特征值与特征向量的计算

求 $Ax = \lambda x$

- 3.1.1 盖氏圆
  - 定义3.1-1 设 $A=[a_{ij}]_{n\times n}$，称由不等式 $|z-a_{ii}|\le\sum_{\substack{j=1\\j\ne i}}^{n}|a_{ij}|$ 所确定的复区域为 A 的第 i 行个盖氏圆，记为 G_i:

$$G_i=\{z:|z-a_{ii}|\le\sum_{\substack{j=1\\j\ne i}}^{n}|a_{ij}|\},\quad i=1,2,\ldots,n.$$

  - 定理3.1-1 若 λ 为 A 的特征值，则 $\lambda\in\bigcup_{i=1}^{n}G_i$

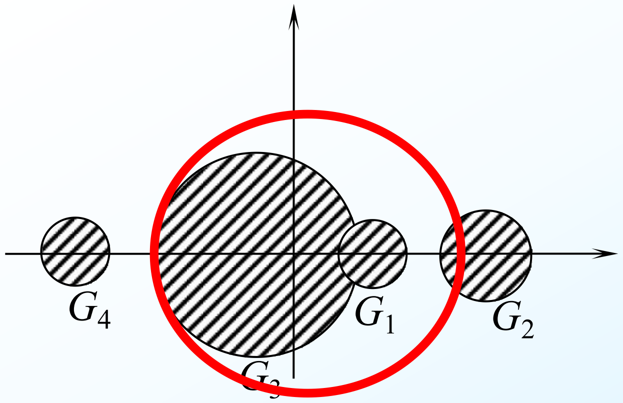
例1 估计方阵特征值的范围

$$A=\begin{bmatrix} 1 & 0.1 & 0.2 & 0.3 \\ 0.5 & 3 & 0.1 & 0.2 \\ 1 & 0.3 & -1 & 0.5 \\ 0.2 & -0.3 & -0.1 & -4 \end{bmatrix}$$

解：

$G_1=\{z: |z-1|\le 0.6\}$； $G2=\{z: |z-3|\le 0.8\}$；

$G_3=\{z: |z+1|\le 1.8\}$； $G4=\{z: |z+4|\le 0.6\}$。



注：定理称 A 的 n 个特征值全落在 n 个盖氏圆上，但未说明每个圆盘内都有一个特征值。

§1 幂法和反幂法

引例 设矩阵 $A=\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}$，用特征方程容易求得 A 的两个特征值为

$$\lambda_1=-1,\quad \lambda_2=3.$$

取初始向量 $x^{(0)}=(1,0)^T$，计算向量序列

$$x^{(k+1)}=Ax^{(k)},\quad k=0,1,\ldots$$

具体结果如下表所示：

引例

幂法计算结果

k	$x_1^{(k)}$	$x_2^{(k)}$
0	1	0
1	1	2
2	5	4
3	13	14
4	41	40
5	121	122
6	365	364
7	1093	1094

引例　　考察两个相邻向量对应分量之比：

$$\frac{x_1^{(2)}}{x_1^{(1)}} = 5 \qquad \frac{x_1^{(3)}}{x_1^{(2)}} = 2.6 \qquad \frac{x_1^{(4)}}{x_1^{(3)}} = 3.154$$

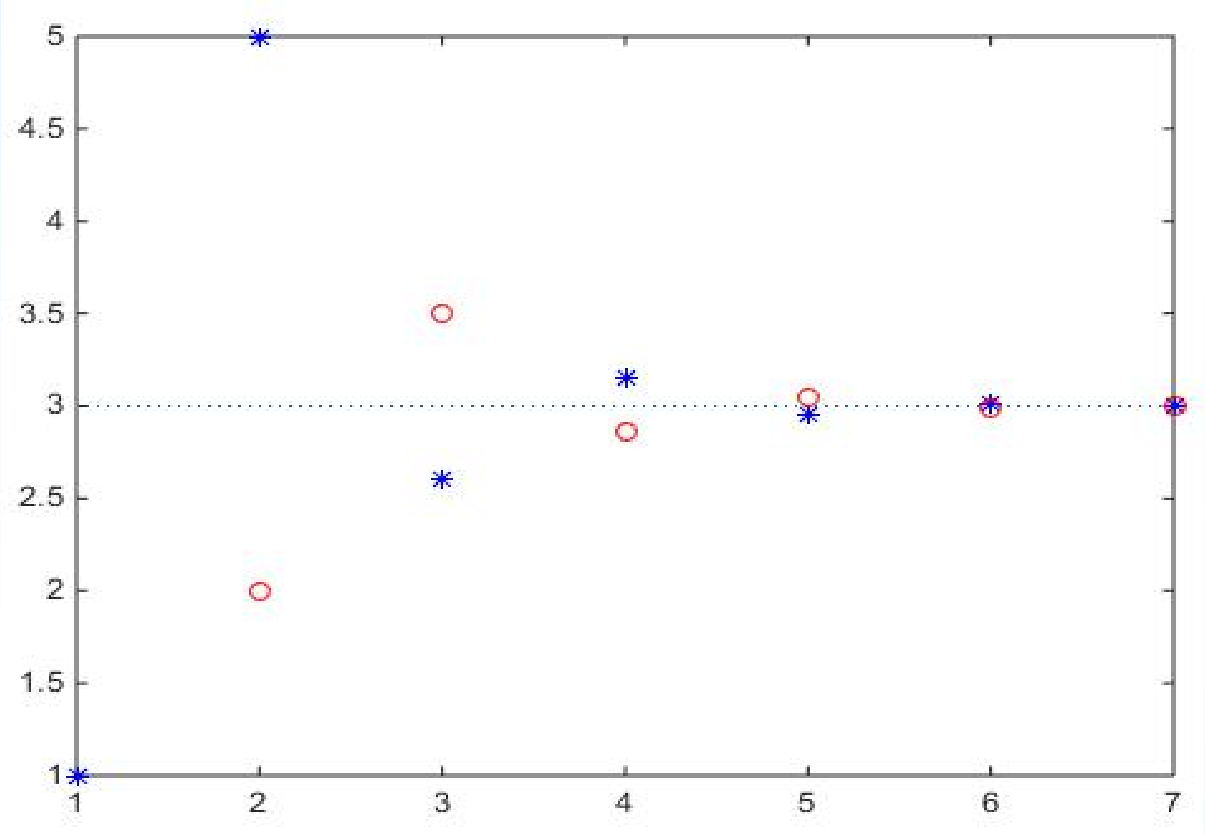
$$\frac{x_1^{(5)}}{x_1^{(4)}} = 2.951 \qquad \frac{x_1^{(6)}}{x_1^{(5)}} = 3.016 \qquad \frac{x_1^{(7)}}{x_1^{(6)}} = 2.994$$

$$\frac{x_2^{(2)}}{x_2^{(1)}} = 2 \qquad \frac{x_2^{(3)}}{x_2^{(2)}} = 3.5 \qquad \frac{x_2^{(4)}}{x_2^{(3)}} = 2.857$$

$$\frac{x_2^{(5)}}{x_2^{(4)}} = 3.05 \qquad \frac{x_2^{(6)}}{x_2^{(5)}} = 2.983 \qquad \frac{x_2^{(7)}}{x_2^{(6)}} = 3.005$$

引例

考察两个相邻向量对应分量之比：



矩阵 $A=\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}$ 的特征值为 -1 和 3.

§1 幂法和反幂法

1.1 幂法

用于求矩阵的按模最大的特征值与相应的特征向量的近似值。

设 A 为 n 阶实矩阵, λ_i , u_i $(i=1,2,\cdots,n)$ 为 A 的特征值和相应的特征向量,

且满足: $|\lambda_1|>|\lambda_2|\geq|\lambda_3|\geq\cdots\geq|\lambda_n|$

$u_1,u_2,\cdots,u_n$,线性无关.

对任意向量 $x^{(0)}$，有$x^{(0)}=\sum_{i=1}^{n}\alpha_i u_i$，$\alpha_i$不全为零 .

$$x^{(k+1)} = Ax^{(k)} = A^{k+1}x^{(0)}$$

$$= \sum_{i=1}^{n} A^{k+1}\alpha_i u_i = \sum_{i=1}^{n}\alpha_i \lambda_i^{k+1} u_i$$

$$= \lambda_1^{k+1}\left[\alpha_1 u_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^{k+1} a_2 u_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^{k+1} a_n u_n\right]$$

$$\approx \lambda_1^{k+1}\alpha_1 u_1$$

定理：设 $A \in R^{n\times n}$，特征值 $\lambda_i\ (i=1,2,\cdots n)$ 满足

$$|\lambda_1| > |\lambda_2| \ge |\lambda_3| \ge \cdots \ge |\lambda_n|,$$

且与 λ_i 对应的特征向量 $u_1, u_2, \cdots u_n$ 线性无关，则对任意非零初始向量 $x^{(0)}(\alpha_1 \ne 0)$, 向量序列

$$x^{(k)} = A^k x^{(0)} \to \lambda_1^k \alpha_1 u_1, \qquad \frac{x_i^{(k+1)}}{x_i^{(k)}} \to \lambda_1 (k \to \infty).$$

相应的特征向量为 $x^{(k+1)}$.

注：$x^{(k+1)} \approx \lambda_1^{k+1} \alpha_1 u_1$, 实际计算时将 $x^{(k+1)}$ 标准化。

标准化： 设与 λ_1 对应的特征向量 u_1

若 $x=(x_1,x_2,\cdots x_n)$, $|x_r|=\max\limits_{1\le i\le n}(|x_i|)$, 记 $x_r=\max(x)$

取初始向量 $x^{(0)}$,将 $x^{(0)}$ 标准化为 $y^{(0)}$

$$\begin{cases} x^{(k)} = Ay^{(k-1)} \\ y^{(k)} = \dfrac{x^{(k)}}{\max(x^{(k)})} \end{cases},\quad k=1,2,\cdots$$

则 $y^{(k)} \to \dfrac{u_1}{\max(u_1)}$, $\quad \max(x^{(k)}) \to \lambda_1 (k\to\infty)$

算法4.1

1. 输入 A, 初始向量 x, 误差限 ε, 最大迭代次数 N;

2. 置 $k=1$, $\lambda=\max(x)$; $y=\dfrac{x}{\lambda}$;

3. 计算 $x=Ay$, $\beta=\max(x)$, $y=\dfrac{x}{\beta}$;

4. 若 $|\lambda-\beta| < \varepsilon$, 输出 β, y, 停机; 否则, 转5;

5. 若 $k < N$, 置 $k=k+1, \lambda=\beta$, 转3, 否则输出失败信息, 停机.

幂法适用范围：

求大型稀疏矩阵按模最大的特征值λ_1 .

矩阵的特征值分布为, $|\lambda_1| > |\lambda_{m+1}| \geq \cdots \geq |\lambda_n|$

λ_1 为单根或重根, 即 $\lambda_1 = \cdots \lambda_m$,

矩阵有n个线性无关的特征向量.

例: 用幂法求矩阵A按模最大特征值和相应的特征向量 ($\varepsilon=10^{-4}$)

$$A=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

$$\begin{cases} m_k=\max\limits_i |x_i^{(k)}|=\left\|x^{(k)}\right\|, \\ y_i^{(k)}=x_i^{(k)}/m_k, \\ x^{(k+1)}=Ay^{(k)}. \end{cases}$$

k	0	1	2	3	4	5	6
$x_i^{(k)}$	1.0000 1.0000 1.0000	1.0000 0.0000 1.0000	2.0000 -2.0000 2.0000	3.0000 -4.0000 3.0000	2.5000 -3.5000 2.5000	2.4286 -3.1286 -2.4286	2.4167 -3.4167 2.4167
m_k	1.0000	1.000	2.0000	4.0000	3.5000	3.428570	3.4167
$y_i^{(k)}$	1.0000 1.0000 1.0000	1.0000 1.0000 1.0000	1.0000 -1.0000 1.0000	0.7500 -1.0000 0.7500	0.7143 -1.0000 0.7143	0.7083 -1.0000 0.7083	0.7073 -1.0000 0.7073

矩阵的最大特征值为 $\lambda_1=2+\sqrt{2}$ 其对应的特征向量 $u_1=(\frac{\sqrt{2}}{2},-1,\frac{\sqrt{2}}{2})^T$

幂法的收敛速度取决于比值 $r=\left|\frac{\lambda_2}{\lambda_1}\right|$, 比值越小, 收敛越快.

1.2 幂法的加速

（一）原点移位法

λ_i 是 A 的特征值，则 $\lambda_i-\lambda_0$ 是 $A-\lambda_0 I$ 的特征值

$$x^{(k+1)} = (A-\lambda_0 I)x^{(k)}$$

$$= (\lambda_1-\lambda_0)^{k+1}[\alpha_1 u_1+(\frac{\lambda_2-\lambda_0}{\lambda_1-\lambda_0})^{k+1}\alpha_2 u_2+\cdots(\frac{\lambda_n-\lambda_0}{\lambda_1-\lambda_0})^{k+1}\alpha_n u_n]$$

设 A 有特征值 λ_i，且 $|\lambda_1| > |\lambda_2| > \cdots$，取 λ_0 使得

$$|\lambda_1 - \lambda_0| > |\lambda_i - \lambda_0| \quad \text{且} \quad \frac{\max\limits_{i \neq 1} |\lambda_i - \lambda_0|}{|\lambda_1 - \lambda_0|} < \left|\frac{\lambda_2}{\lambda_1}\right|$$

用幂法求 $A - \lambda_0 I$ 的按模最大的特征值 λ_1^*，则 $\lambda_1 = \lambda_1^* + \lambda_0$，这种方法称为原点移位法.

注：实际应用时，A的特征值不知道，λ 无法确定，当收敛速度慢时，可以适当移动原点.

例：计算矩阵 $A=\begin{bmatrix}1.0 & 1.0 & 0.5\\1.0 & 1.0 & 0.25\\0.5 & 0.25 & 2.0\end{bmatrix}$ 按模取最大的特征值

先用规范化的幂法计算：

k	$y^{(k)}=x^{(k)}/\max(x^{(k)})$	$\lambda_1\approx\max(x^{(k)})$
0	(1， 1， 1)	
1	$(0.9091,\ 0.8182,\ 1)^T$	2.75
19	$(0.7482,\ 0.6497,\ 1)^T$	2.5365374
20	$(0.7482,\ 0.6497,\ 1)^T$	2.5365323

采用原点位移的加速法求解,取 $\lambda_0=0.75$, 矩阵 $B=A-\lambda_0 I$

k	$y^{(k)}=x^{(k)}/\max(x^{(k)})$	$\lambda_1\approx\max(x^{(k)})$
0	$(1,\ 1,\ 1)^T$	
9	$(0.7483,\ 0.6497,\ 1)^T$	1.7866587
10	$(0.7483,\ 0.6497,\ 1)^T$	1.7865914

由以上两表得到：$\lambda_1=\mu_1+\lambda_0\approx 2.5365914$

(二)幂法的埃特肯（Aitken）加速

若$\{a_k\}$收敛与a，且$\lim\limits_{k\to\infty}\dfrac{a_{k+1}-a}{a_k-a}=c\neq 0$ 即$\{a_k\}$线性收敛，

当k充分大时,有 $\dfrac{a_{k+1}-a}{a_k-a}\approx\dfrac{a_{k+2}-a}{a_{k+1}-a}$

$$\boxed{y_{n+2}=x_{n+2}-\frac{(x_{n+2}-x_{n+1})^2}{x_{n+2}-2x_{n+1}+x_n}}$$

$$\Rightarrow\quad a\approx a_k-\frac{(a_{k+1}-a_k)^2}{a_{k+2}-2a_{k+1}+a_k}:=\hat{a}_k$$

用$\hat{a}_k$ 逼近a,这种方法称为 *Aitken* 加速法.

算法4.2

1. 输入$A=(a_{ij})$,初始向量x,误差限ε,最大迭代次数N,

2. 置$k=1,\ \alpha_0=0,\ \alpha_1=0,\ \lambda_0=1.0,\ y=\dfrac{x}{\max(x)}$,

3. 计算 $x=Ay$, 置$\max(x)\Rightarrow\alpha_2$, $\dfrac{x}{\max(x)}\Rightarrow y$,

4. 计算 $\lambda=\alpha_0-\dfrac{(\alpha_1-\alpha_0)^2}{\alpha_2-2\alpha_1+\alpha_0}$,

5. 若$|\lambda-\lambda_0|<\varepsilon$,输出$\lambda,y$停机,否则转6,

6. 若$k<N$,置$\alpha_1\Rightarrow\alpha_0$, $\alpha_2=\alpha_1$, $\lambda\Rightarrow\lambda_0$, $k+1\Rightarrow k$,转3,否则,输出失败信息,停机.

(三) 对称矩阵的Rayleigh商加速法

- 定义： 设A对称，$x \neq 0$，则称 $R(x)=\dfrac{x^T Ax}{x^T x}$ 为 x 关于A的Rayleigh商。

$$R(y^{(k)})=\frac{(A^k x^{(0)})^T A^{k+1} x^{(0)}}{(A^k x^{(0)})^T A^k x^{(0)}}=\frac{\sum_{j=1}^{n}\alpha_j^2\lambda_j^{2k+1}}{\sum_{j=1}^{n}\alpha_j^2\lambda_j^{2k}}\approx\lambda_1+O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right)$$

• 公式

$$\begin{cases} y^{(k)} = \dfrac{x^{(k)}}{\max(x^{(k)})} = \dfrac{A^k x^{(0)}}{\max(A^k x^{(0)})} \\ x^{(k+1)} = Ay^{(k)} = \dfrac{A^{k+1} x^{(0)}}{\max(A^k x^{(0)})} \\ R(y^{(k)}) = \dfrac{(y^{(k)})^T A y^{(k)}}{(y^{(k)})^T (y^{(k)})} = \dfrac{(A^k x^{(0)})^T A^{k+1} x^{(0)}}{(A^k x^{(0)})^T A^k x^{(0)}} \end{cases}$$

称为Rayleigh商加速法。

此外 $R(y^{(k)}) \to \lambda_1, \quad y^{(k)} \to \dfrac{v_1}{\max(v_1)}$

• 注：有了$R(x^{(k)})$，$R(x^{(k+1)})$，$R(x^{(k+2)})$，的值，可再用**Aitken**加速法得到的一个更好的近似值.

因为

$$\frac{R(x^{(k+2)})-\lambda_1}{R(x^{(k+1)})-\lambda_1}\approx\frac{R(x^{(k+1)})-\lambda_1}{R(x^{(k)})-\lambda_1}$$

所以

$$\lambda_1\approx R(x^{(k+2)})-\frac{[R(x^{(k+2)})-R(x^{(k+1)})]^2}{R(x^{(k+2)})-2R(x^{(k+1)})+R(x^{(k)})}=\lambda_1^{(k+2)}$$

例：设$A=\begin{pmatrix}1 & 2 & 1\\ 2 & 3 & 1\\ 1 & 1 & 3\end{pmatrix}$，用瑞利商加速法求 A 的最大模特征值及特征向量，并与幂法相比较.

解：编程计算结果如下：

原始的幂法				Aitken 加速法		Rayleigh 加速法	
1	5.1667	6	5.1334	1	5.1613	1	5.1299
2	5.1613	7	5.1330	2	5.1690	2	5.1322
3	5.1437	8	5.1328	3	5.1336	3	5.1326
4	5.1373	9	5.1327	4	5.1326	4	5.1326
5	5.1346			5	5.1326		

反幂法：

基本思想：$Ax=\lambda x \Rightarrow x=A^{-1}(\lambda x)$，则$A^{-1}x=\dfrac{1}{\lambda}x$

(1)A与A^{-1}的特征值互为倒数,特征向量不变,求A的按模最小的特征值λ_n

$\Leftrightarrow$ 求A^{-1}的按模最大的特征值$\dfrac{1}{\lambda_n}$.

(2)计算$x^{(k+1)}=A^{-1}y^{(k)} \Leftrightarrow$ 解方程组$Ax^{(k+1)}=y^{(k)}$

算法4.3（反幂法）

1.输入矩阵A,近似值λ^*,初始向量x,误差限ε,最大迭代次数N,

2.置$k=1, \lambda_0=1, y=\dfrac{x}{\max(x)}$,

3.作三角分解$(A-\lambda^* I)=LU$

4.解方程组$LUx=y(Lz=y, Ux=z)$,

5.$\mu=\max(x), y=\dfrac{x}{\max(x)}, \lambda=\lambda^*+\dfrac{1}{\mu}$

6.若$|\lambda-\lambda_0|<\varepsilon$,输出$\lambda, y$,停机,否则转7,

7.若$k<N$,置$k+1\Rightarrow k, \lambda\Rightarrow\lambda_0$,转4;否则输出失败信息,停机.

例：用反幂法求矩阵$A=\begin{bmatrix}2&-1&0\\0&2&-1\\0&-1&2\end{bmatrix}$求按模最小的特征值及特征向量，取$x^{(0)}=(0,0,1)^T$；

解：求A按模最小的特征值及其特征向量

用反幂法

$$x^{(k+1)}=A^{-1}y^{(k)} \quad \Leftrightarrow \quad Ax^{(k+1)}=y^{(k)}$$

将A进行LU分解，

$$A=\begin{bmatrix}1&0&0\\0&1&0\\0&-\frac{1}{2}&1\end{bmatrix}\begin{bmatrix}2&-1&0\\0&2&-1\\0&0&\frac{3}{2}\end{bmatrix}=LU$$

$$y^{(0)} = x^{(0)} = (0,0,1)^T$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & \frac{3}{2} \end{bmatrix}$$

$$x^{(1)} = A^{-1} y^{(0)} \Leftrightarrow LUx^{(1)} = y^{(0)}$$

$$Lz = y^{(0)} \Rightarrow z = (0,0,1)^T$$

$$Ux^{(1)} = z \Rightarrow x^{(1)} = (\frac{1}{6}, \frac{1}{3}, \frac{2}{3})^T \qquad \mu = \frac{2}{3} \quad \lambda^{(1)} = \frac{1}{\mu} = 1.5 \quad y^{(1)} = (\frac{1}{4}, \frac{1}{2}, 1)^T$$

$$Lz = y^{(1)} \Rightarrow z = (\frac{1}{4}, \frac{1}{2}, \frac{5}{4})^T$$

$$Ux^{(2)} = z \Rightarrow x^{(2)} = (\frac{11}{24}, \frac{2}{3}, \frac{5}{6})^T \qquad \mu = \frac{5}{6} \qquad \lambda^{(2)} = \frac{1}{\mu} = 1.2 \quad y^{(2)} = (\frac{11}{20}, \frac{4}{5}, 1)^T$$

k	lambda	x1	x2	x3	err
1.0000	1.5000	0.2500	0.5000	1.0000	1.5000
2.0000	1.2000	0.5500	0.8000	1.0000	0.3000
3.0000	1.0714	0.7589	0.9286	1.0000	0.1286
4.0000	1.0244	0.8765	0.9756	1.0000	0.0470
5.0000	1.0082	0.9378	0.9918	1.0000	0.0162
6.0000	1.0027	0.9688	0.9973	1.0000	0.0055
7.0000	1.0009	0.9844	0.9991	1.0000	0.0018
8.0000	1.0003	0.9922	0.9997	1.0000	0.0006

反幂法求在 $\tilde{\lambda}$ 附近的特征值:

由于$A-\tilde{\lambda}I$的最小特征值是与$\tilde{\lambda}$最为接近的特征值. 若假设 $A-\tilde{\lambda}I$ 的最小特征值为μ，则与其最为接近的特征值近似值为 $\tilde{\lambda}+\mu$.设与$\tilde{\lambda}$最接近的特征值为λ_i

$$|\lambda_i-\tilde{\lambda}| \ll |\lambda_j-\tilde{\lambda}|, j\neq i, j=1,\cdots,n$$

作矩阵 $A-\tilde{\lambda}I$,它的特征值及相应特征向量为:

$$\mu_j=\lambda_j-\tilde{\lambda}\text{和}u_j, j=1,\cdots,n$$

若用反幂法求 $A-\tilde{\lambda}I$ 的最小特征值，则有如下计算关系:

$$(A-\tilde{\lambda}I)x_{k+1}=x_k, k=0,1,\cdots$$

于是得矩阵A在 $\tilde{\lambda}$ 附近特征值和相应特征向量为:

$$\begin{cases} \lambda_i = \tilde{\lambda} + \mu_i \approx \tilde{\lambda} + \dfrac{x_i^{(k)}}{x_i^{(k+1)}}, \\ u_i \approx x_{k+1}. \end{cases}$$

引例： 用反幂法求矩阵

$$A=\begin{pmatrix}2 & -2 & 0\\ 0 & 2 & -1\\ 0 & -1 & 2\end{pmatrix}$$

在2.93附近的特征值及相应特征向量（精度$\varepsilon=10^{-5}$）.

解： 对矩阵 $A-2.93I$ 作三角分解:

$$A-2.93I=\begin{pmatrix}-0.93 & -1 & 0\\ 0 & -0.93 & -1\\ 0 & -1 & -0.93\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & \dfrac{1}{0.93} & 1\end{pmatrix}\begin{pmatrix}-0.93 & -1 & 0\\ 0 & -0.93 & -1\\ 0 & 0 & -0.93+\dfrac{1}{0.93}\end{pmatrix}.$$

取 $x_0=(0,0,1)^T$，由迭代公式计算结果如下表：

表 4.8 反幂法的计算

k	0	1	2	3
$x_i^{(k)}$	0.00000	7.95906	12.69231	14.27843
	1.00000	−7.40192	−12.80385	−14.26763
	1.00000	6.88379	12.83758	14.26627
m_k	1.00000	7.95906	12.83758	14.26627
$y_i^{(k)}$	0.00000	1.00000	0.98868	1.00000
	0.00000	−0.93000	−0.99737	−0.99924
	1.00000	0.86490	1.00000	0.99915
$z_i^{(k)}$	0.00000	1.00000	0.98868	1.00000
	0.00000	−0.93000	−0.99737	−0.99924
	1.00000	1.86490	2.07244	2.07360

求矩阵全部特征值的QR方法

60年代出现的QR算法是目前计算中小型矩阵的全部特征值与特征向量的最有效方法。 理论依据：任一非奇异实矩阵都可分解成一个正交矩阵Q和一个上三角矩阵R的乘积，而且当R的对角元符号取定时，分解是唯一的。

QR方法的基本思想是利用矩阵的QR分解通过迭代格式

$$\begin{cases} A_k = Q_k R_k \\ A_{k+1} = R_k Q_k \end{cases} \qquad (k = 1, 2, \cdots).$$

将$A=A_1$化成相似的上三角阵（或分块上三角阵)，从而求出矩阵A的全部特征值与特征向量。

由 $A=A_1=Q_1R_1$，即 $Q_1^{-1}A=R_1$。

于是 $A_2=R_1Q_1=Q^{-1}AQ_1$，即 A_2 与 A 相似。

同理可得，$A_k \sim A \ (k=2,3,\cdots)$。

故它们有相同的特征值。

可证，在一定条件下，基本QR方法产生的矩阵序列 $\{A_k\}$ "基本" 收敛于一个上三角阵（或分块上三角阵）。即主对角线（或主对角线子块）及其以下元素均收敛，主对角线（或主对角线子块）以上元素可以不收敛。

特别的，如果A是实对称阵，则 $\{A_k\}$ "基本" 收敛于对角矩阵。

§2　吉文斯(Givens)旋转变换

正交变换，进行矩阵对角化，求解一般矩阵所有特征值。

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

例　设旋转矩阵R为

$$R=\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$$B = RAR^T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \triangleq \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

比较系数，有

$$b_{11} = a_{11}\cos^2\theta + a_{22}\sin^2\theta + a_{12}\sin 2\theta$$

$$b_{12} = b_{21} = \frac{1}{2}(a_{22} - a_{11})\sin 2\theta + a_{12}\cos 2\theta$$

$$b_{22} = a_{11}\sin^2\theta + a_{22}\cos^2\theta - a_{12}\cos 2\theta$$

为使A的相似矩阵B成为对角阵，只须适当选取θ，使

$$b_{12} = b_{21} = \frac{1}{2}(a_{22} - a_{11})\sin 2\theta + a_{12}\cos 2\theta = 0$$

若 $a_{11}=a_{22}$ 时，取 $|\theta|\le\dfrac{\pi}{4}$；

若 $a_{11}\ne a_{22}$，则有 $\mathrm{tg}2\theta=\dfrac{2a_{12}}{a_{11}-a_{22}}$，此时 $|\theta|\le\dfrac{\pi}{4}$.

从而旋转矩阵R确定. A的特征值为：

$$\lambda_1=b_{11},\qquad \lambda_2=b_{22}$$

计算可得对应于λ_1，λ_2 的特征向量是

$$x_1=(\cos\theta,\sin\theta)^T,\ x_2=(-\sin\theta,\cos\theta)^T$$

n阶矩阵的吉文斯旋转变换：给定

$$
P(i,j)=\begin{array}{c}
\begin{pmatrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& \cos\theta &&&& \sin\theta &&& \\
&&&& 1 &&&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
&&& -\sin\theta &&&& \cos\theta &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{pmatrix} \\
\begin{matrix} \qquad\qquad\quad i & \qquad\qquad\qquad j \qquad\qquad\qquad \end{matrix}
\end{array}
$$

一般的吉文斯旋转变换矩阵：

$$
\begin{cases}
P_k(p,p)=P_k(q,q)=\cos\theta, \\
P_k(p,q)=-P_k(q,p)=\sin\theta, \\
P_k(i,i)=1,\ \ P_k(i,j)=0,\quad i,j\neq p,q.
\end{cases}
$$

高效计算sin(θ)和cos(θ)：

当 $|\theta| \le \dfrac{\pi}{4}$ 时，由三角恒等式有：

$$2\cos^2\theta - 1 = \cos 2\theta = \frac{1}{\sqrt{1+\mathrm{tg}^2 2\theta}} = \frac{y}{\sqrt{x^2+y^2}} \qquad 2\cos^2\theta = 1 + \frac{y}{\sqrt{x^2+y^2}}$$

$\cos\theta$始终取正值.

为避免精度不高，可采用

$$\sin 2\theta = 2\sin\theta\cos\theta = \mathrm{tg}2\theta \cdot \cos 2\theta = \frac{x}{\sqrt{x^2+y^2}}$$

即

$$\sin\theta = \frac{\sin 2\theta}{2\sqrt{(1+\cos 2\theta)/2}}.$$

可以证明:

设矩阵 $A \in \mathrm{R}^{n\times n}$是**对称矩阵**，记$A_0 = A$，对$A$作一系列旋转相似变换，即

$$A_k = P_k A_{k-1} P_k^T, (k = 1, 2, \cdots)$$

其中 A_k $(k = 1，2，...)$ 仍是对称矩阵，且逐渐趋于对角矩阵.

引例 设矩阵 $A=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$，用吉文斯旋转变换求$A$的所有特征值和特征向量.

解： A实对称，使用 Givens 变换，取旋转矩阵 $P_{12}(\pi/4)$.

因为对 $|a_{21}|=\max\{|a_{21}|,|a_{31}|\}$，则有

$$A^{(1)}=P_{12}^T A P_{12}=\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 3 & -1 \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 \end{pmatrix}$$

同理取 $P_{13}(\theta)$ ，其中 $\theta = \frac{1}{2}\arctan\frac{2a_{13}^{(1)}}{a_{11}^{(1)} - a_{33}^{(1)}} = \frac{1}{2}\arctan\sqrt{2}$

得

$$A^{(2)} = P_{13}^T A P_{13} = \begin{pmatrix} 0.63398 & -0.32505 & 0 \\ -0.32505 & 3 & -0.62797 \\ 0 & -0.62797 & 2.36603 \end{pmatrix}$$

反复计算有

$$A^{(k)} = \begin{pmatrix} 0.58578 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3.41421 \end{pmatrix}$$

故A的所有特征值为:

$$\lambda_1 = 0.58578,\ \lambda_2 = 2,\ \lambda_3 = 3.41421.$$

此外

$$P = P_{12}P_{13}\cdots P_k = \begin{pmatrix} 0.5 & 0.7071 & 0.5 \\ 0.7071 & 0 & -0.7071 \\ 0.5 & -0.7071 & 0.5 \end{pmatrix}$$

于是λ_1，λ_2，λ_3对应的3个特征向量为:

$$x_1 = \begin{pmatrix} 0.5 \\ 0.7071 \\ 0.5 \end{pmatrix},\ x_1 = \begin{pmatrix} 0.7071 \\ 0 \\ -0.7071 \end{pmatrix},\ x_3 = \begin{pmatrix} 0.5 \\ -0.7071 \\ 0.5 \end{pmatrix}$$

A的准确特征值为: $\lambda_1 = 2-\sqrt{2}\,,\lambda_2 = 2\,,\lambda_3 = 2+\sqrt{2}$

§3　豪斯霍尔德（Househoulder）变换

定义: 设 $v \in R_*^n = R^n - \{0\}$，$n$ 阶豪斯霍尔德矩阵 $H = H(v)$

$$H = I - \frac{2}{v^T v} v v^T$$

I为n阶单位阵.　显然，对任意 $x \in R^n$

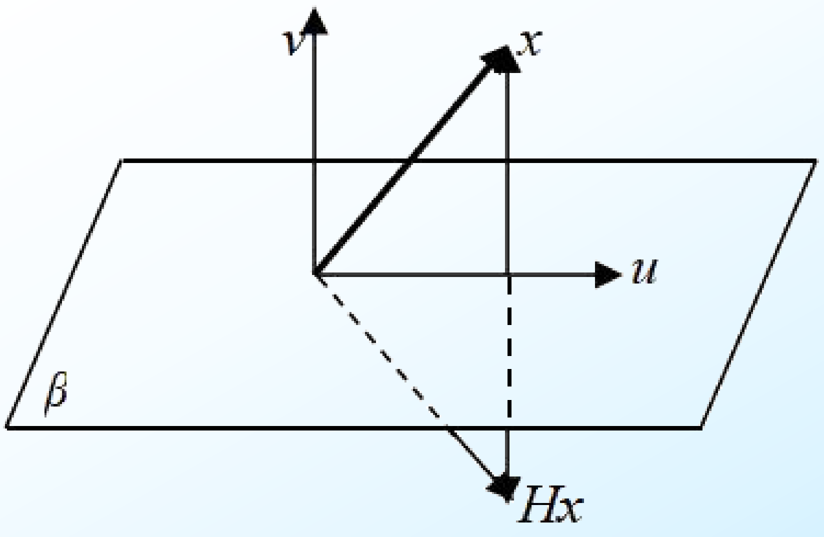


图 4.3 镜面反射图

$$Hx = x - 2\frac{v^T x}{v^T v} v \text{，} \quad v^T Hx = -v^T x$$

定理：设 $b,u\in R_*^n$, $b\neq u$, $\|b\|_2=\|u\|_2$ ，则存在豪斯霍尔德阵H,

使得 $Hb=u$.

证明： 取 $v=b-u\in R_*^n$ 令 $H=I-\dfrac{2}{v^Tv}vv^T$

且 $$Hb=b-\frac{2(b-u)(b-u)^T}{(b-u)^T(b-u)}b$$

$$=b-\frac{2(b-u)^T b\ (b-u)}{b^Tb-b^Tu-u^Tb+u^Tu}.$$

因为 $\|b\|_2=\|u\|_2$ 所以 $b^Tb=u^Tu.$

又因为 $b^T u = \sum_{i=1}^{n} b_i u_i = u^T b$ 故

$$Hb = b - \frac{2\left(b^T b - u^T b\right)}{2\left(b^T b - u^T b\right)}(b-u) = b-(b-u) = u.$$

如何确定H? 取

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \\ b_{r+1} \\ \vdots \\ b_n \end{pmatrix} \qquad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \\ u_{r+1} \\ \vdots \\ u_n \end{pmatrix} \qquad v = b - u = \begin{pmatrix} b_1 - u_1 \\ b_2 - u_2 \\ \vdots \\ b_r - u_r \\ b_{r+1} - u_{r+1} \\ \vdots \\ b_n - u_n \end{pmatrix} \qquad H = I - \frac{2}{v^T v} v v^T$$

比如

$$b=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \\ b_{r+1} \\ \vdots \\ b_n \end{pmatrix} \qquad u=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \\ u_{r+1} \\ \vdots \\ 0 \end{pmatrix}$$

令

$$u_i=\begin{cases} b_i & i=1,\cdots,r, \\ -sign\left(b_{r+1}\right)\sqrt{b_{r+1}^2+\cdots+b_n^2}, & i=r+1, \\ 0, & i=r+2,\cdots,n. \end{cases}$$

满足$u\neq b\in R_*^n$且$\|u\|_2=\|b\|_2$.

定理：设 $A \in R^{n\times n}$ 为对称矩阵，则存在豪斯霍尔德阵 $H_1, H_2, \cdots, H_k \ (k \le n-2)$，使由递推公式

$$A_1 = A \qquad A_{i+1} = H_i A_i H_i \qquad (i = 1, \cdots, k)$$

得到的对称矩阵 A_{k+1} 为三对角阵.

引例：用豪斯霍尔德变换将对称阵 $A = \begin{pmatrix} 6 & 2 & 3 & 1 \\ 2 & 5 & 4 & 8 \\ 3 & 4 & 9 & 1 \\ 1 & 8 & 1 & 7 \end{pmatrix}$

化为三对角矩阵.

解：当 $n=4$，则至多通过 $n-2=4-2=2$ 次豪斯霍尔德变换可将对称阵A化为三对角阵.

令 $A_1=A$，记 $b_1=\begin{pmatrix}6\\2\\3\\1\end{pmatrix}$ 取 $u_1=\begin{pmatrix}6\\-\sqrt{2^2+3^2+1^2}\\0\\0\end{pmatrix}=\begin{pmatrix}6\\-\sqrt{14}\\0\\0\end{pmatrix}$

则 $v_1=b_1-u_1=\begin{pmatrix}0\\2+\sqrt{14}\\3\\1\end{pmatrix}$, $\|v_1\|_2^2=42.9666.$

$$H_1 = I - 2\frac{v_1 v_1^T}{\|v_1\|_2^2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.5345 & -0.8018 & -0.2673 \\ 0 & -0.8018 & -0.5811 & -0.1396 \\ 0 & -0.2673 & -0.1396 & 0.9535 \end{pmatrix}$$

所以

$$H_1 A_1 H_1 = \begin{pmatrix} 6 & -3.7417 & 0 & 0 \\ -3.7417 & 13.8756 & 1.8079 & -3.1394 \\ 0 & 1.8079 & 4.2912 & -6.0079 \\ 0 & -3.1394 & -6.0079 & 2.8512 \end{pmatrix} = A_2$$

记　　　　　　　　　　　取

$$b_2 = \begin{pmatrix} -3.7417 \\ 13.8756 \\ 1.8079 \\ -3.1394 \end{pmatrix} \qquad u_2 = \begin{pmatrix} -3.7417 \\ 13.8756 \\ -\sqrt{1.8079^2 + (-3.1394)^2} \\ 0 \end{pmatrix} = \begin{pmatrix} -3.7417 \\ 13.8756 \\ -3.6228 \\ 0 \end{pmatrix}$$

则

$$v_2 = b_2 - u_2 = \begin{pmatrix} 0 \\ 0 \\ 5.4307 \\ -3.1394 \end{pmatrix}, \qquad \|v_2\|_2^2 = 39.3483.$$

从而

$$H_2 = I - \frac{2v_2 v_2^T}{\|v_2\|_2^2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -0.499\,0 & 0.866\,6 \\ 0 & 0 & 0.866\,6 & 0.499\,0 \end{pmatrix}$$

所以 $$H_2 A_2 H_2 = \begin{pmatrix} 6 & -3.7417 & 0 & 0 \\ -3.7417 & 13.8576 & -3.6227 & 0 \\ 0 & -3.6227 & 8.4058 & -3.6387 \\ 0 & 0 & -3.6387 & -1.2634 \end{pmatrix} = A_3$$

显然，A_3 为所求对称三对角阵.

定义： 对于方阵A，如果当$i > j+1$时，有$a_{ij} = 0$，则称A为上海森伯格矩阵，即

$$A = \begin{bmatrix} a_{11} & a_{11} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \ddots & \ddots & \vdots \\ & & a_{nn-1} & a_{nn} \end{bmatrix}$$

§4　QR 方法

定理：$A \in R^{n\times n}$可逆，则存在正交阵Q使

$$Q^T A Q = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & \cdots & R_{2m} \\ & & \cdots & \cdots \\ & & & R_{mm} \end{pmatrix}$$

R_{ii}至多2阶. 若1阶，其元素即A的特征值；若2阶其特征值为A的一对共轭复特征值.

定理：设 $A \in R^{n\times n}$，则A正交相似于一个n阶上海森伯格矩阵

$$\begin{pmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & * & * \end{pmatrix} \quad (i > j+1 \Rightarrow a_{ij} = 0).$$

QR方法的基本公式

设$A_1 = Q_1R_1$（QR分解），令$A_2 = R_1Q_1$，再次将矩阵A_2作QR分解，不妨设$A_2 = Q_2R_2$，令$A_3 = R_2Q_2$，一般地，有

$$\begin{cases} A_k = Q_k R_k, & (QR\text{分解}) \\ A_{k+1} = R_k Q_k. & (\text{迭代定义}) \end{cases} \qquad k = 1,2,\ldots$$

QR基本算法：

1. 对 $A^{(k)}$ 作 QR 分解 $A^{(k)}=Q_kR_k$.

2. 逆序相乘 $A^{(k)}$ 的分解矩阵，$A^{(k)}=\underline{R_k}Q_{\underline{k}}$.

3. 判别 $A^{(k+1)}$ 是否为主对角线为1×1或2×2的子块形式的分块上三角形矩阵，若是对角线上各子块的特征值为所求特征值，终止，否则 $k+1\rightarrow k$，转1.

例：用QR方法求的A特征值，其中 $A=\begin{pmatrix} 5 & -2 & -5 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 2 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix}$

解：记$A^{(0)}=A$，采用正交变换将$A^{(0)}$做QR分解，得$A^{(0)}=Q_0R_0$，其中

$$Q_0=\begin{pmatrix} 0.9806 & -0.0377 & 0.1923 & -0.1038 \\ 0.1961 & 0.1887 & -0.8804 & -1.4192 \\ 0 & 0.9813 & 0.1761 & 0.0740 \\ 0 & 0 & 0.3962 & -0.8998 \end{pmatrix} \quad R_0=\begin{pmatrix} 5.0992 & -1.9612 & -5.4912 & -0.3922 \\ 0 & 2.0381 & 1.5852 & -2.5288 \\ 0 & 0 & 2.5242 & -3.2736 \\ 0 & 0 & 0 & 0.7822 \end{pmatrix}$$

做逆序相乘

$$A^{(1)} = R_0Q_0 = \begin{pmatrix} 4.6157 & 5.9508 & 1.5992 & 0.2390 \\ 0.3997 & 1.9401 & -2.5717 & 1.5361 \\ 0 & 2.4770 & -0.8525 & 3.1294 \\ 0 & 0 & 0.3099 & -0.7031 \end{pmatrix}$$

继续做下去，经过11次计算后，有

$$A^{(11)} = \begin{pmatrix} 4.0000 & * & * & * \\ 0 & 1.8789 & -3.5910 & * \\ 0 & 1.3290 & 0.1211 & * \\ 0 & 0 & 0 & -1.0000 \end{pmatrix}$$

于是A有两个实特征值为 $\lambda_1 = 4.0000$ ，$\lambda_2 = -1.0000$及一对共轭复特征根 $\lambda_{3,4} = 1 \pm 2i$，它来自2×2矩阵块$\begin{pmatrix} 1.8789 & -3.5910 \\ 1.3290 & 0.1211 \end{pmatrix}$的特征根.

改进的QR方法

（1）对于一般的矩阵在使用QR方法之前，可先A将作正交相似变换化为上海森伯格矩阵H，然后对H作QR迭代，可大量节省运算量.

（2）对上海森伯格矩阵用吉文斯变换作QR分解

$$R(i,i+1,\theta_i)H=\begin{bmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 &&&&& \\ &&& c_i & s_i &&& \\ &&& -s_i & c_i &&& \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{bmatrix}\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1i} & \cdots & h_{1,n-1} & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2i} & \cdots & h_{2,n-1} & h_{2n} \\ & \ddots & \ddots & \cdots & \cdots & \vdots & \vdots \\ &&& h_{ii} & \cdots & h_{i,n-1} & h_{i,n} \\ &&& h_{i+1,i} & \cdots & h_{i+1,n-1} & h_{i+1,} \\ &&&& \ddots & \vdots & \vdots \\ &&&&& h_{n,n-1} & h_{n,n} \end{bmatrix}$$

（3）记$H_1 = H$，设$H_1= U_1R_1$ 令$H_2= R_1U_1$. 由于R_1为上三角，U_1为上海森伯格阵，所以H_2为上海森伯格阵. 此变换约需$4n^2$次乘除和加减运算.