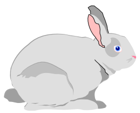
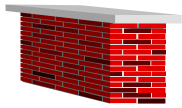
6. 矩阵特征值的数值计算





6.1 引言

工程技术的许多实际问题，例如振动问题，稳定问题的求解，有时会归结为求矩阵的特征值 λ 和对应的特征向量 x 。学过线性代数后，我们已知求矩阵**A**的特征值 λ 和特征向量 x 的解法，即先求出A的特征多项式：

$$\det(A-\lambda E)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

建立特征方程： $\det(A-\lambda E)=0$

求解特征方程，所得根 λ_0 即为矩阵A的特征值，然后求解方程组（ $A-\lambda_0 E$ ）$X=0$，就可得出矩阵A对应于特征值 λ_0 的特征向量X。

但众所周知，高次代数方程求根是相当困难的，而且重根的计算精度较低。同时，矩阵A求特征多项式系数的过程对舍入误差十分敏感，这对最后计算结果影响很大。因此，从数值计算角度来看，上述方法缺乏实用价值。

目前，求矩阵特征值问题实际采用的是迭代法。这里将介绍两种方法：幂法、反幂法以及QR方法。

引入Gerishgorin圆盘

$$D_i=|z-a_{ii}|\le\Lambda_i, i=1,2,\cdots,n$$

定理 1.1.6 第一圆盘定理(Gerischgorin) $A\in C^{n\times n}$, λ_i 为其特征值, $\Lambda_i=\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ij}|$,在复平面上

定义 $D_i=|z-a_{ii}|\le\Lambda_i, i=1,2,\cdots,n$ 为 A 的 Gerishgorin 圆盘.则 $\forall\lambda_i$, $\lambda_i\in\bigcup_{i=1}^{n}D_i$

定理 1.1.7 第二圆盘定理　由定理 1.1.6 的所有圆组成的连通部分中任意取一个,如它由 k 个圆组成,那们这个连通部分必有且仅由 A 的 k 个特征值(特征值相同时重复计算)

P128

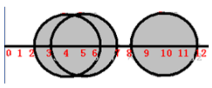
例 1用Gerschgorin定理估计矩阵 $A=\begin{bmatrix}4&1&1\\1&10&1\\1&1&5\end{bmatrix}$ 的特征值的范围。

解

$C_1=\{z:|z-4|<2\}$　$C_2=\{z:|z-10|<2\}$　$C_3=\{z:|z-5|<2\}$

其中C_1与C_3为连通，并与C_2分离，

A的特征值在右图阴影内部.



在C_1与C_3为连通分支中有**2**个特征根，在C_2中有**1**个特征根。

A的3个特征值分别约为3.36375086761336，5.22592119496650和 10.41032793742014.

6.2 幂法与反幂法

1. 幂法(求矩阵按模最大的特征值)

定理6-2：设矩阵**A**的特征值为 $|\lambda_1|>|\lambda_2|\ge\cdots\ge|\lambda_n|$

并设A有完全的特征向量系 $x_1,x_2,\cdots,x_n$ (它们线性无关)，则对任意一个非零向量$\mathbf{V}_0\in\mathbf{R}^n$ 所构造的向量序列 $V_k=AV_{k-1}$,有 $\lim_{k\to\infty}\frac{(V_k)_j}{(V_{k-1})_j}=\lambda_1$

其中 $(\mathbf{V}_k)_j$ 表示向量$\mathbf{V}_k$的第**j**个分量. $V_k=\begin{pmatrix}v_1^{(k)}\\ \vdots\\ v_j^{(k)}\\ \vdots\\ v_n^{(k)}\end{pmatrix}$, $(V_k)_j=v_j^{(k)}$

P149

证 仅就λ_1为实数的情况来证明.假定

$$V_0=\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n(\alpha_1\neq 0)$$

$$V_1=AV_0=\alpha_1Ax_1+\alpha_2Ax_2+\cdots+\alpha_nAx_n \qquad Ax_i=\lambda_ix_i$$
$$=\alpha_1\lambda_1x_1+\alpha_2\lambda_2x_2+\cdots+\alpha_n\lambda_nx_n$$

$$V_2=AV_1=A^2V_0=\alpha_1\lambda_1^2x_1+\alpha_2\lambda_2^2x_2+\cdots+\alpha_n\lambda_n^2x_n$$

…………..

$$V_k=AV_{k-1}=A^kV_0=\alpha_1\lambda_1^kx_1+\alpha_2\lambda_2^kx_2+\cdots+\alpha_n\lambda_n^kx_n$$

$$=\lambda_1^k[\alpha_1x_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^kx_i]$$

$$V_k=\lambda_1^k[\alpha_1x_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^kx_i]$$

同理可得 $$V_{k-1}=\lambda_1^{k-1}[\alpha_1x_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^{k-1}x_i]$$

假定 $(x_1)_j\neq 0$,因为 $\left|\frac{\lambda_i}{\lambda_1}\right|<1(i=2,3,\cdots,n)$,故得

$$\lim_{k\to\infty}\frac{(V_k)_j}{(V_{k-1})_j}=\lambda_1\lim_{k\to\infty}\frac{\alpha_1(x_1)_j+[\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^kx_i]_j}{\alpha_1(x_1)_j+[\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^{k-1}x_i]_j}=\lambda_1$$

$$|m_k|=\|V_k\|_\infty$$

幂法(计算矩阵A的按模最大特征值)

(1)任取一非零向量$V_0\in R^n$,一般可取$V_0=(1,1,\ldots,1)^T$

(2)计算$V_k=AV_{k-1}$

(3)当k足够大时,即可得到: $\lambda_1\approx\frac{(V_k)_j}{(V_{k-1})_j}$

按上述计算过程，有一严重缺点，当$|\lambda_1|>>1$ （或$|\lambda_1|<<1$时）$\{V_k\}$中不为零的分量将随k的增大而无限增大，计算机就可能出现上溢（或随k的增大而很快出现下溢），因此，在实际计算时，须按规范法计算，每步先对向量V_k进行“规范化”，即取V_k中绝对值最大的一个分量记作$m_k=\max(V_k)$，用m_k遍除的向量V_k的所有分量，得到规范化向量u_k ($u_k=V_k/m_k$)。

$$V_1=\begin{pmatrix}1\\-2\\0\end{pmatrix},m_1=\max(V_1)=-2,u_1=\frac{V_1}{m_1}=\begin{pmatrix}-1/2\\1\\0\end{pmatrix}$$

定理6-3：在定理1 的条件下,规范化向量序列$\{u_k\}$收敛于矩阵A按模最大的特征值λ_1对应的特征向量,而向量序列$\{V_k\}$的绝对值最大的分量m_k收敛于λ_1,即

$$\lim_{k\to\infty}u_k=\frac{x_1}{\max(x_1)} \qquad \lim_{k\to\infty}m_k=\lambda_1$$

幂法规范化算法

(1) 任取一非零向量$u_0=V_0\in R^n$,一般可取$V_0=(1,1,\ldots,1)^T$

(2) 计算$V_k=Au_{k-1}$

(3) $m_k=\max(V_k)$, $u_k=V_k/m_k$

当k足够大时,即可得到：$\lambda_1\approx m_k$ $\qquad u_k\approx\frac{x_1}{\max(x_1)}$

例1 用幂法求矩阵

$$A=\begin{bmatrix}133&6&135\\44&5&46\\-88&-6&-90\end{bmatrix}$$

按模最大特征值λ_1和对应的特征向量x_1

解： 取初始向量$V_0=u_0=(1,1,1)^T$，计算出V_k,u_k和m_k,迭代7次的结果列于下表

k	V_k			u_k		
0	1	1	1	1	1	1
1	274	95	-184	1	0.34672	-0.67153
2	44.43277	14.84322	-29.64262	1	0.33413	-0.66727
3	44.92333	14.97623	-29.95048	1	0.33337	-0.66670
4	44.99572	14.99865	-29.99722	1	0.33334	-0.66667
5	44.99959	14.99988	-29.99974	1	0.33333	-0.66667
6	44.99953	14.99983	-29.99968	1	0.33333	-0.66667
7	44.99953	14.99983	-29.99968	1	0.33333	-0.66667

$$m_2=44.42377,m_3=44.92333,m_4=44.99572$$
$$m_5=44.99959,m_6=44.99953,m_7=44.99953$$

由上可见经过7次迭代, m_7的值已稳定到小数后5位，故所求的按模最大特征值和对应的特征向量可取作：

$$\lambda_1\approx 44.9995,\overline{x_1}\approx(1,0.333,-0.6667)^T$$

A的三个特征值与特征向量分别是： $\overline{x_1}\approx\frac{x_1}{\max(x_1)}$

$$\lambda_1=45,\lambda_2=2,\lambda_3=1$$

$$x_1=(3,1,-2)^T,x_2=(3,2,-3)^T,x_3=(2,1,-2)^T$$

当 $\left|\frac{\lambda_2}{\lambda_1}\right|<<1$ 时 幂法收敛快，当 $\left|\frac{\lambda_2}{\lambda_1}\right|\approx 1$ 时 幂法收敛慢。

小结： 幂法适用求矩阵的按模最大特征值及相应的特征向量，其优点是算法简单，容易编写程序在计算机上实现，缺点是收敛速度慢，其有效性依赖与矩阵特征值的分布情况. 为了加速幂法，人们提出若干措施：（1）原点平移法，（2）Aitken加速法，Rayleigh商加速，见 P155。

2.反幂法（求矩阵的按模最小特征值）

基本思路： 设A没有零特征值，则A非奇异，即A的逆阵存在。设A的特征值为 $|\lambda_1| \geq |\lambda_2| \geq \cdots > |\lambda_n| > 0$

其对应的特征向量为 $x_1, x_2, \cdots, x_n$

因为 $A x_k = \lambda_k x_k$ 所以 $A^{-1} x_k = \lambda_k^{-1} x_k$

故λ_k^{-1} 就是矩阵A^{-1}的特征值，它们满足

$$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \geq \cdots \geq \left|\frac{1}{\lambda_1}\right|$$

对应的特征向量仍为x_k。因此，求矩阵A 的按模最小特征值λ_n，就相当于求其逆阵A^{-1}的按模最大特征值λ_n^{-1}，这只需应用幂法于A^{-1}，即可求得。

反幂法规范化算法

(1) 任取一非零向量$u_0=V_0\in R^n$,一般可取$V_0=(1,1,\ldots,1)^T$

(2) 计算$V_k=A^{-1} u_{k-1}$ ⟺ 求解$AV_k=u_{k-1}$

(3) $m_k=\max(V_k)$，$u_k= V_k / m_k$

当k足够大时, 即可得到：$\lambda_n\approx 1/ m_k$ $\quad u_k \approx \dfrac{x_n}{\max(x_n)}$

有时为了加速反幂法，将(2)中

$$V_k = A^{-1}u_{k-1} \quad 改为 \quad V_k = (A-qE)^{-1}u_{k-1}$$

这时求得特征值$(\lambda-q)^{-1}$，这里λ为所有λ_i中与q最接近者。

注意：

由于求逆非常费时。故在求V_k时，$V_k = A^{-1}u_{k-1}$

可采用解方程组 $AV_k = u_{k-1}$ 的办法。

由于每次解方程组的系数矩阵都相同，可预先作三角分解，这样每次迭代仅仅求解两个三角方程组就可以了。特别当n较大时，将大大地节省计算量。

反幂法的适用范围是求矩阵的按模最小特征值及对应的特征向量。

6.3 矩阵的正交分解

1、Householder变换阵 P160

定义 设非零向量 $W \in R^n, W = (w_1, w_2, \cdots, w_n)^T$，且满足条件 $\| W \|_2 = 1$，形如

$$H = E(W, W, 2) = I - 2WW^T$$

的 n 阶方阵称为初等反射阵，或称为 Householder 变换（矩阵）。

例1：设 $u = (1 \quad 0 \quad 1)^T \in R^3$,则$\| u \|_2 = \sqrt{2}$

取：$W = \dfrac{u}{\| u \|_2} = \left(\dfrac{1}{\sqrt{2}} \quad 0 \quad \dfrac{1}{\sqrt{2}}\right)^T \in R^3, \| W \|_2 = 1$

$$H = E - 2WW^T = E - 2\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}\left(\frac{1}{\sqrt{2}} \quad 0 \quad \frac{1}{\sqrt{2}}\right)$$

$$= \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

H为Householder变换阵

H阵的性质：P160,Th6-4

(1)对称 $H = H^T$ $\quad H = E - 2WW^T, \|W\|_2 = 1$

(2)正交 $HH^T = H^2 = (E-2WW^T)(E-2WW^T)$

$= E - 4WW^T + 4WW^TWW^T = E.$ $\quad H^{-1} = H^T = H$

(3) 若 $Hx = y$,则$\|x\|_2 = \|y\|_2$

H变换下，向量的长度保持不变。

(4)镜面反射性质

(i) $Hw = -w$

-1为H的一个特征值，w为对应的特征向量;

(ii) H的其余$n-1$个特征值为1。

对应于1的特征向量是与w正交的非零向量。

（4）镜映射－几何意义 $H=E-2WW^T, \|W\|_2=1$

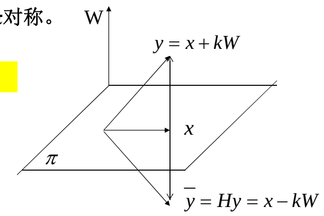
平面π 方程 $W^Tx=0 \quad \forall x\in\pi$ 平面π称为。

若 $x\in\pi$ ，$Hx=(E-2WW^T)x=x-2WW^Tx=x$

若 $y\notin\pi$ ， $y=x+kW$， $Hy=H(x+kW)=x-kW=\bar{y}$

$\bar{y}=Hy$ 与y关于平面π对称。

H称为镜面反射阵。



Householder变换阵的一个重要性质是可在一个向量中引入零分量.

定理6-5 对任意非零向量, $x=(x_1,x_2,\ldots,x_n)^T$,存在一个Householder变换阵 H,使 $Hx=\sigma e_1$

其中, $e_1=(1,0,\ldots,0)^T$, $\sigma=\pm\|x\|_2$

$$H\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\sigma\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$$

证明 令 $U=x-\sigma e_1$, $W=U/\|U\|_2$

构造Householder变换阵

$$H=E-2WW^T=E-2\frac{UU^T}{\|U\|_2^2}=E-\frac{1}{\rho}UU^T$$

其中 $\rho=\frac{1}{2}\|U\|_2^2=\frac{1}{2}((x_1-\sigma)^2+x_2^2+\ldots+x_n^2)$

$$U=x-\sigma e_1=\begin{pmatrix}x_1-\sigma\\x_2\\\vdots\\x_n\end{pmatrix}$$

$=\frac{1}{2}(2\sigma^2-2x_1\sigma)=\sigma(\sigma-x_1)$

取 $\sigma=-sign(x_1)\|x\|_2$

$$Hx=(E-\frac{1}{\rho}UU^T)x=x-\frac{1}{\rho}UU^Tx=x-U=\sigma e_1$$

$$U=x-\sigma e_1=\begin{pmatrix}x_1-\sigma\\x_2\\\vdots\\x_n\end{pmatrix}$$

$U^Tx=\sigma^2-x_1\sigma$

$\rho=\sigma(\sigma-x_1),\ \sigma^2=\|x\|_2^2$

$U^Tx=\rho$ 证毕

$$\sigma=-sign(x_1)\|x\|_2$$

例2 设 $x=(3,5,1,1)^T$,求 Householder变换阵 H,使 $Hx=\sigma e_1$ 其中, $e_1=(1,0,0,0)^T$。

$$\sigma=-sign(x_1)\|x\|_2$$

解： $\because\ \sigma=-\|x\|_2=-6,\quad \therefore\ U=x-\sigma e_1=(9,5,1,1)^T$

$\rho=\sigma(\sigma-x_1)=54$

$$H=E-\frac{1}{\rho}UU^T=\frac{1}{54}\begin{bmatrix}-27&-45&-9&-9\\-45&29&-5&-5\\-9&-5&53&-1\\-9&-5&-1&53\end{bmatrix}\qquad Hx=\begin{pmatrix}-6\\0\\0\\0\end{pmatrix}$$

对非零向量 $x\in R^n$,存在 Householder变换阵 H,使 $Hx=\sigma e_1$。

推论 存在Householder变换阵 H_j,使 P162

$$H_j\begin{pmatrix}x_1\\\vdots\\x_{j-1}\\x_j\\x_{j+1}\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}x_1\\\vdots\\x_{j-1}\\\sigma\\0\\\vdots\\0\end{pmatrix},$$

$H_j=E-\frac{1}{\rho}uu^T$,

$u=(0,\cdots,0,x_j-\sigma,x_{j+1}\cdots,x_n)^T$

$\sigma=-sign(x_j)(x_j^2+\cdots+x_n^2)^{1/2}$

$\rho=\sigma(\sigma-x_j)$

$$x=\begin{pmatrix}x_1\\\vdots\\x_p\\0\\\vdots\\0\end{pmatrix},\ p\le j-1$$

性质 $H_jx=x$

$u^Tx=0$

例3 设 $x=(3,-5,1,1)^T$,求 Householder变换阵 H,使 $Hx=(3,\ \sigma,0,0)^T$。

解：由推论， $\sigma=-sign(x_2)(x_2^2+x_3^2+x_4^2)^{1/2}=3\sqrt{3}$

$u=(0,x_2-\sigma,x_3,x_4)^T=(0,-5-3\sqrt{3},1,1)^T$

$\rho=\sigma(\sigma-x_2)=3\sqrt{3}(5+3\sqrt{3})$

$$H=E-\frac{1}{\rho}uu^T=\frac{1}{3\sqrt{3}(5+3\sqrt{3})}\begin{pmatrix}3\sqrt{3}(5+3\sqrt{3})&0&0&0\\0&-5(5+3\sqrt{3})&5+3\sqrt{3}&5+3\sqrt{3}\\0&5+3\sqrt{3}&26+15\sqrt{3}&-1\\0&5+3\sqrt{3}&-1&26+15\sqrt{3}\end{pmatrix}$$

$$Hx=\begin{pmatrix}3\\3\sqrt{3}\\0\\0\end{pmatrix}$$

若 $H\begin{pmatrix}3\\0\\0\\0\end{pmatrix}=\begin{pmatrix}3\\0\\0\\0\end{pmatrix}$ $Hx=\begin{pmatrix}3\\-5\\\sigma\\0\end{pmatrix}$, $\sigma=?,\ H=?$

3、平面旋转阵（Givens变换） P163

定义　n阶方阵

$$
R(i,j)=\begin{bmatrix}
1 &&&&&&&& \\
& \ddots &&&&&&& \\
&& 1 &&&&&& \\
&&& \cos\theta &&& \sin\theta && \\
&&&& 1 &&&& \\
&&&&& \ddots &&& \\
&&&&&& 1 && \\
&&& -\sin\theta &&& \cos\theta && \\
&&&&&&& 1 & \\
&&&&&&&& \ddots \\
&&&&&&&&& 1
\end{bmatrix}
\begin{array}{l} \\ \\ \\ \text{第 } i \text{ 行} \\ \\ \\ \\ \text{第 } j \text{ 行} \\ \\ \\ (j>i) \end{array}
=\boldsymbol{J(i,j,\theta)}
$$

称为平面旋转阵，或称为Givens变换（矩阵）,也记为$J(i,j,\theta)$。

1. $R_{i,j}$ 阵的性质

① $R_{ij}^T \cdot R_{ij} = I$，因此 $R_{ij}^{-1} = R_{ij}^T$，

平面旋转是非对称的正交阵；

② R_{ij}^T 也是一个平面旋转阵；

③ $\det(R_{ij}) = 1$。

2. $R_{i,j}$ 阵的作用

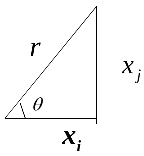
(1) R_{ij} 左乘向量 $x = (x_1, x_2, \cdots, x_n)^T$ 只改变 x 的第 i 个和第 j 个分量

若令 $y = R_{ij}x$，有

$$y_i = x_i\cos\theta + x_j\sin\theta$$
$$y_j = -x_i\sin\theta + x_j\cos\theta$$
$$y_k = x_k \qquad (k = 1,2,\cdots,n; k \neq i,j)$$

调整 θ 使 $y_j = 0$, $tg\theta = \dfrac{x_j}{x_i}$,

$$S=\sin\theta=\frac{x_j}{r},\quad C=\cos\theta=\frac{x_i}{r},\quad r=\sqrt{x_i^2+x_j^2}$$



于是

$$y_i = Cx_i + Sx_j = r = \sqrt{x_i^2 + x_j^2}$$
$$y_j = -Sx_i + Cx_j = 0$$

有 $R_{ij}x = (x_1,\cdots,x_{i-1},r,x_{i+1},\cdots,x_{j-1},0,x_{j+1},\cdots,x_n)^T$

4. 定义　　P165

上*Hessenberg*(海森伯格)阵 $A \in R^{n\times n}$

当 $i > j+1$ 时, $a_{ij} = 0$

$$
A=\begin{pmatrix}
* & * & \cdots & \cdots & * \\
* & * & & & * \\
0 & \ddots & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & * \\
0 & \cdots & 0 & * & *
\end{pmatrix}
$$

定理6-6 设 $A\in R^{n\times n}$，则总存在正交阵Q，使 **QAQ^{-1}** 为上Hessenberg矩阵。

例4　用Givens变换将A化为上三角阵R. $A=\begin{pmatrix}1&1&3\\2&3&4\\0&-2&1\end{pmatrix}$

解 设 $A = A^{(1)} = \left(\alpha_1^{(1)} \quad \alpha_2^{(1)} \quad \alpha_3^{(1)}\right)$

取 $J(1,2,\theta)=\begin{pmatrix}\cos\theta&\sin\theta&0\\-\sin\theta&\cos\theta&0\\0&0&1\end{pmatrix}$ 使 $J(1,2,\theta)\alpha_1^{(1)}=\begin{pmatrix}*\\0\\0\end{pmatrix}$

$$\Rightarrow \sin\theta=\frac{2}{\sqrt5},\quad \cos\theta=\frac{1}{\sqrt5}$$

$$J(1,2,\theta)A^{(1)}=\left(\alpha_1^{(2)} \quad \alpha_2^{(2)} \quad \alpha_3^{(2)}\right)=A^{(2)}=\begin{pmatrix}\sqrt5&\frac{7}{\sqrt5}&\frac{11}{\sqrt5}\\0&\frac{1}{\sqrt5}&\frac{-2}{\sqrt5}\\0&-2&1\end{pmatrix},$$

取 $J(2,3,\theta)=\begin{pmatrix}1&0&0\\0&\cos\theta&\sin\theta\\0&-\sin\theta&\cos\theta\end{pmatrix}$ 使 $J(2,3,\theta)\alpha_2^{(2)}=\begin{pmatrix}*\\*\\0\end{pmatrix}$

$$\Rightarrow \sin\theta=-\frac{2\sqrt5}{\sqrt{21}},\quad \cos\theta=\frac{1}{\sqrt{21}}$$

$$J(2,3,\theta)J(1,2,\theta)A^{(1)}=J(2,3,\theta)A^{(2)}=\begin{pmatrix}\sqrt5&\frac{7}{\sqrt5}&\frac{11}{\sqrt5}\\0&-\frac{4}{\sqrt{21}}&\frac{5}{\sqrt{21}}\\0&0&-\frac{3}{\sqrt{21}}\end{pmatrix}=R$$

6.4 QR方法

1. 矩阵的QR分解

定理6-7 (QR分解定理）

设 $A\in R^{n\times n}$非奇异，则存在正交阵Q与上三角阵R，使 **A=QR**

且当R对角元素为正时，分解是唯一的。

P168

化矩阵A 为上三角阵

$$A=\begin{bmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{bmatrix} \to \begin{bmatrix} * & \cdots & * \\ & \ddots & \vdots \\ & & * \end{bmatrix}$$

$$\boldsymbol{H}_1=\boldsymbol{E}-\frac{1}{\rho_1}\boldsymbol{U}^{(1)}\boldsymbol{U}^{(1)T},$$

$$\boldsymbol{U}^{(1)}=\boldsymbol{\alpha}_1^{(1)}-\sigma_1\boldsymbol{e}_1$$

$$\sigma_1=-sign(a_{11}^{(1)})\|\alpha_1^{(1)}\|_2$$

设 $A = A^{(1)} = [\alpha_1^{(1)} \quad \alpha_2^{(1)} \quad \cdots \quad \alpha_n^{(1)}]$ $\quad \rho_1=\sigma_1(\sigma_1-a_{11}^{(1)})$

对$\alpha_1^{(1)}$，构造$Householder$阵 H_1,使得 $\boldsymbol{H}_1\boldsymbol{\alpha}_1^{(1)}=\sigma_1\boldsymbol{e}_1$

$$\boldsymbol{H}_1\boldsymbol{A}^{(1)}=\begin{bmatrix}\boldsymbol{H}_1\boldsymbol{\alpha}_1^{(1)} & \boldsymbol{H}_1\boldsymbol{\alpha}_2^{(1)} & \cdots & \boldsymbol{H}_1\boldsymbol{\alpha}_n^{(1)}\end{bmatrix}$$

$$=\begin{bmatrix} \sigma_1 & \alpha_{12}^{(2)} & \cdots & \alpha_{1n}^{(2)} \\ 0 & \alpha_{22}^{(2)} & \cdots & \alpha_{2n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \alpha_{n2}^{(2)} & \cdots & \alpha_{nn}^{(2)} \end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1^{(2)} & \boldsymbol{\alpha}_2^{(2)} & \cdots & \boldsymbol{\alpha}_n^{(2)}\end{bmatrix}=\boldsymbol{A}^{(2)}$$

对$A^{(2)}$的第二列$\alpha_2^{(2)}$ 构造H_2使$H_2\alpha_2^{(2)}=(*,\sigma_2,0,\cdots,0)^T$

$$H_2A^{(2)}=\left[H_2\alpha_1^{(2)},H_2\alpha_2^{(2)},\cdots,H_2\alpha_n^{(2)}\right]$$

$$H_2\alpha_1^{(2)}=\alpha_1^{(2)}$$

$$=\begin{bmatrix} \sigma_1 & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{14}^{(2)} \\ 0 & \sigma_2 & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} \\ \vdots & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}\triangleq A^{(3)}=\left[\alpha_1^{(2)},\alpha_2^{(3)},\cdots,\alpha_n^{(3)}\right]$$

$$A^{(3)}=H_2A^{(2)}=H_2H_1A^{(1)}$$

$$\boldsymbol{H}_2=\boldsymbol{E}-\frac{1}{\rho_2}\boldsymbol{U}^{(2)}\boldsymbol{U}^{(2)T},\qquad \sigma_2=-sign(a_{22}^{(2)})[(a_{22}^{(2)})^2+\cdots+(a_{n2}^{(2)})^2]^{1/2}$$

$$\boldsymbol{U}^{(2)}=(0,a_{22}^{(2)}-\sigma_2,a_{32}^{(2)},\cdots,a_{n2}^{(2)})^T\qquad \rho_2=\sigma_2(\sigma_2-a_{22}^{(2)})$$

一般的,设 A(k) 按列分块为 $A^{(k)}=[\alpha_1^{(2)} \quad \alpha_2^{(3)} \quad \cdots \quad \alpha_{k-1}^{(k)} \quad \cdots \quad \alpha_n^{(k)}]$

构造H_k，使

$$H_kA^{(k)}=\left[H_k\alpha_1^{(2)} \quad H_k\alpha_2^{(3)} \quad \cdots \quad H_k\alpha_{k-1}^{(k)} \quad \cdots \quad H_k\alpha_n^{(k)}\right]$$

$$=\begin{bmatrix} \sigma_1 & \cdots & a_{1,k}^{(2)} & a_{1,k+1}^{(2)} & \cdots & a_{1,n}^{(2)} \\ 0 & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \sigma_k & a_{k,k+1}^{(k+1)} & \cdots & a_{k,n}^{(k+1)} \\ 0 & \cdots & 0 & a_{k+1,k+1}^{(k+1)} & \cdots & a_{k+1,n}^{(k+1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,k+1}^{(k+1)} & \cdots & a_{n,n}^{(k+1)} \end{bmatrix}=\boldsymbol{A}^{(k+1)}$$

$$=[\alpha_1^{(2)} \quad \alpha_2^{(3)} \quad \cdots \quad \alpha_{k-1}^{(k)} \quad \alpha_k^{(k+1)} \quad \cdots \quad \alpha_n^{(k+1)}]$$

$$A^{(k+1)}=H_kA^{(k)}=H_kH_{k-1}\cdots H_1A^{(1)}$$

$$H_k=\boldsymbol{E}-\frac{1}{\rho_k}U^{(k)}(U^{(k)})^T$$

$$\sigma_k=-sgin(a_{kk}^{(k)})\left(\sum_{i=k}^{n}(a_{ik}^{(k)})^2\right)^{\frac{1}{2}}$$

如此下去，经过n-1步

$$\rho_k=\sigma_k(\sigma_k-a_{kk}^{(k)})$$

$$U^{(k)}=(\underbrace{0,\ldots,0}_{k-1},a_{kk}^{(k)}-\sigma_k,a_{k+1,k}^{(k)},\ldots,a_{nk}^{(k)})$$

$$A=\begin{bmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{bmatrix} \to \begin{bmatrix} * & \cdots & * \\ & \ddots & \vdots \\ & & * \end{bmatrix}$$

Householder变换对A作QR分解

$\forall A\in R^{n\times n}$非奇异

构造$Householder$阵 $H_k\in R^{n\times n}(k=1,2,\cdots,n-1)$

使$H_{n-1}H_{n-2}\cdots H_2H_1A=R$（上三角阵）

$$A=H_1^{-1}H_2^{-1}\cdots H_{n-1}^{-1}R=H_1H_2\cdots H_{n-1}R=QR$$

其中 $Q=H_1H_2\cdots H_{n-1}\in R^{n\times n}$为正交阵

$$R=Q^TA=H_{n-1}H_{n-2}\cdots H_2H_1A$$

$$H_k^{-1}=H_k^T=H_k,Q^{-1}=Q^T=H_{n-1}H_{n-2}\cdots H_2H_1$$

例1 用Householder变换对A作QR分解

$$A=\begin{bmatrix} 3 & -14/13 & -19/13 \\ 4 & 0 & 5/13 \\ 12 & -3/13 & -11/13 \end{bmatrix}\qquad \alpha_1^{(1)}=\begin{pmatrix}3\\4\\12\end{pmatrix}$$

解 设 $A=A^{(1)}=[\alpha_1^{(1)} \quad \alpha_2^{(1)} \quad \alpha_3^{(1)}]$

对$\alpha_1^{(1)}$，构造**Householder**阵 $\boldsymbol{H}_1$,使得 $\boldsymbol{H}_1\alpha_1^{(1)}=\sigma_1\boldsymbol{e}_1$

$$\because \sigma_1=-sign(a_{11}^{(1)})\|\alpha_1^{(1)}\|_2=-13\qquad \rho_1=\sigma_1(\sigma_1-a_{11}^{(1)})=208$$

$$U^{(1)}=\alpha_1^{(1)}-\sigma_1e_1=\begin{pmatrix}16\\4\\12\end{pmatrix}\qquad H_1=I-\frac{1}{\rho_1}U^{(1)}U^{(1)T}=\begin{bmatrix} -3/13 & -4/13 & -12/13 \\ -4/13 & 12/13 & -3/13 \\ -12/13 & -3/13 & 4/13 \end{bmatrix}$$

$$A^{(2)}=H_1A^{(1)}=\begin{bmatrix}-13&6/13&1\\0&5/13&1\\0&12/13&1\end{bmatrix}=\begin{bmatrix}\alpha_1^{(2)}&\alpha_2^{(2)}&\alpha_3^{(2)}\end{bmatrix}\quad \alpha_2^{(2)}=\begin{pmatrix}6/13\\5/13\\12/13\end{pmatrix}$$

对$\alpha_2^{(2)}$，构造H_2，使得 $H_2\alpha_2^{(2)}=(6/13,\sigma_2,0)^T$

$\sigma_2=-sign(a_{22}^{(2)})[(a_{22}^{(2)})^2+(a_{32}^{(2)})^2]^{1/2}=-1$ $\quad\rho_2=\sigma_2(\sigma_2-a_{22}^{(2)})=18/13$

$U^{(2)}=(0,a_{22}^{(2)}-\sigma_2,a_{32}^{(2)})^T=(0,18/13,12/13)^T$

$$H_2=I-\frac{1}{\rho_2}U^{(2)}U^{(2)T}=\begin{bmatrix}1&0&0\\0&-5/13&-12/13\\0&-12/13&5/13\end{bmatrix}$$

$$A^{(3)}=H_2A^{(2)}=H_2H_1A^{(1)}=\begin{bmatrix}-13&6/13&1\\0&-1&-17/13\\0&0&-7/13\end{bmatrix}=R$$

$A=A^{(1)}=H_1H_2R \qquad \Rightarrow\ Q=H_1H_2$

$$Q=\begin{bmatrix}-\frac{39}{169}&\frac{164}{169}&-\frac{12}{169}\\-\frac{52}{169}&-\frac{24}{169}&-\frac{159}{169}\\-\frac{156}{169}&-\frac{33}{169}&\frac{56}{169}\end{bmatrix}\qquad A=QR$$

2. QR方法

P171

QR方法是求n阶实矩阵的全部特征值的最有效方法之一

1. QR方法的原理及收敛性

记 $A^{(1)}=A$

对$A^{(1)}$作QR分解：$A^{(1)}=Q_1R_1$，Q_1为正交矩阵，R_1为上三角阵。

构造矩阵 $A^{(2)}=R_1Q_1=Q_1^{-1}A^{(1)}Q_1=Q_1^TA^{(1)}Q_1$，则$A^{(2)}$相似于$A^{(1)}$

对$A^{(2)}$作QR分解 $A^{(2)}=Q_2R_2$

构造矩阵 $A^{(3)}=R_2Q_2=Q_2^{-1}A^{(2)}Q_2=Q_2^TA^{(2)}Q_2$

$A^{(3)}$相似于$A^{(2)}$相似于$A^{(1)}$

依此类推

对A_k作QR分解 $A^{(k)}=Q_kR_k$

构造矩阵 $A^{(k+1)}=R_kQ_k=Q_k^{-1}A^{(k)}Q_k=Q_k^TA^{(k)}Q_k$

$A^{(k+1)}$相似于$A^{(k)}$相似于A, $\quad k=1,2,\dots$

从矩阵A开始，得到一个矩阵序列 $A=A^{(1)},A^{(2)},\dots,A^{(k)},A^{(k+1)},\dots$

它们彼此相似，它们有相同的特征值。

可以证明：若A可逆，则

$$\lim_{k\to\infty}A^{(k)}=\begin{bmatrix}\lambda_1&*&\cdots&*\\&\lambda_2&&\vdots\\&&\ddots&*\\&&&\lambda_n\end{bmatrix}\ \text{或}\ \begin{pmatrix}*&\cdots&\cdots&\cdots&*\\&*&&&\vdots\\&&*\ *&&\vdots\\&&*\ *&&\vdots\\&&&*&\vdots\\&&&&*\end{pmatrix}，\text{分块上三角阵}$$

A的特征值全部是实数 $\qquad$ A的特征值有复数

QR方法的关键是QR分解. QR分解可用Householder变换来实现

Matlab调用格式： [q,r]=qr(a)

例2 用QR算法求矩阵A的特征值

$$A=\begin{bmatrix}3&-14/13&-19/13\\4&0&5/13\\12&-3/13&-11/13\end{bmatrix}$$

解 由例4，

$$A=Q_1R_1=\begin{bmatrix}-\frac{39}{169}&\frac{164}{169}&-\frac{12}{169}\\-\frac{52}{169}&-\frac{24}{169}&-\frac{159}{169}\\-\frac{156}{169}&-\frac{33}{169}&\frac{56}{169}\end{bmatrix}\cdot\begin{bmatrix}-13&\frac{6}{13}&1\\0&-1&-\frac{17}{13}\\0&0&-\frac{7}{13}\end{bmatrix}$$

$$A^{(2)}=R_1Q_1=\begin{bmatrix}1.9349&-1.28762&0.8202\\1.5148&0.3974&0.5075\\0.4970&0.1051&-0.1784\end{bmatrix}.$$

用$A^{(2)}$代替$A^{(1)}$,依此类推

k=5

$$A^{(5)}=\begin{pmatrix}6.3110&-5.3384&2.8061\\8.3014&-3.8117&3.6675\\0.0000&-0.0000&-0.3455\end{pmatrix}$$

$\lambda_1\approx 1.2497+4.3242i$

$\lambda_2\approx 1.2497-4.3242i$

$\lambda_3\approx -0.3455$

$$\begin{vmatrix}6.3110-\lambda&-5.3384\\8.3014&-3.8117-\lambda\end{vmatrix}=0$$

3. QR方法的实际计算步骤

第一步

$$A \xrightarrow[\text{用}Householder\text{阵作正交相似变换}]{} \text{上}Hessenberg\text{阵}B \qquad \begin{pmatrix} * & \cdots & \cdots & \cdots & * \\ * & \ddots & & & \vdots \\ & \ddots & * & & \vdots \\ & & * & \ddots & \vdots \\ & & & \ddots & \vdots \\ & & & * & * \end{pmatrix}$$

即，QAQ^{-1} =B　P172

第二步

$$B \xrightarrow[\text{用}Given\text{变换产生迭代序列}]{} \begin{cases} B_k = Q_k R_k \\ B_{k+1} = R_k Q_k \end{cases} \rightarrow \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix}$$

用Givens变换矩阵对上Hessenberg矩阵做QR分解。　P176 P177

作业

习题 6
P186:
1（2），4

上机作业4

1. 用规范的幂法与反幂法求矩阵A的按模最大、最小特征值与对应的特征向量。

$$A=\begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}, e=10^{-5}.$$

2.用Householder变换求矩阵A的QR 分解，并用**QR** 方法做3次迭代。

$$A=\begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}.$$