\neq i}^{n}|a_{ij}|$为半径的圆盘。

证: 设λ是任意特征值, x是对应的特征向量, 则

$$(\lambda I-A)x=0.$$

定理1.3 (Gerschgorin圆盘定理)

设$A=(a_{ij})_{n\times n}$, 则A的每一个特征值必属于下述某个圆盘之中

$$|\lambda - a_{ii}| \leqslant \sum_{j=1,j\neq i}^{n} |a_{ij}|, \qquad i=1,2,\cdots,n,$$

其中上式表示复平面上以a_{ii}为中心, 以 $\sum\limits_{j=1,j\neq i}^{n} |a_{ij}|$为半径的圆盘。

证: 设λ是任意特征值, x是对应的特征向量, 则

$$(\lambda I - A)x = 0.$$

记$|x_i| = \max\limits_{j} |x_j| \neq 0$, 第$i$个方程为

$$-\sum_{j\neq i} a_{ij}x_j + (\lambda - a_{ii})x_i = 0,$$

定理1.3 (Gerschgorin圆盘定理)

设$A=(a_{ij})_{n\times n}$, 则A的每一个特征值必属于下述某个圆盘之中

$$|\lambda-a_{ii}|\leqslant\sum_{j=1,j\neq i}^{n}|a_{ij}|,\qquad i=1,2,\cdots,n,$$

其中上式表示复平面上以a_{ii}为中心, 以$\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|$为半径的圆盘。

证: 记$|x_i|=\max\limits_j|x_j|\neq 0$, 第$i$个方程为

$$-\sum_{j\neq i}a_{ij}x_j+(\lambda-a_{ii})x_i=0,$$

$$|\lambda-a_{ii}|\leqslant\sum_{j\neq i}|a_{ij}|\left|\frac{x_j}{x_i}\right|\leqslant\sum_{j\neq i}|a_{ij}|.$$

例1.2

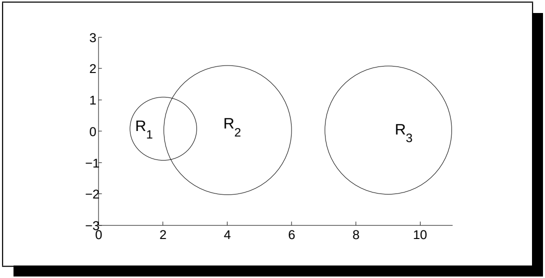
对于矩阵 $\begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix}$，圆盘定理中的圆盘为：

$$R_2 = \{\lambda \in Z : |\lambda - 4| \leqslant 2\} \qquad R_1 = \{\lambda \in Z : |\lambda - 2| \leqslant 1\}$$
$$R_3 = \{\lambda \in Z : |\lambda - 9| \leqslant 2\}$$

例1.2

对于矩阵 $\begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix}$，圆盘定理中的圆盘为：

$$R_2 = \{\lambda \in Z : |\lambda - 4| \leqslant 2\} \qquad R_1 = \{\lambda \in Z : |\lambda - 2| \leqslant 1\}$$
$$R_3 = \{\lambda \in Z : |\lambda - 9| \leqslant 2\}$$



注：
$\lambda_1 = 8.4853$
$\lambda_2 = 4.6318$
$\lambda_3 = 1.8828$

定义1.5

设A为n阶实对称矩阵，对于任意非零向量x，称$R(x)=\dfrac{(Ax,x)}{(x,x)}$为关于向量$x$的瑞雷(*Rayleigh*)商，其中$(x,y)=\sum_{i=1}^{n}x_iy_i$为$R^n$中两向量$x$，$y$的内积。

定义1.5

设A为n阶实对称矩阵，对于任意非零向量x，称$R(x)=\dfrac{(Ax,x)}{(x,x)}$为关于向量$x$的瑞雷(*Rayleigh*)商，其中$(x,y)=\sum\limits_{i=1}^{n}x_iy_i$为$R^n$中两向量$x$，$y$的内积。

定理1.4

设A为n阶实对称矩阵，$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$为其全部特征值，则

(1) $\lambda_n\leqslant\dfrac{(Ax,x)}{(x,x)}\leqslant\lambda_1,\quad\forall x\neq 0\in R^n;$

(2) $\lambda_n=\min\limits_{x\neq 0}\dfrac{(Ax,x)}{(x,x)}\qquad\lambda_1=\max\limits_{x\neq 0}\dfrac{(Ax,x)}{(x,x)}.$

证: 因为A是实对称矩阵，故有完全的特征向量系。设$x^{(1)}, x^{(2)}, \cdots, x^{(n)}$是$A$的规范化正交特征向量组，即$(x^{(i)}, x^{(j)}) = \delta_{ij}$。任一非零向量$x$都能表示为$x = \sum_{i=1}^{n} a_i x^{(i)}$，且

$$(x, x) = \left(\sum_{1 \leqslant i \leqslant n} \alpha_i x^{(i)}, \sum_{1 \leqslant i \leqslant n} \alpha_i x^{(i)} \right) = \sum_{1 \leqslant i \leqslant n} \alpha_i^2,$$

$$(Ax, x) = \left(\sum_{1 \leqslant i \leqslant n} A\alpha_i x^{(i)}, \sum_{1 \leqslant i \leqslant n} \alpha_i x^{(i)} \right) = \sum_{1 \leqslant i \leqslant n} \lambda_i \alpha_i^2.$$

故

$$\frac{(Ax, x)}{(x, x)} = \frac{\sum_{i=1}^{n} \lambda_i \alpha_i^2}{\sum_{i=1}^{n} \alpha_i^2} \leqslant \frac{\lambda_1 \sum_{i=1}^{n} \alpha_i^2}{\sum_{i=1}^{n} \alpha_i^2} = \lambda_1 \quad x \neq 0.$$

故

$$\frac{(Ax,x)}{(x,x)}=\frac{\sum\limits_{i=1}^{n}\lambda_i\alpha_i^2}{\sum\limits_{i=1}^{n}\alpha_i^2}\leqslant\frac{\lambda_1\sum\limits_{i=1}^{n}\alpha_i^2}{\sum\limits_{i=1}^{n}\alpha_i^2}=\lambda_1 \quad x\neq 0.$$

同理可证

$$\frac{(Ax,x)}{(x,x)}\geqslant\lambda_n. \quad x\neq 0.$$

在瑞雷商中，分别取$x=x^{(1)}$和$x=x^{(n)}$，可达到瑞雷商的最大值λ_1和最小值λ_n，故

$$\lambda_1=\max_{x\neq 0}\frac{(Ax,x)}{(x,x)}, \qquad \lambda_n=\min_{x\neq 0}\frac{(Ax,x)}{(x,x)}.$$

乘幂法是一种计算矩阵按模最大特征值的方法，该方法简单有效并能给出对应的特征向量。

乘幂法是一种计算矩阵按模最大特征值的方法，该方法简单有效并能给出对应的特征向量。

乘幂法需对实矩阵A做如下假设：

(1) 具有n个线性无关的特征向量$x^{(i)}, i = 1, 2, \cdots, n$;

(2) $Ax^{(i)} = \lambda_i x^{(i)}$, 其中$\lambda_i$是实特征值且满足$|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$.

乘幂法是一种计算矩阵按模最大特征值的方法，该方法简单有效并能给出对应的特征向量。

乘幂法需对实矩阵A做如下假设：

(1) 具有n个线性无关的特征向量$x^{(i)}$, $i = 1, 2, \cdots, n$;

(2) $Ax^{(i)} = \lambda_i x^{(i)}$, 其中$\lambda_i$是实特征值且满足$|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$.

算法描述（未考虑停止迭代条件）

- 选定初始向量$v^{(0)} \in R^n$;
- 利用关系式$v^{(k+1)} = Av^{(k)}$迭代计算。

算法2.1 (原始乘幂法)

(1) 给定初始非零向量$v^{(0)}$及控制参数ε;

(2) $v^{(1)} = Av^{(0)}$, $v^{(2)} = Av^{(1)}$;

(3) $\lambda_2 = \frac{1}{N_2} \sum_{j \in M_2} \frac{v_j^{(2)}}{v_j^{(1)}}$;

(4) 对$k = 2, 3, \cdots$, 做

- $v^{(k+1)} = Av^{(k)}$;
- $\lambda_{k+1} = \frac{1}{N_{k+1}} \sum_{j \in M_{k+1}} \frac{v_j^{(k+1)}}{v_j^{(k)}}$;
- $err_{k+1} = |\lambda_{k+1} - \lambda_k|$;
- 当$err_{k+1} < \varepsilon$时, 停止迭代;

(5) 返回近似按模最大特征值λ_{k+1}和近似特征向量$v^{(k+1)}$.

M_{k+1}为向量$v^{(k)}$中非零分量所对应的指标集,

N_{k+1}为向量$v^{(k)}$中非零分量的个数

例2.1

用乘幂法计算矩阵$A=\begin{bmatrix} 12 & 6 & -6 \\ 6 & 16 & 2 \\ -6 & 2 & 16 \end{bmatrix}$的按模最大特征值及其对应的特征向量。

解:取初始向量$[1.0, 0.5, -0.5]^T$和控制参数$\varepsilon = 1.0 \times 10^{-4}$。利用迭代关系式$v^{(k+1)} = Av^{(k)}$进行计算

例2.1

用乘幂法计算矩阵$A=\begin{bmatrix} 12 & 6 & -6 \\ 6 & 16 & 2 \\ -6 & 2 & 16 \end{bmatrix}$的按模最大特征值及其对应的特征向量。

k	$v^{(k)}$	λ_k	err_k
0	$[1.0, 0.5, -0.5]$	$--$	$--$
1	$[18, 13, -13]$	23.3333	$--$
2	$[372, 290, -290]$	21.7607	1.5726
3	$[7944, 6292, -6292]$	21.5826	0.1780
4	$[170832, 135752, -135752]$	21.5517	0.0309
5	$[3679008, 2925520, -2925520]$	21.5456	0.0061
6	$[79254336, 63031328, -63031328]$	21.5443	0.0013
7	$[1707427968, 135796408, -1357964608]$	21.5441	0.0003
8	$[36784710912, 2925607232, -2925607232]$	21.5440	<0.0001

例2.1

用乘幂法计算矩阵$A=\begin{bmatrix} 12 & 6 & -6 \\ 6 & 16 & 2 \\ -6 & 2 & 16 \end{bmatrix}$的按模最大特征值及其对应的特征向量。

解: 取按模最大特征值$\lambda_1 \approx 21.5440$, 对应的特征向量为

$$x^{(1)} \approx [36784710912, 2925607232, -2925607232]^T.$$

注：特征向量$x^{(1)}$的分量随着k的增大而不断增大！

- 取初始向量$v^{(0)} \in R^n$，$x^{(1)}, x^{(2)}, \cdots, x^{(n)}$线性无关，故

$$v^{(0)} = \sum_{i=1}^{n} \alpha_i x^{(i)} = \alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}.$$

§2.2 幂方法——算法分析

- 取初始向量$v^{(0)} \in R^n$，$x^{(1)}, x^{(2)}, \cdots, x^{(n)}$线性无关, 故

$$v^{(0)} = \sum_{i=1}^{n} \alpha_i x^{(i)} = \alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}.$$

- 利用$v^{(k+1)} = Av^{(k)}$计算得向量序列$\{v^{(k)}\}$.

- 取初始向量$v^{(0)} \in R^n$，$x^{(1)}, x^{(2)}, \cdots, x^{(n)}$线性无关，故

$$v^{(0)} = \sum_{i=1}^{n} \alpha_i x^{(i)} = \alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}.$$

- 利用$v^{(k+1)} = Av^{(k)}$计算得向量序列$\{v^{(k)}\}$.
- 对$\{v^{(k)}\}_{k=0}^{\infty}$，因为$Ax^{(i)} = \lambda_i x^{(i)}$，所以

$$\begin{aligned} v^{(1)} &= Av^{(0)} = \lambda_1 \alpha_1 x^{(1)} + \lambda_2 \alpha_2 x^{(2)} + \cdots + \lambda_n \alpha_n x^{(n)}, \\ v^{(2)} &= Av^{(1)} = \lambda_1^2 \alpha_1 x^{(1)} + \lambda_2^2 \alpha_2 x^{(2)} + \cdots + \lambda_n^2 \alpha_n x^{(n)}. \end{aligned}$$

- 取初始向量$v^{(0)} \in R^n$，$x^{(1)}, x^{(2)}, \cdots, x^{(n)}$线性无关, 故

$$v^{(0)} = \sum_{i=1}^{n} \alpha_i x^{(i)} = \alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}.$$

- 利用$v^{(k+1)} = Av^{(k)}$计算得向量序列$\{v^{(k)}\}$.
- 对$\{v^{(k)}\}_{k=0}^{\infty}$，因为$Ax^{(i)} = \lambda_i x^{(i)}$, 所以

$$\begin{aligned} v^{(1)} &= Av^{(0)} = \lambda_1 \alpha_1 x^{(1)} + \lambda_2 \alpha_2 x^{(2)} + \cdots + \lambda_n \alpha_n x^{(n)}, \\ v^{(2)} &= Av^{(1)} = \lambda_1^2 \alpha_1 x^{(1)} + \lambda_2^2 \alpha_2 x^{(2)} + \cdots + \lambda_n^2 \alpha_n x^{(n)}. \end{aligned}$$

- 归纳法得

$$\begin{aligned} v^{(k)} &= Av^{(k-1)} = \lambda_1^k \alpha_1 x^{(1)} + \lambda_2^k \alpha_2 x^{(2)} + \cdots + \lambda_n^k \alpha_n x^{(n)} \\ &= \lambda_1^k \left(\alpha_1 x^{(1)} + \left(\frac{\lambda_2}{\lambda_1}\right)^k \alpha_2 x^{(2)} + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k \alpha_n x^{(n)} \right). \end{aligned}$$

- 由$\left|\frac{\lambda_i}{\lambda_1}\right| < 1$, 当$k$充分大时, 所以

$$v^{(k)} \approx \lambda_1^k \alpha_1 x^{(1)}.$$

- 由$\left|\frac{\lambda_i}{\lambda_1}\right| < 1$, 当$k$充分大时, 所以

$$v^{(k)} \approx \lambda_1^k \alpha_1 x^{(1)}.$$

- 于是

$$v^{(k+1)} \approx \lambda_1^{k+1} \alpha_1 x^{(1)} \approx \lambda_1 v^{(k)}.$$

- 由 $\left|\frac{\lambda_i}{\lambda_1}\right| < 1$, 当$k$充分大时, 所以

$$v^{(k)} \approx \lambda_1^k \alpha_1 x^{(1)}.$$

- 于是

$$v^{(k+1)} \approx \lambda_1^{k+1} \alpha_1 x^{(1)} \approx \lambda_1 v^{(k)}.$$

- 注意到

$$v^{(k+1)} = Av^{(k)},$$

- 由$\left|\frac{\lambda_i}{\lambda_1}\right| < 1$, 当$k$充分大时, 所以

$$v^{(k)} \approx \lambda_1^k \alpha_1 x^{(1)}.$$

- 于是

$$v^{(k+1)} \approx \lambda_1^{k+1} \alpha_1 x^{(1)} \approx \lambda_1 v^{(k)}.$$

- 注意到

$$v^{(k+1)} = Av^{(k)},$$

- 则有

$$Av^{(k)} \approx \lambda_1 v^{(k)}.$$

§2.2 幂方法——算法分析

- 由$\left|\frac{\lambda_i}{\lambda_1}\right| < 1$, 当$k$充分大时, 所以

$$v^{(k)} \approx \lambda_1^k \alpha_1 x^{(1)}.$$

- 于是

$$v^{(k+1)} \approx \lambda_1^{k+1} \alpha_1 x^{(1)} \approx \lambda_1 v^{(k)}.$$

- 注意到

$$v^{(k+1)} = A v^{(k)},$$

- 则有

$$A v^{(k)} \approx \lambda_1 v^{(k)}.$$

- 这表明$v^{(k)}$可以近似地作为矩阵A按模最大特征值λ_1所对应的特征向量。

§2.2 幂方法——算法分析

- 由$\left|\frac{\lambda_i}{\lambda_1}\right| < 1$, 当$k$充分大时, 所以
$$v^{(k)} \approx \lambda_1^k \alpha_1 x^{(1)}.$$
- 于是
$$v^{(k+1)} \approx \lambda_1^{k+1} \alpha_1 x^{(1)} \approx \lambda_1 v^{(k)}.$$
- 注意到
$$v^{(k+1)} = A v^{(k)},$$
- 则有
$$A v^{(k)} \approx \lambda_1 v^{(k)}.$$
- 这表明$v^{(k)}$可以近似地作为矩阵A按模最大特征值λ_1所对应的特征向量。若将向量$v^{(k)}$的第j个分量记为$v_j^{(k)}$, 则
$$\lambda_1 \approx \frac{v_j^{(k+1)}}{v_j^{(k)}} \quad \text{或} \quad \lambda_1 \approx \frac{1}{n}\sum_{j=1}^{n} \frac{v_j^{(k+1)}}{v_j^{(k)}}, \qquad v_j^{(k)} \neq 0.$$

“上溢”和“下溢”现象

从上例可以看出，随着迭代步数k的增加，向量$v^{(k)}$的分量逐渐增大。事实上，当$|\lambda_1|>1$或$|\lambda_1|<1$时，向量序列$\{v^{(k)}\}$中，各个不为零的分量将会随着k的增大，而或无限增大或趋于零。

规范化

规范化的方法为用$v^{(k)}$的按模最大分量，记为$\max[v^{(k)}]$，去除所有分量，得到向量$u^{(k)}=\dfrac{v^{(k)}}{\max[v^{(k)}]}$. 利用规范化的方法可将向量$u^{(k)}$的各个分量控制在$[-1,1]$中.

“上溢”和“下溢”现象

从上例可以看出，随着迭代步数k的增加，向量$v^{(k)}$的分量逐渐增大。事实上，当$|\lambda_1| > 1$或$|\lambda_1| < 1$时，向量序列$\{v^{(k)}\}$中，各个不为零的分量将会随着k的增大，而或无限增大或趋于零。

规范化

规范化的方法为用$v^{(k)}$的按模最大分量，记为$\max[v^{(k)}]$，去除所有分量，得到向量$u^{(k)} = \dfrac{v^{(k)}}{\max[v^{(k)}]}$. 利用规范化的方法可将向量$u^{(k)}$的各个分量控制在$[-1, 1]$中.

例如，若$v^{(k)} = [-2, 3, -5, 1]^T$，则$\max[v^{(k)}] = -5$，
$u^{(k)} = [0.4, -0.6, 1, -0.2]^T$.

算法2.2 (改进乘幂法)

(1) 给定初始非零向量$v^{(0)}$及控制参数ε;

(2) λ_0为向量$v^{(0)}$的按模最大分量;

(3) $u^{(0)} = v^{(0)}/\lambda_0$;

(4) 对$k = 1, 2, \cdots$, 做

- $v^{(k+1)} = Au^{(k)}$;
- λ_{k+1}为向量$v^{(k+1)}$的按模最大分量;
- $u^{(k+1)} = v^{(k+1)}/\lambda_{k+1}$;
- $err_k = |\lambda_{k+1} - \lambda_k|$;
- 当$err_k < \varepsilon$时, 停止迭代;

(5) 返回近似按模最大特征值λ_{k+1}和近似特征向量$u^{(k+1)}$。

例2.2

用改进乘幂法计算例*2.1*中矩阵A的按模最大特征值及其对应的特征向量。

解:取初始向量$v^{(0)} = [1.0, 0.5, -0.5]^T$和控制参数$\varepsilon = 1.0 \times 10^{-4}$.
则$\lambda_0 = 1$和$u^{(0)} = v^{(0)}$.

例2.2

用改进乘幂法计算例*2.1*中矩阵A的按模最大特征值及其对应的特征向量。

k	$v^{(k)}$	λ_k	$u^{(k)}$
0	$[1.0, 0.5, -0.5]$	1.0000	$[1.0000, 0.5000, -0.5000]$
1	$[18, 13, -13]$	18.0000	$[1.0000, 0.7222, -0.7222]$
2	$[20.6667, 16.1111, -16.1111]$	20.6667	$[1.0000, 0.7796, -0.7796]$
3	$[21.3548, 16.9140, -16.9140]$	21.3548	$[1.0000, 0.7920, -0.7920]$
4	$[21.5045, 17.0886, -17.0886]$	21.5045	$[1.0000, 0.7947, -0.7947]$
5	$[21.5358, 17.1251, -17.1251]$	21.5358	$[1.0000, 0.7952, -0.7952]$
6	$[21.5423, 17.1327, -17.1327]$	21.5423	$[1.0000, 0.7953, -0.7953]$
7	$[21.5437, 17.1343, -17.1343]$	21.5437	$[1.0000, 0.7953, -0.7953]$
8	$[21.5439, 17.1346, -17.1346]$	21.5439	$[1.0000, 0.7953, -0.7953]$
9	$[21.5440, 17.1347, -17.1347]$	21.5440	$[1.0000, 0.7953, -0.7953]$

例2.2

用改进乘幂法计算例*2.1*中矩阵A的按模最大特征值及其对应的特征向量。

解: 当$k=9$时$err_9<\varepsilon$, 迭代停止。
矩阵A的按模最大特征值$\lambda_1 \approx 21.5440$,
对应的特征向量$x^{(1)} \approx [1.0000, 0.7953, -0.7953]^T$.

§2.4 幂方法——算法分析

- 取初始向量$v^{(0)}$, 由$u^{(k)} = \dfrac{v^{(k)}}{\max[v^{(k)}]}, v^{(k+1)} = Au^{(k)}$得$\{u^{(k)}\}_{k=0}^{\infty}$和$\{v^{(k)}\}_{k=0}^{\infty}$.

- 取初始向量$v^{(0)}$, 由$u^{(k)} = \frac{v^{(k)}}{\max[v^{(k)}]}, v^{(k+1)} = Au^{(k)}$

  得$\{u^{(k)}\}_{k=0}^{\infty}$和$\{v^{(k)}\}_{k=0}^{\infty}$.

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

- 取初始向量$v^{(0)}$，由$u^{(k)} = \dfrac{v^{(k)}}{\max[v^{(k)}]}, v^{(k+1)} = Au^{(k)}$
  得$\{u^{(k)}\}_{k=0}^{\infty}$和$\{v^{(k)}\}_{k=0}^{\infty}$.

$$u^{(0)}=\frac{v^{(0)}}{\max[v^{(0)}]} = \frac{\alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}}{\max[\alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}]},$$

$$v^{(1)}=Au^{(0)} = \frac{\lambda_1\alpha_1 x^{(1)} + \lambda_2\alpha_2 x^{(2)} + \cdots + \lambda_n\alpha_n x^{(n)}}{\max[\alpha_1 x^{(1)} + \alpha_2 x^{(2)} + \cdots + \alpha_n x^{(n)}]}.$$

- 取初始向量$v^{(0)}$，由$u^{(k)}=\dfrac{v^{(k)}}{\max[v^{(k)}]}, v^{(k+1)}=Au^{(k)}$
  得$\{u^{(k)}\}_{k=0}^{\infty}$和$\{v^{(k)}\}_{k=0}^{\infty}$.

$$u^{(0)}=\frac{v^{(0)}}{\max[v^{(0)}]}=\frac{\alpha_1x^{(1)}+\alpha_2x^{(2)}+\cdots+\alpha_nx^{(n)}}{\max[\alpha_1x^{(1)}+\alpha_2x^{(2)}+\cdots+\alpha_nx^{(n)}]},$$

$$v^{(1)}=Au^{(0)}=\frac{\lambda_1\alpha_1x^{(1)}+\lambda_2\alpha_2x^{(2)}+\cdots+\lambda_n\alpha_nx^{(n)}}{\max[\alpha_1x^{(1)}+\alpha_2x^{(2)}+\cdots+\alpha_nx^{(n)}]}.$$

- 则

$$u^{(k)}=\frac{\lambda_1^k\alpha_1x^{(1)}+\lambda_2^k\alpha_2x^{(2)}+\cdots+\lambda_n^k\alpha_nx^{(n)}}{\max[\lambda_1^k\alpha_1x^{(1)}+\lambda_2^k\alpha_2x^{(2)}+\cdots+\lambda_n^k\alpha_nx^{(n)}]},$$

$$v^{(k+1)}=Au^{(k)}=\frac{\lambda_1^{k+1}\alpha_1x^{(1)}+\lambda_2^{k+1}\alpha_2x^{(2)}+\cdots+\lambda_n^{k+1}\alpha_nx^{(n)}}{\max[\lambda_1^k\alpha_1x^{(1)}+\lambda_2^k\alpha_2x^{(2)}+\cdots+\lambda_n^k\alpha_nx^{(n)}]}.$$

所以,

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

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

所以,

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

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

由$\left|\frac{\lambda_i}{\lambda_1}\right|<1\ (i=2,3,\cdots,n)$,当$k$充分大时,

$$u^{(k+1)}\approx u^{(k)}\approx\frac{x^{(1)}}{\max[x^{(1)}]},\qquad \max[v^{(k+1)}]\approx\lambda_1.$$

从而

$$u^{(k+1)} = \frac{v^{(k+1)}}{\max[v^{(k+1)}]} \approx \frac{Au^{(k)}}{\lambda_1} \approx \frac{Au^{(k+1)}}{\lambda_1}.$$

则$\max[v^{(k+1)}]$可作为按模最大特征值λ_1的近似, $u^{(k+1)}$可作为对应的近似特征向量。

```
function [t,y] = eigIPower(a,xinit,ep)
   v0 = xinit;
   [tv,ti] = max(abs(v0));
   lam0 = v0(ti);
   u0 = v0/lam0;   flag = 0;
   while (flag==0)
      v1 = a*u0;
      [tv,ti] = max(abs(v1));
      lam1 = v1(ti);
      u0 = v1/lam1;
      err = abs(lam0-lam1);
      if (err<=ep)
         flag = 1;
      end
      lam0 = lam1;
   end
   t = lam1;   y = u0;
```

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A=\begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A=\begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

$$v_0=\begin{bmatrix} 1 \\ 1 \end{bmatrix},\quad u_0=\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$
$$v_1=\begin{bmatrix} 5 \\ 9 \end{bmatrix},\quad \|v_1\|_\infty=9$$

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A=\begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

$$v_1=\begin{bmatrix} 5 \\ 9 \end{bmatrix}, u_1=\begin{bmatrix} 0.5556 \\ 1.0000 \end{bmatrix},$$

$$v_2=\begin{bmatrix} 3.6667 \\ 7.2222 \end{bmatrix}, \|v_2\|_\infty=7.2222$$

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A = \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 3.6667 \\ 7.2222 \end{bmatrix}, u_2 = \begin{bmatrix} 0.5077 \\ 1.0000 \end{bmatrix},$$

$$v_3 = \begin{bmatrix} 3.5231 \\ 7.0308 \end{bmatrix}, \|v_3\|_\infty = 7.0308$$

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A = \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

$$v_3 = \begin{bmatrix} 3.5231 \\ 7.0308 \end{bmatrix}, u_3 = \begin{bmatrix} 0.5011 \\ 1.0000 \end{bmatrix},$$

$$v_4 = \begin{bmatrix} 3.5033 \\ 7.0006 \end{bmatrix}, \|v_4\|_\infty = 7.0006$$

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A = \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

$$v_4 = \begin{bmatrix} 3.5033 \\ 7.0006 \end{bmatrix}, u_4 = \begin{bmatrix} 0.5000 \\ 1.0000 \end{bmatrix},$$
$$v_5 = \begin{bmatrix} 3.5001 \\ 7.0000 \end{bmatrix}, \|v_5\|_\infty = 7.0000$$

例2.3

用改进的乘幂法求下面矩阵的按模最大的特征值和对应的特征向量，精度为10^{-4}.

$$A = \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$$

$$v_5 = \begin{bmatrix} 3.5001 \\ 7.0000 \end{bmatrix}, u_5 = \begin{bmatrix} 0.5000 \\ 1.0000 \end{bmatrix},$$

$$v_6 = \begin{bmatrix} 3.5001 \\ 7.0000 \end{bmatrix}, \|v_6\|_\infty = 7.0000$$

反幂法可用来计算非奇异矩阵的按模最小特征值及其对应的特征向量。

反幂法可用来计算非奇异矩阵的按模最小特征值及其对应的特征向量。

反幂法需对实矩阵A作如下假设:

(1) 具有n个线性无关的特征向量$x^{(i)}$;

(2) $Ax^{(i)} = \lambda_i x^{(i)}$, 其中$\lambda_i$是实特征值且满足$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots > |\lambda_n| > 0$.

反幂法可用来计算非奇异矩阵的按模最小特征值及其对应的特征向量。

反幂法需对实矩阵A作如下假设:

(1) 具有n个线性无关的特征向量$x^{(i)}$;

(2) $Ax^{(i)} = \lambda_i x^{(i)}$, 其中$\lambda_i$是实特征值且满足$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots > |\lambda_n| > 0$.

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

反幂法可用来计算非奇异矩阵的按模最小特征值及其对应的特征向量。

反幂法需对实矩阵A作如下假设:

(1) 具有n个线性无关的特征向量$x^{(i)}$;

(2) $Ax^{(i)} = \lambda_i x^{(i)}$, 其中$\lambda_i$是实特征值且满足$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots > |\lambda_n| > 0$.

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

所以A^{-1}的特征值满足

$$|\lambda_1^{-1}| \leqslant |\lambda_2^{-1}| \leqslant \cdots < |\lambda_n^{-1}|.$$

反幂法可用来计算非奇异矩阵的按模最小特征值及其对应的特征向量。

反幂法需对实矩阵A作如下假设:

(1) 具有n个线性无关的特征向量$x^{(i)}$;

(2) $Ax^{(i)} = \lambda_i x^{(i)}$, 其中$\lambda_i$是实特征值且满足$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots > |\lambda_n| > 0$.

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

所以A^{-1}的特征值满足

$$|\lambda_1^{-1}| \leqslant |\lambda_2^{-1}| \leqslant \cdots < |\lambda_n^{-1}|.$$

- 乘幂法求得矩阵A^{-1}的按摸最大特征值μ_n及特征向量$x^{(n)}$。
- 得矩阵A的按摸最小特征值$\lambda_n = \mu_n^{-1}$及特征向量$x^{(n)}$。

算法2.3 (反幂法)

(1) 给定初始非零向量$v^{(0)}$及控制参数ε;

(2) λ_0为向量$v^{(0)}$的按模最大分量;

(3) $u^{(0)} = v^{(0)}/\lambda_0$;

(4) 对$k = 1, 2, \cdots$, 做

- $v^{(k+1)} = A^{-1}u^{(k)}$;
- λ_{k+1}为向量$v^{(k+1)}$的按模最大分量;
- $u^{(k+1)} = v^{(k+1)}/\lambda_{k+1}$;
- $err_k = |\lambda_{k+1}^{-1} - \lambda_k^{-1}|$;
- 当$err_k < \varepsilon$时, 停止迭代;

(5) 返回近似按模最小特征值λ_{k+1}^{-1}和近似特征向量$u^{(k+1)}$.

算法2.3 (反幂法)

(1) 给定初始非零向量$v^{(0)}$及控制参数ε;

(2) λ_0为向量$v^{(0)}$的按模最大分量;

(3) $u^{(0)} = v^{(0)}/\lambda_0$;

(4) 对$k = 1, 2, \cdots,$ 做

- $Av^{(k+1)} = u^{(k)}$;
- λ_{k+1}为向量$v^{(k+1)}$的按模最大分量;
- $u^{(k+1)} = v^{(k+1)}/\lambda_{k+1}$;
- $err_k = |\lambda_{k+1}^{-1} - \lambda_k^{-1}|$;
- 当$err_k < \varepsilon$时, 停止迭代;

(5) 返回近似按模最小特征值λ_{k+1}^{-1}和近似特征向量$u^{(k+1)}$.

例2.4

用反幂法计算例*2.1*中矩阵A的按模最小特征值及其对应的特征向量。

解:取初始向量$v^{(0)} = [1.0, -0.5, 0.5]^T$和控制参数$\varepsilon = 1.0 \times 10^{-4}$.
则$\lambda_0 = 1$和$u^{(0)} = v^{(0)}$.

例2.4

用反幂法计算例*2.1*中矩阵A的按模最小特征值及其对应的特征向量。

解:取初始向量$v^{(0)} = [1.0, -0.5, 0.5]^T$和控制参数$\varepsilon = 1.0 \times 10^{-4}$.
则$\lambda_0 = 1$和$u^{(0)} = v^{(0)}$.

为避免矩阵求逆，线性方程组$Av^{(k+1)} = u^{(k)}$可以采用三角分解法求解。由于矩阵三角分解仅需计算一次，所以计算量可大为减少。

$$A = \begin{bmatrix} 1.0000 & 0 & 0 \\ 0.5000 & 1.0000 & 0 \\ -0.5000 & 0.3846 & 1.0000 \end{bmatrix} \begin{bmatrix} 12.0000 & 6.0000 & -6.0000 \\ 0 & 13.0000 & 5.0000 \\ 0 & 0 & 11.0769 \end{bmatrix}$$

例2.4

用反幂法计算例*2.1*中矩阵*A*的按模最小特征值及其对应的特征向量。

k	$v^{(k)}$	λ_k^{-1}	$u^{(k)}$
0	$[1.0, -0.5, 0.5]$	1	$[1, -0.5, 0.5]$
1	$[0.2083, -0.1250, 0.1250]$	4.8000	$[1.0000, -0.6000, 0.6000]$
2	$[0.2208, -0.1375, 0.1375]$	4.5283	$[1.0000, -0.6226, 0.6226]$
3	$[0.2237, -0.1403, 0.1403]$	4.4710	$[1.0000, -0.6274, 0.6274]$
4	$[0.2243, -0.1409, 0.1409]$	4.4591	$[1.0000, -0.6284, 0.6284]$
5	$[0.2244, -0.1411, 0.1411]$	4.4566	$[1.0000, -0.6286, 0.6286]$
6	$[0.2244, -0.1411, 0.1411]$	4.4561	$[1.0000, -0.6287, 0.6287]$
7	$[0.2244, -0.1411, 0.1411]$	4.4560	$[1.0000, -0.6287, 0.6287]$
8	$[0.2244, -0.1411, 0.1411]$	4.4560	$[1.0000, -0.6287, 0.6287]$

例2.4

用反幂法计算例*2.1*中矩阵A的按模最小特征值及其对应的特征向量。

解:因此, 矩阵A的按模最小特征值约为4.4561, 对应的特征向量约为$[1.0000, -0.6287, 0.6287]^T$.

§2.7 幂方法——结合原点平移的反幂法

若已知矩阵A的某特征值的粗略近似值p, 则可利用结合原点平移技巧的反幂法来计算该矩阵与数p最接近的特征值及其对应的特征向量。

若已知矩阵A的某特征值的粗略近似值p, 则可利用结合原点平移技巧的反幂法来计算该矩阵与数p最接近的特征值及其对应的特征向量。

首先作如下假设:

(1) 数p是矩阵A的第i个特征值λ_i的近似, 且满足$0 < |\lambda_i - p| < |\lambda_j - p|, \quad j \neq i.$

若已知矩阵A的某特征值的粗略近似值p, 则可利用结合原点平移技巧的反幂法来计算该矩阵与数p最接近的特征值及其对应的特征向量。

首先作如下假设:

(1) 数p是矩阵A的第i个特征值λ_i的近似, 且满足$0 < |\lambda_i - p| < |\lambda_j - p|, \quad j \neq i.$

由算法假设可知:

- 矩阵$B = A - pI$满足反幂法所要求的假设,且$\mu_i = \lambda_i - p$是矩阵B的按摸最小特征值。
- 对B应用反幂法求出μ_i及其对应特征向量$x^{(i)}$的近似值, 而矩阵A与数p最接近的特征值$\lambda_i = \mu_i + p$, 其对应的特征向量为$x^{(i)}$.

算法2.4 (带原点平移的反幂法)

(1) 给定初始非零向量$v^{(0)}$, 控制参数ε以及数p;

(2) λ_0为$v^{(0)}$中按模最大的那个分量;

(3) $u^{(0)} = v^{(0)}/\lambda_0$;;

(4) 对$k = 1, 2, \cdots,$ 做

- $(A - pI)v^{(k+1)} = u^{(k)}$;
- λ_{k+1}为向量$v^{(k+1)}$的按模最大分量;
- $u^{(k+1)} = v^{(k+1)}/\lambda_{k+1}$;
- $err_k = |\lambda_{k+1}^{-1} - \lambda_k^{-1}|$;
- 当$err_k < \varepsilon$时, 停止迭代;

(5) 返回近似特征值$\lambda_{k+1}^{-1} + p$和近似特征向量$u^{(k+1)}$.

例2.5

求矩阵$A=\begin{bmatrix}-12 & 3 & 3\\ 3 & 1 & -2\\ 3 & -2 & 7\end{bmatrix}$与数$p=-13$最接近的特征值及其对应的特征向量。

解:对矩阵$B=A-pI$进行LU分解，有

$$B=\begin{bmatrix}1 & 3 & 3\\ 3 & 14 & -2\\ 3 & -2 & 20\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 3 & 1 & 0\\ 3 & -\frac{11}{5} & 1\end{bmatrix}\begin{bmatrix}1 & 3 & 3\\ 0 & 5 & -11\\ 0 & 0 & -\frac{66}{5}\end{bmatrix}.$$

取初始向量$v^{(0)}=u^{(0)}=[1,1,1]^T$，利用结合原点平移的反幂法进行计算。

例2.5

求矩阵$A=\begin{bmatrix} -12 & 3 & 3 \\ 3 & 1 & -2 \\ 3 & -2 & 7 \end{bmatrix}$与数$p=-13$最接近的特征值及其对应的特征向量。

k	$v^{(k)}$	$u^{(k)}$	$\lambda_{k+1}^{-1}+p$
0	$[1,1,1]$	$[1,1,1]$	-12
1	$[-2.4545, 0.6667, 0.4848]$	$[1.0000, -0.2716, -0.1975]$	-13.4074
2	$[-4.5971, 1.0782, 0.7875]$	$[1.0000, -0.2345, -0.1713]$	-13.2175
3	$[-4.5409, 1.0676, 0.7793]$	$[1.0000, -0.2351, -0.1716]$	-13.2202
4	$[-4.5418, 1.0678, 0.7795]$	$[1.0000, -0.2351, -0.1716]$	-13.2202

例2.5

求矩阵$A=\begin{bmatrix} -12 & 3 & 3 \\ 3 & 1 & -2 \\ 3 & -2 & 7 \end{bmatrix}$与数$p=-13$最接近的特征值及其对应的特征向量。

解:因此，矩阵A与$p=-13$最接近的特征值约为-13.2202, 对应的特征向量约为$[1.0000, -0.2351, -0.1716]^T$.

§3 QR方法

QR方法以矩阵的QR分解为基础，是计算中小型矩阵全部特征值的最有效方法之一，且具有收敛速度快，算法稳定等特点。

QR方法的算法描述如下:

算法3.1 (QR方法)

(1) 令$A_1 = A$, 并给定控制参数ε.

(2) 对$k = 1, 2, \cdots$, 做

- 计算矩阵A_k的QR分解$A_k = Q_k R_k$;
- 计算$A_{k+1} = R_k Q_k$;
- $err_k = |\mathrm{diag}(A_{k+1} - A_k)|$;
- 当$err_k < \varepsilon$时，停止迭代并返回A_{k+1}。

这里$\mathrm{diag}(A_k)$表示由矩阵A_k对角线元素构成的列向量。

例3.1

用QR方法计算矩阵$A=\begin{bmatrix} 5 & -2 & -5 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 2 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$的全部特征值.

解:令$A_1=A$, 利用QR分解算法对矩阵A_1进行QR分解。

$$\begin{aligned} A_1 &= Q_1R_1 \\ &= \begin{bmatrix} 0.9806 & -0.0377 & -0.1923 & -0.1038 \\ 0.1961 & 0.1887 & -0.8804 & -0.4192 \\ 0.0000 & 0.9813 & 0.1761 & 0.0740 \\ 0.0000 & 0.0000 & 0.3962 & -0.8989 \end{bmatrix} \\ &\times \begin{bmatrix} 5.0992 & -1.9612 & -5.4912 & -0.3922 \\ 0.0000 & 2.0381 & 1.5852 & -2.5288 \\ 0.0000 & 0.0000 & 2.5242 & -3.2736 \\ 0.0000 & 0.0000 & 0.0000 & 0.7822 \end{bmatrix}. \end{aligned}$$

例3.1

用QR方法计算矩阵$A=\begin{bmatrix} 5 & -2 & -5 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 2 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$的全部特征值.

解:令$A_2=R_1Q_1$，利用QR分解算法对矩阵A_2进行QR分解。

$$\begin{aligned} A_2 &= Q_2R_2 \\ &= \begin{bmatrix} 0.9963 & -0.0607 & 0.0596 & 0.0142 \\ 0.0863 & 0.7011 & -0.6885 & -0.1643 \\ 0 & 0.7104 & 0.6845 & 0.1633 \\ 0 & 0 & 0.2321 & -0.9727 \end{bmatrix} \\ &\times \begin{bmatrix} 4.6327 & -5.7616 & 1.2055 & 1.3985 \\ 0 & 3.4762 & -2.4634 & 3.3559 \\ 0 & 0 & 1.3093 & 0.9518 \\ 0 & 0 & 0 & 0.9486 \end{bmatrix}. \end{aligned}$$

例3.1

用QR方法计算矩阵$A=\begin{bmatrix} 5 & -2 & -5 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 2 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$的全部特征值.

解:迭代20次后得到

$$A_{20} = \begin{bmatrix} 4.0000 & 0.7773 & 6.1847 & -0.8453 \\ 0.0000 & -0.4261 & -2.5937 & -3.4272 \\ 0 & 2.3263 & 2.4261 & 2.1161 \\ 0 & 0 & 0.0000 & -1.0000 \end{bmatrix}$$

例3.1

用QR方法计算矩阵$A=\begin{bmatrix} 5 & -2 & -5 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 2 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$的全部特征值.

解:所以A的两个特征值为4.0000和-1.0000. 其它两个特征值满足方程

$$\begin{vmatrix} -0.4261-\lambda & -2.5937 \\ 2.3263 & 2.4261-\lambda \end{vmatrix} = 0,$$

即A具有两个共轭复特征值

$$\lambda \approx 1 \pm 2\mathrm{i}.$$

事实上, 矩阵A的四个特征值分别为$-1, 4, 1 \pm 2\mathrm{i}$.

§4 Matlab简介：eig函数

对于矩阵特征值问题, Matlab软件提供了eig函数。其格式如下:

```
>> [V,D] = eig(A)
```

其中对角阵D表示矩阵A的全部特征值, 矩阵V的第i列为D中第i列对角线元素所对应的特征向量。

示例： 利用eig函数计算矩阵$A = \begin{bmatrix} 5 & -2 & -5 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 2 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$的全部特征值及其对应的特征向量.

§4 Matlab简介：eig函数

可在Matlab软件的命令窗口输入如下命令：

```
>>[V,D]=eig([5,-2,-5,-1;1,0,-3,2;0,2,2,-3;0,0,1,-2])
```

在命令窗口可看到计算结果：

```
>>[V,D]=eig([5,-2,-5,-1;1,0,-3,2;0,2,2,-3;0,0,1,-2])

V =


   0.9787    0.6602              0.6602                0.5774
   0.1596    0.2330 + 0.3884i    0.2330 - 0.3884i    -0.0000
   0.1277    0.4272 - 0.3884i    0.4272 + 0.3884i     0.5774
   0.0213    0.0388 - 0.1553i    0.0388 + 0.1553i     0.5774
D =


   4.0000         0                   0                   0
        0    1.0000 + 2.0000i         0                   0
        0         0              1.0000 - 2.0000i         0
        0         0                   0             -1.0000
>>
```