数值分析

主讲教师：贺慧霞

北京航空航天大学数学科学学院

第三章 矩阵特征值与特征向量的算法

----3.3.1 矩阵的QR分解

矩阵的QR分解

定理：设A是一个n阶实方阵，那么A可分解为一个正交矩阵Q和一个上三角矩阵R的乘积，即**正交三角分级**，简称**QR分解**：

$$A = QR.$$

3.3.1 矩阵的QR分解

定义2 设列向量$w\in R^n$，且$w^Tw=1$，称矩阵 $H(w)=I-2ww^T$ 为**初等反射矩阵**，矩阵也称为**豪斯霍尔德**(Householder)**矩阵**. 如果记$w=(w_1,w_2,\ldots,w_n)$，则有

$$H(w)=\begin{pmatrix} 1-2w_1^2 & -2w_1w_2 & \cdots & -2w_1w_n \\ -2w_2w_1 & 1-2w_2^2 & \cdots & -2w_2w_n \\ \vdots & \vdots & & \vdots \\ -2w_nw_1 & -2w_nw_2 & \cdots & 1-2w_n^2 \end{pmatrix}. \qquad (3.1)$$

初等反射矩阵的几何意义：

考虑以w为法向量且过原点O的超平面$S: w^T x = 0$.

对任意向量$v \in R^n$，则$v = x + y$，其中$x \in S, y \in S^{\perp}$，于是

$$Hx = (I - 2ww^T)x = x - 2ww^T x = x.$$

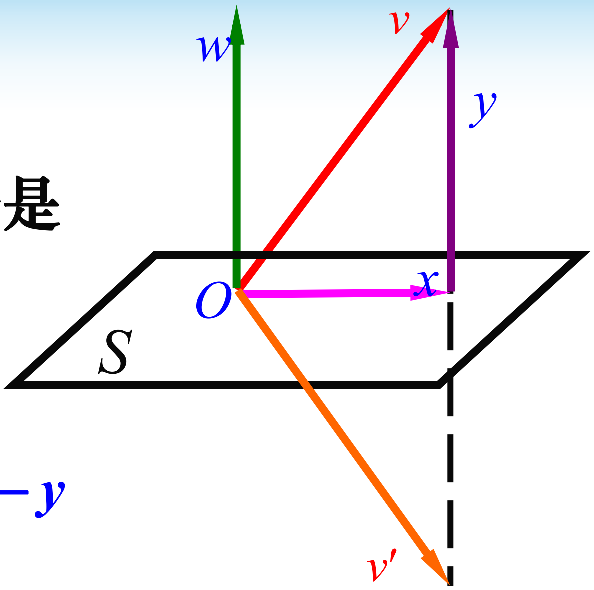
对于$y \in S^{\perp}$，易知 $y // w$

$$Hy = (I - 2ww^T)y = y - 2w\|y\| = y - 2y = -y$$

从而对任意向量$v \in R^n$，总有

$$Hv = H(x - y) = Hx - Hy = v'.$$

其中v'为v关于平面S的镜面反射.



通过Householder变换可以把一个矩阵上三角化和拟上三角化

定理1 设有初等反射矩阵，其中$H=I-2ww^T$，其中$w^Tw=1$，则

(1) H是对称矩阵，即$H^T=H$.

(2) H是正交矩阵，即$H^{-1}=H$.

(3) 设A为对称矩阵，那么 $A_1=H^{-1}AH=HAH$, 即亦是对称矩阵.

证明: 对称性显然,只证H的正交性.

$$H^TH=H^2=(I-2ww^T)(I-2ww^T)=I-4ww^T+4w(w^Tw)w^T=I.$$

设向量$u\neq 0$，则显然 $H=I-2\dfrac{uu^T}{\|u\|_2^2}$ 是一个初等反射矩阵.

定理2 设 x,y 为两个不相等的 n 维向量, $\|x\|_2=\|y\|_2$,则存在一个初等反射矩阵H，使$Hx=y$.

证明: 令 $w=\dfrac{x-y}{\|x-y\|_2}$, 则得到一个初等反射矩阵

$$H=I-2ww^T=I-2\frac{x-y}{\|x-y\|_2^2}(x^T-y^T).$$

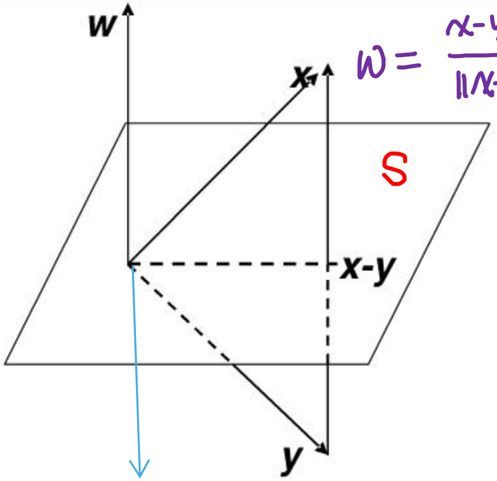
$$x\cdot y=y\cdot x \Leftrightarrow x^T y=y^T x$$

而且 $$Hx=x-2\frac{x-y}{\|x-y\|_2^2}(x^T-y^T)x=x-2\frac{(x-y)(x^Tx-y^Tx)}{\|x-y\|_2^2}.$$

因为 $$\|x-y\|_2^2=(x-y)^T(x-y)=2(x^Tx-y^Tx).$$

所以 $$Hx=x-(x-y)=y.$$

$\omega \perp S$



$$w = \frac{x-y}{\|x-y\|_2}$$

$$y = Hx = (I - 2ww^T)x = x - 2(w^Tx)w$$

$$\Rightarrow x-y = 2(w^Tx)w$$

$$\Rightarrow x-y \parallel \omega.$$

$\therefore$ x和y关于S对称.

显然，*w*是使*Hx=y*成立的唯一长度等于1的向量(不计符号).

引理3.1： 设有非零向量s和单位向量e，则必存在*Householder*矩阵H使得$Hs=\alpha e$，其中α是实数且$|\alpha|=\sqrt{s^Ts}$.

证明： 取单位向量$w=\dfrac{1}{\rho}(s-\alpha e)$.其中$\rho=\sqrt{(s-\alpha e)^T(s-\alpha e)}$,

则有 $$Hs=(I-2ww^T)s=s-\frac{2}{\rho}w(s^T-\alpha e^T)s$$

$$=s-\frac{2}{\rho}w(\alpha^2-\alpha e^Ts)$$

因为 $$\rho^2=(s-\alpha e)^T(s-\alpha e)=2(\alpha^2-\alpha e^Ts)$$

所以 $$Hs=s-\rho w=\alpha e$$

$\|S\|_2=\|\alpha e\|_2=|\alpha|$

$\exists\ H_w$, s.t. $H_w S=\alpha e$

并且 $w=\dfrac{s-\alpha e}{\|s-\alpha e\|_2}$

$\forall\ s=(s_1,s_2,\cdots,s_n)\neq 0$，单位向量$e_i=(0, 0,\cdots, 1, 0,\cdots, 0)^T$，

第i个位置为1，其余位置为0.

$c=-sgn(s_i)\|s\|_2$;(若$s_i=0$，$sgn(s_i)=1$);

$u=s-ce_i$　如果c与si同号,则 s-cei 时有效数字可能会损失,所以取c与si异号.

则　$H=I-\dfrac{1}{\rho}uu^T$;

使得　$Hs=ce_i$.

(**)
$$\begin{cases} H=I-\dfrac{1}{\rho}uu^T; \\ c=-sgn(s_i)\|s\|_2;(\text{若}s_i=0,\ \text{取}c=\|s\|_2); \\ u=s-ce_i \\ \rho=\dfrac{1}{2}\|u\|_2^2 \end{cases}$$

$w=\dfrac{x-y}{\|x-y\|_2}$，$H=I-2ww^T$

【例】 设$x=(3, 5, 1, 1)^T$，则$\|x\|_2=6$. 构造镜面反射H,使得

$$Hx = -\mathrm{sgn}(x_1)\,\|x\|_2\,(1,0,0,0)^T.$$

【解】 $\|x\|_2 = 6,\quad c = -\mathrm{sgn}(x_1)\,\|x\|_2 = -6,\ \ u = x - c\vec{e}_1 = (9,5,1,1)^T,$

$\|u\|_2^2 = 108,\quad \rho = 54,$

所以 $H = I - \dfrac{uu^T}{\rho} =$

$$(**)\begin{cases} H = I - \dfrac{1}{\rho}uu^T; \\ c = -sgn(s_i)\|s\|_2; (若 s_i = 0，取 c = \|s\|_2); \\ u = s - ce_i \\ \rho = \dfrac{1}{2}\|u\|_2^2 \end{cases}$$

$Hx = (-6,0,0,0)^T.$

$$A=(s_1,s_2,s_3,\cdots,s_n)$$

$S_2^{(2)}=(0,a_{22}^{(2)},a_{32}^{(2)},\cdots,a_{n2}^{(2)})^T$，则 $\exists\ H_2$, s.t.

$H_2S_2^{(2)}=(0,c_2,0,\cdots,0)=c_2e_2$

如果$s_1\neq\vec{0}$，则存在H_1, s.t.

$$H_1s_1=c_1e_1,\quad e_1=(1,0,\cdots,0)^T$$

$$A_2=H_1A_1=\begin{bmatrix} c_1 & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix} \xrightarrow{H_2} \begin{bmatrix} c_1 & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & c_2 & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}$$

定理：设A是一个n阶实方阵，那么A可分解为一个正交矩阵Q和一个上三角矩阵R的乘积，$A=QR$

证明：$A=[a_{ij}]_{n\times n}$　$e_r=(0,\ldots,0,1,0,\ldots,0)^T$　（1在第r个位置上）

第 1 步：(1) 设$a_{i1}(i=2,3,\ldots,n)$不全为零，令

$$s_1=(a_{11},a_{21},\ldots,a_{n1})^T \qquad c_1=-sign(a_{11})\sqrt{s_1^T s_1}$$

$$u_1=s_1-c_1e_1 \qquad (\text{若}a_{11}=0,\text{则取}c_1=\sqrt{s_1^T s_1})$$

则　$$H_1=I-2u_1u_1^T/(u_1^T u_1)$$

$$H_1 = I - 2u_1u_1^T/(u_1^Tu_1)$$

由引理知 $\quad H_1s_1 = c_1e_1 = (c_1, 0, \ldots, 0)^T$

所以有 $\quad A_2 = H_1A_1 = \begin{bmatrix} c_1 & a_{12}^{(2)} & \ldots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \ldots & a_{2n}^{(2)} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{n2}^{(2)} & \ldots & a_{nn}^{(2)} \end{bmatrix}$

（2）若$a_{i1} = 0$ $\quad (i = 2, 3, \ldots, n)$ 则取$H_1 = I$ 此时有$A_2 = H_1A = A$

第 2 步：（1） 设$a_{i2}^{(2)}$ $(i=3,4,\ldots,n)$不全为零，令

$$s_2=(0,a_{22}^{(2)},\ldots,a_{n2}^{(2)})^T \qquad c_2=-sign(a_{22}^{(2)})\sqrt{s_2^Ts_2}$$

$$u_2=s_2-c_2e_2 \qquad (\text{若}a_{22}^{(2)}=0,\text{则取}c_2=\sqrt{s_2^Ts_2})$$

构成的Householder阵为：$H_2=I-2u_2u_2^T/(u_2^Tu_2)=\begin{bmatrix}1 & 0\\ 0 & W_1\end{bmatrix}$

其中W_1是$(n-1)\times(n-1)$的矩阵，

并有 $H_2S_2=c_2e_2=(0,c_2,0,\ldots,0)^T$

所以有 $A_3 = H_2 A_2 = \begin{bmatrix} c_1 & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & c_2 & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}$

(2) 若$a_{i2}^{(2)} = 0 \quad (i = 3,4,\ldots,n)$

则取$H_2 = I$ 此时有$A_3 = H_2 A_2 = A_2$

第 3 步：一般情形，设上述方法已经过得到A_r,即

$$A_r=\begin{pmatrix} c_1 & \cdots & a_{1,r-1}^{(2)} & a_{1,r}^{(2)} & a_{1,r+1}^{(2)} & \cdots & a_{1,n}^{(2)} \\ & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & c_{r-1} & a_{r-1,r}^{(r)} & a_{r-1,r+1}^{(r)} & \cdots & a_{r-1,n}^{(r)} \\ & & & a_{r,r}^{(r)} & a_{r,r+1}^{(r)} & \cdots & a_{r,n}^{(r)} \\ & & & \cdots & \cdots & \cdots & \cdots \\ & & & \cdots & \cdots & \cdots & \cdots \\ & & & a_{n,r}^{(r)} & a_{n,r+1}^{(r)} & \cdots & a_{n,n}^{(r)} \end{pmatrix}$$

（1）设$a_{ir}^{(r)}(i=r,r+1,\ldots,n)$不全为零，令

$$s_r=(0,\ldots,0,a_{rr}^{(r)},\ldots,a_{nr}^{(2)})^T \qquad c_r=-sign(a_{rr}^{(r)})\sqrt{s_r^T s_r}$$

$$u_r=s_r-c_re_r \qquad （若a_{rr}^{(r)}=0,则取c_r=\sqrt{s_r^T s_r}）$$

构成的Householder阵为:$H_r=I-2u_ru_r^T/(u_r^Tu_r)=\begin{bmatrix} I_{r-1} & 0 \\ 0 & W_{r-1} \end{bmatrix}$

其中I_{r-1}是$(r-1)\times(r-1)$的单位矩阵，

W_{r-1}是$(n-r+1)\times(n-r+1)$的矩阵，

并有 $H_rs_r=c_re_r=(0,\ldots,0,c_r,0,\ldots,0)^T$

那么，

$$A_{r+1}=H_rA_r=\begin{pmatrix} c_1 & \cdots & a^{(2)}_{1,r-1} & a^{(2)}_{1,r} & a^{(2)}_{1,r+1} & \cdots & a^{(2)}_{1,n} \\ & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & c_{r-1} & a^{(r)}_{r-1,r} & a^{(r)}_{r-1,r+1} & \cdots & a^{(r)}_{r-1,n} \\ & & & c_r & a^{(r+1)}_{r,r+1} & \cdots & a^{(r+1)}_{r,n} \\ & & & \cdots & \cdots & \cdots & \cdots \\ & & & \cdots & \cdots & \cdots & \cdots \\ & & & 0 & a^{(r+1)}_{n,r+1} & \cdots & a^{(r+1)}_{n,n} \end{pmatrix}。$$

（2）若$a^{(r)}_{ir}=0 \quad (i=r+1,\ldots,n)$

则取$H_r=I$　此时有$A_{r+1}=H_rA_r=A_r$

第 4 步：于是，当$r=n-1$时，$A_n=H_{n-1}H_{n-2}\dots H_1A$是上三角阵。

于是得 $A=H_1H_2\dots H_{n-2}H_{n-1}A_n$

令 $\boxed{Q=H_1H_2\dots H_{n-2}H_{n-1},\quad R=A_n}$

则有 $A=QR.$ 完成证明。

矩阵的QR分解

定理：设A是一个n阶实方阵，那么A可分解为一个正交矩阵Q和一个上三角矩阵R的乘积，即：

$$A=QR.$$

当R对角线上的元素为正数时，分解唯一.

【例】用初等反射矩阵将矩阵$A=\begin{array}{c} \quad s_1 \\ \begin{pmatrix} \boxed{\begin{matrix}1\\2\\2\end{matrix}} & \begin{matrix}1\\-1\\-4\end{matrix} & \begin{matrix}1\\-1\\5\end{matrix} \end{pmatrix} \end{array}$分解为QR形式.

【解】第 1 步：设将A的第一列变为与e_1平行的向量，

$$\begin{cases} d_r=\sqrt{\sum_{i=r}^{n}(a_{ir}^{(r)})^2}; \\ c_r=-sign(a_{rr}^{r})d_r;(\text{若}a_{rr}^{r}=0,\text{取}c_r=d_r); \\ u_r=s_r-c_re_r, \\ h_r=\|u_r\|_2^2/2 \qquad H_r=I-\frac{1}{h_r}u_ru_r^T \end{cases}$$

$$d_1=\sqrt{1^2+2^2+2^2}=3,$$

$$c_1=-3,$$

$$u_1=s_1+3e_1=(4,2,2)^T,$$

$$h_1=\frac{\|u_1\|_2^2}{2}=12,$$

$$h_1 = \frac{\| u_1 \|_2^2}{2} = 12,$$

$$H_1 = I - \frac{u_1 u_1^T}{h_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{12} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

$$H_1 A = \begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ 2 & -4 & 5 \end{pmatrix} = \begin{pmatrix} -3 & 3 & -3 \\ 0 & 0 & -3 \\ 0 & -3 & 3 \end{pmatrix}$$

$$s_2 = (0, 0, -3)^T,$$

$$d_2 = \sqrt{0^2 + 0^2 + (-3)^2} = 3, \quad c_2 = 3, \quad u_2 = s_2 - 3e_2 = (0, -3, -3)^T,$$

$$h_2 = \frac{\| u_2 \|_2^2}{2} = 9, \quad H_2 = I - h_2^{-1} u_2 u_2^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

$$H_2H_1A=\begin{pmatrix}1&0&0\\0&0&-1\\0&-1&0\end{pmatrix}\begin{pmatrix}-3&3&-3\\0&0&-3\\0&-3&3\end{pmatrix}=\begin{pmatrix}-3&3&-3\\0&3&-3\\0&0&3\end{pmatrix}$$

故 $A=QR,$

其中$Q=H_1H_2=\begin{pmatrix}\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{1}{3}&-\frac{2}{3}\\\frac{2}{3}&-\frac{2}{3}&\frac{1}{3}\end{pmatrix}$ $\quad R=\begin{pmatrix}-3&3&-3\\0&3&-3\\0&0&3\end{pmatrix}$

避免矩阵相乘的QR分解：

记$A_1 = A$，并记$A_r = [a_{ij}^{(r)}]_{n\times n}$，令$Q_1 = I$

对于 $r = 1,2,\ldots,n-1$执行

1.若$a_{ir}^{(r)}$ $(i = r+1,\ldots,n)$全为零，则令$Q_{r+1} = Q_r, A_{r+1} = A_r$转(5)；

否则转（2）；

$\|u_r\|_2^2 = (a_{rr}^r - c_r)^2 + \sum_{i=r+1}^{n}(a_{ir}^r)^2 = \sum_{i=r}^{n}(a_{ir}^r)^2 - 2c_r a_{rr}^r + c_r^2$

相当

$= 2c_r^2 - 2c_r a_{rr}^r$

$= 2h_r$

2.计算

$$\begin{cases} d_r = \sqrt{\sum_{i=r}^{n}(a_{ir}^{(r)})^2}; \\ c_r = -sign(a_{rr}^r)d_r;(若a_{rr}^r = 0，取c_r = d_r); \\ h_r = c_r^2 - c_r a_{rr}^{(r)} = \frac{1}{2}\|u_r\|_2^2 \end{cases}$$

3.令$u_r = (0,\ldots,0,a_{rr}^r - c_r, a_{r+1,r}^r,\ldots,a_{nr}^r)^T$

4.计算$(Q_{r+1}=Q_rH_r, A_{r+1}=H_rA_r)$

分析过程

结论

$$\begin{cases}\omega_r = Q_r u_r \\ Q_{r+1} = Q_r - \omega_r u_r^T / h_r\end{cases}$$

$$Q_{r+1} = Q_rH_r = Q_r\left[I - 2\frac{u_ru_r^T}{\|u_r\|_2^2}\right]$$

$$= Q_r - 2\frac{Q_ru_ru_r^T}{\|u_r\|_2^2} = Q_r - \underbrace{\boxed{Q_ru_r}}_{\omega_r}\cdot\frac{u_r^T}{h_r}$$

$$\begin{cases}p_r = A_r^T u_r / h_r \\ A_{r+1} = A_r - u_r p_r^T\end{cases}$$

$$A_{r+1} = H_rA_r = \left[I - \frac{2u_ru_r^T}{\|u_r\|_2^2}\right]A_r$$

$$= A_r - \frac{u_ru_r^TA_r}{h_r} = A_r - u_r\cdot\underbrace{\boxed{\frac{(A_r^Tu_r)^T}{h_r}}}_{p_r}$$

5. 继续.

二、实矩阵Schur(舒尔)分解

定理：设 $A \in R^{n\times n}$，则存在正交矩阵$Q \in R^{n\times n}$，

满足： $$Q^T AQ = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & \cdots & R_{2m} \\ & & \ddots & \vdots \\ & & & R_{mm} \end{bmatrix}$$

其中R_{jj} 为实数或具有一对复共轭特征值的2阶方阵

$$R_{jj}=\begin{bmatrix}\alpha_j & \beta_j\\ -\beta_j & \alpha_j\end{bmatrix},\quad \left|\lambda I-R_{jj}\right|=(\lambda-\alpha_j)^2+\beta_j^2=0$$

特征值为$\lambda=\alpha_j\pm i\beta_j$，其中$i$为虚单位

矩阵 $\begin{bmatrix}R_{11} & R_{12} & \cdots & R_{1m}\\ & R_{22} & \cdots & R_{2m}\\ & & \ddots & \vdots\\ & & & R_{mm}\end{bmatrix}$ 称为A的Schur标准形

上面定理说明：只要求得矩阵A的Schur标准形，就很容易求得矩阵A的全部特征值。

三、矩阵的拟上三角化

定义： 称$H \in \mathbf{R}^{n\times n}$为上**Hessenberg**矩阵**(拟上三角矩阵)**，当且仅当$i > j+1$时，$h_{ij} = 0$.

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1,n-2} & h_{1,n-1} & h_{1,n} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2,n-2} & h_{2,n-1} & h_{2,n} \\ 0 & h_{32} & h_{33} & \cdots & h_{3,n-2} & h_{3,n-1} & h_{3,n} \\ 0 & 0 & h_{43} & \cdots & h_{4,n-2} & h_{4,n-1} & h_{4,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & h_{n-1,n-2} & h_{n-1,n-1} & h_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & h_{n,n-1} & h_{n,n} \end{pmatrix}。$$

Householder矩阵将矩阵约化为上Hessenberg矩阵

定理：设A是一个n阶实方阵，则存在Householder矩阵$U_1, U_2, \ldots\ldots, U_{n-2}$使得：

$$U_{n-2}\ldots U_2 U_1 A U_1 U_2 \ldots U_{n-2} = U_0 A U_0 = H.$$

H为上Hessenberg矩阵.

证明：矩阵的拟上三角化的步骤:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & & a_{nn} \end{pmatrix}$$

1.设$a_{i1}\quad(i=3,4,\ldots,n)$不全为零，令

$$S_1=(0,a_{21},a_{31},\ldots,a_{n1})^T \qquad c_1=-sign(a_{21})\sqrt{S_1^TS_1}$$

$$u_1=S_1-c_1e_2 \qquad (若a_{21}=0,则取c_1=\sqrt{S_1^TS_1})$$

定义 H_1, st $S_1 \parallel \vec{e}_2$

构成的Householder阵为：$H_1=I-2u_1u_1^T/(u_1^Tu_1)=\begin{bmatrix}1 & 0\\ 0 & W_1\end{bmatrix}$

其中W_1是$(n-1)\times(n-1)$的矩阵，

并有 $H_1S_1=c_1e_2=(0,c_1,0,\ldots,0)^T$

做的是相似变换，保特征值

因而得 $A^{(2)} = H_1 A H_1 = \begin{bmatrix} a_{11} & a_{12}^{(2)} & \dots & a_{1n}^{(2)} \\ c_1 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} \end{bmatrix}$

若$a_{i1} = 0 \quad (i = 3, 4, \dots, n)$ 则取$H_1 = I$ 此时有 $A^{(2)} = H_1 A H_1 = A$

$S_2 = (0, 0, a_{32}^{(2)}, \cdots, a_{n2}^{(2)}) \parallel \vec{e}_3$

$\exists H_2$

2.设$a_{i2}^{(2)}$　$(i=4,5,\dots,n)$不全为零，令

$$S_2=(0,0,a_{32}^{(2)},\dots,a_{n2}^{(2)})^T \qquad c_2=-sign(a_{32}^{(2)})\sqrt{S_2^T S_2}$$

$$u_2=S_2-c_2e_3 \qquad (\text{若}a_{32}^{(2)}=0\text{，则取}c_2=\sqrt{S_2^T S_2})$$

构成的Householder阵为：$H_2=I-2u_2u_2^T/(u_2^Tu_2)=\begin{bmatrix} I_2 & 0 \\ 0 & W_2 \end{bmatrix}$

其中I_2是2×2的单位阵，W_2是$(n-2)\times(n-2)$的矩阵。

可得　$H_2S_2=c_2e_3=(0,0,c_2,\dots,0)^T$

于是得到 $A^{(3)} = H_2 A^{(2)} H_2 = \begin{bmatrix} a_{11} & a_{12}^{(2)} & a_{13}^{(3)} & \dots & a_{1n}^{(3)} \\ c_1 & a_{22}^{(2)} & a_{23}^{(3)} & \dots & a_{2n}^{(3)} \\ 0 & c_2 & a_{33}^{(3)} & \dots & a_{3n}^{(3)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{n3}^{(3)} & \dots & a_{nn}^{(3)} \end{bmatrix}$

若 $a_{i2}^{(2)} = 0 \quad (i = 4, 5, \dots, n)$

则取 $H_2 = I$ 此时有 $A^{(3)} = H_2 A^{(2)} H_2 = A^{(2)}$

3.一般情形，设上述方法已经得到$A_r \quad (r \geq 2)$

设$a_{ir}^{(r)} \quad (i = r+1, r+2, \ldots, n)$不全为零，令

$$S_r = (0, \ldots, 0 a_{r+1r}^{(2)}, \ldots, a_{nr}^{(2)})^T \quad c_r = -sign(a_{r+1r}^{(r)})\sqrt{S_r^T S_r}$$

$$u_r = S_r - c_r e_{r+1} \quad (\text{若} a_{r+1r}^{(r)} = 0, \text{则取} c_r = \sqrt{S_r^T S_r})$$

构成的Householder阵为：$H_r = I - 2u_r u_r^T / (u_r^T u_r) = \begin{bmatrix} I_r & 0 \\ 0 & W_r \end{bmatrix}$

其中I_r是$r \times r$的单位矩阵，W_r是$(n-r)\times(n-r)$的矩阵， 关键求W_r

并有 $H_r S_r = c_r e_{r+1} = (0, \ldots, 0, c_r, 0, \ldots, 0)^T$

$$\boldsymbol{A}^{(r+1)} = \boldsymbol{H}_r \boldsymbol{A}^{(r)} \boldsymbol{H}_r = \begin{pmatrix} a_{11} & \cdots & a_{1,r}^{(r)} & a_{1,r+1}^{(r+1)} & \cdots & \cdots & a_{1,n}^{(r+1)} \\ c_1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & a_{r,r}^{(r)} & a_{r,r+1}^{(r+1)} & \cdots & \cdots & a_{r,n}^{(r+1)} \\ & & c_r & a_{r+1,r+1}^{(r+1)} & \cdots & \cdots & a_{r+1,n}^{(r+1)} \\ & & & \cdots & \cdots & \cdots & \cdots \\ & & & \cdots & \cdots & \cdots & \cdots \\ & & & a_{n,r+1}^{(r+1)} & \cdots & \cdots & a_{n,n}^{(r+1)} \end{pmatrix}$$

若$a_{ir}^{(r)} = 0 \quad (i = r+2, \ldots, n)$ 则取$\boldsymbol{H}_r = \boldsymbol{I}$ 此时有$\boldsymbol{A}^{(r+1)} = \boldsymbol{A}^{(r)}$

4.当$r=n-1$时，就得到拟上三角矩阵A^{n-1}

$$A=H_{n-2}H_{n-3}\ldots H_2H_1AH_1H_2\ldots H_{n-3}H_{n-2}=P^TAP$$

其中 $P=H_1H_2\ldots H_{n-3}H_{n-2}$为正交矩阵。

特别当A是实对称矩阵时，$A^{(n-1)}$也是对称的，因而是三对角的。

$$\mathbf{A}^{(n-1)}=\begin{bmatrix} a_{11} & c_1 & & & \\ c_1 & a_{22}^{(2)} & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-2} & a_{n-1,n-1}^{(n-1)} & a_{n,n-1}^{(n-1)} \\ & & & a_{n,n-1}^{(n-1)} & a_{nn}^{(n-1)} \end{bmatrix}$$

$$A^{(r)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(2)} & \cdots & a_{1,r-1}^{(r-1)} & a_{1r}^{(r)} & a_{1,r+1}^{(r)} & \cdots & a_{1n}^{(r)} \\ c_1 & a_{22}^{(2)} & \cdots & a_{2,r-1}^{(r-1)} & a_{2r}^{(r)} & a_{2,r+1}^{(r)} & \cdots & a_{2n}^{(r)} \\ & & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & c_{r-1} & a_{rr}^{(r)} & a_{r,r+1}^{(r)} & \cdots & a_{rn}^{(r)} \\ & & & & a_{r+1,r}^{(r)} & a_{r+1,r+1}^{(r)} & \cdots & a_{r+1,n}^{(r)} \\ & & & & \vdots & \vdots & & \vdots \\ & & & & a_{n,r}^{(r)} & a_{n,r+1}^{(r)} & \cdots & a_{nn}^{(r)} \end{pmatrix} = \begin{pmatrix} A_{11}^{(r)} & A_{12}^{(r)} \\ 0 \quad \vec{y}_r & A_{22}^{(r)} \end{pmatrix}$$

$$H_r = \begin{bmatrix} I & 0 \\ 0 & W_r \end{bmatrix}$$

$$A^{(r+1)} = H_r A^{(r)} H_r = \begin{bmatrix} A_{11}^{(r)} & A_{12}^{(r)} W_k \\ 0 \quad W_r y_r & W_r A_{22}^{(r)} W_r \end{bmatrix} = \begin{bmatrix} A_{11}^{(r+1)} & A_{12}^{(r+1)} \\ 0 \quad y_{r+1} & A_{22}^{(r+1)} \end{bmatrix}$$

- 把A变成Hessenberg矩阵（拟上三角矩阵）的目是减少QR方法的计算量；
- 把A变成Hessenberg矩阵（拟上三角矩阵）能够减少QR方法计算量的主要原因是

1. 对拟上三角矩阵作QR分解时，Q一定是拟上三角矩阵；
2. RQ（$=A_{k+1}$）的乘积为拟上三角矩阵。

【例】

用Householder方法把矩阵A= $\begin{pmatrix} 5 & 1 & 2 & 2 \\ -2 & 6 & -1 & 3 \\ 1 & 3 & 4 & -1 \\ 2 & -4 & -2 & 6 \end{pmatrix}$ 约化为

y_1

上Hessenberg矩阵.

【解】第一步约化. (1)构造 $W_1 = I - h_1^{-1} u_1 u_1^T$，使得 $W_1 y_1 = c_1 e_1$.

$$c_1 = -\mathrm{sgn}(a_{21}) \| y_1 \|_2 = 3, \quad u_1 = y_1 - c_1 e_1 = (-5,1,2)^T,$$

$$h_1 = \frac{1}{2} \| u_1 \|_2^2 = 15, \quad W_1 = \frac{1}{15} \begin{pmatrix} -10 & 5 & 10 \\ 5 & 14 & -2 \\ 10 & -2 & 11 \end{pmatrix} \quad W_1 y_1 = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix},$$

(2)约化计算：令$H_1=\begin{pmatrix}1 & 0\\ 0 & W_1\end{pmatrix}$,

$$A^{(2)}=H_1AH_1=\begin{pmatrix}5 & 1.33333333 & 1.93333333 & 1.96666667\\ 3 & 5.11111111 & -1.48888889 & -2.64444444\\ 0 & -2.82222222 & 5.29777778 & 2.52888889\\ 0 & 3.02222222 & -3.13777778 & 5.95111111\end{pmatrix}$$

第二步约化.

y_2

(1)构造$W_2=I-h_2^{-1}u_2u_2^T$,使得$W_2y_2=c_2e_1$. $=(c_2,0)^T$

$$c_2=-\mathrm{sgn}(a_{32})\|y_2\|_2=4.13506534,$$

$$u_2=y_2-c_2e_1=(-6.95728757,3.02222222)^T,$$

$$\boldsymbol{h}_2=\frac{1}{2}\|\boldsymbol{u}_2\|_2^2=28.7688387,$$

$$\boldsymbol{W}_2=\boldsymbol{I}-\boldsymbol{h}_2^{-1}\boldsymbol{u}_2\boldsymbol{u}_2^T=\begin{pmatrix}-0.682509702 & 0.730876532\\ 0.730876532 & 0.682509702\end{pmatrix},$$

(2)约化计算：令$\boldsymbol{H}_2=\begin{pmatrix}\boldsymbol{I}_2 & 0\\ 0 & \boldsymbol{W}_2\end{pmatrix}$,

$$\boldsymbol{A}^{(3)}=\boldsymbol{H}_2\boldsymbol{A}^{(2)}\boldsymbol{H}_2=\begin{pmatrix}5 & 1.33333333 & 0.04478410 & 2.68704607\\ 3 & 5.11111111 & -0.91658127 & -2.89305284\\ 0 & 4.13506535 & 5.75820296 & 2.70782190\\ 0 & 0 & 2.95884478 & 5.13068592\end{pmatrix}$$

同QR分解一样，在计算$A^{(n-1)}$时，不必要计算出H_i

记 $\mathbf{A}^{(1)}=\mathbf{A}$，并记 $\mathbf{A}^{(r)}$ 的第 r 列至第 n 列的元素为 $a_{ij}^{(r)}(i=1,2,\cdots,n;j=r,r+1,\cdots,n)$。

对于 $r=1,2,\cdots,n-2$ 执行

(1) 若 $a_{ir}^{(r)}(i=r+2,r+3,\cdots,n)$全为零，则令 $\mathbf{A}^{(r-1)}=\mathbf{A}^{(r)}$，转(5)；否则转(2)。

(2) 计算

$$d_r=\sqrt{\sum_{i=r+1}^{n}(a_{ir}^{(r)})^2}$$

$$c_r=-\operatorname{sgn}(a_{r+1,r}^{(r)})\,d_r \quad (\text{若 } a_{r+1,r}^{(r)}=0,\text{则取 } c_r=d_r)$$

$$h_r=c_r^2-c_r a_{r+1,r}^{(r)}$$

(3) 令 $\boldsymbol{u}_r=(0,\cdots,0,a_{r+1,r}^{(r)}-c_r,a_{r+2,r}^{(r)},\cdots,a_{nr}^{(r)})^{\mathrm{T}}\in\mathbf{R}^n$。

(4) 计算

$$
\begin{aligned}
\boldsymbol{p}_r &= \mathbf{A}^{(r)\mathrm{T}}\boldsymbol{u}_r/h_r \\
\boldsymbol{q}_r &= \mathbf{A}^{(r)}\boldsymbol{u}_r/h_r \\
t_r &= \boldsymbol{p}_r^{\mathrm{T}}\boldsymbol{u}_r/h_r \\
\boldsymbol{\omega}_r &= \boldsymbol{q}_r - t_r\boldsymbol{u}_r \\
\mathbf{A}^{(r+1)} &= \mathbf{A}^{(r)} - \boldsymbol{\omega}_r\boldsymbol{u}_r^{\mathrm{T}} - \boldsymbol{u}_r\boldsymbol{p}_r^{\mathrm{T}}
\end{aligned}
$$

(5) 继续。

当此算法执行完后，就得到与原矩阵 $\mathbf{A}$ 相似的拟上三角矩阵 $\mathbf{A}^{(n-1)}$

小　结

一、矩阵的QR分解

定理： 设A是一个n阶实方阵，那么A可分解为一个正交矩阵Q和一个上三角矩阵R的乘积，即：

$$A=QR.$$

二、矩阵的拟上三角分解

定理： Householder(豪斯霍尔德)约化矩阵为上Hessenberg矩阵

设$A\in R^{n\times n}$是一个n阶实方阵,则存在初等反射矩阵$H_1,H_2,\cdots,H_{n-1}$,使得

$$H_{n-1}\cdots H_2H_1AH_1H_2\cdots H_{n-1}=H_0^TAH_0=H \text{ (上Hessenberg矩阵)}.$$

作业

1. ①设 A 是对称矩阵，λ 和 x $(\|x\|_2=1)$ 是 A 的特征值及对应的特征向量，又设 H 为一个正交矩阵，使得

$$Hx = e_1 = (1,0,\cdots,0)^T$$

证明：$B=HAH^T$ 的第一行第一列除了 λ 外其余均为 0.

② 对于矩阵

$$A=\begin{pmatrix} 2 & 10 & 2 \\ 10 & 5 & -8 \\ 2 & -8 & 11 \end{pmatrix},$$

$\lambda=9$ 是其特征值，$x=(\frac{2}{3},\frac{1}{3},\frac{2}{3})^T$ 为其对应特征向量，

求一初等反射矩阵 H，s.t. $Hx=e_1$，并计算 $B=HAH^T$.

2. 用初等反射变换将

$$A=\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ 2 & -4 & 5 \end{pmatrix}$$

分解为 QR 形式. Q 正交，R 上三角.

Householder 变换.