第十一讲 矩阵的 QR 分解

一．Givens 矩阵与 Givens 变换

1．定义：设实数 c 与 s 满足$c^2+s^2=1$，称

$$
T_{ij}=\begin{array}{c}\left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & c & & & & s & & & \\
 & & & & 1 & & & & & & \\
 & & & \vdots & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & -s & & & & c & & & \\
 & & & & & & & & 1 & & \\
 & & & \vdots & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]\end{array}
\begin{array}{l}
\\ \\ \\ \leftarrow(i) \\ \\ \\ \\ \leftarrow(j) \\ \\ \\ \\
\end{array}
$$

$(i<j)$

为 Givens 矩阵（初等旋转矩阵），也记作$T_{ij}=T_{ij}(c,s)$。由 Givens 矩阵所确定的线性变换称为 Givens 变换（初等旋转变换）。

说明：（1）实数$c^2+s^2=1$，故存在θ，使$c=\cos(\theta),s=\sin(\theta)$。

（2）$y=T_{ij}x$中T_{ij}确定了将向量 x 变成 y 的一种变换，正是 Givens 变换。二阶情况下，$y=\begin{bmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{bmatrix}x$ 确定的正是平面直角坐标系中绕原点的一个旋转变换（旋转θ度）。

（3）以上实 Givens 矩阵也可推广称为复初等旋转矩阵。

$$
U_{ik} = \left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & ce^{j\theta_1} & & & & se^{j\theta_2} & & & \\
& & & & 1 & & & & & & \\
& & & \vdots & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & -se^{j\theta_3} & & & & ce^{j\theta_4} & & & \\
& & & & & & & & 1 & & \\
& & & \vdots & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{array}\right]
\begin{array}{l}
\\
\\
\\
\leftarrow (i) \\
\\
\\
\\
\leftarrow (k) \\
\\
\\
\\
\end{array}
$$

其中 c 与 s 仍为满足$c^2+s^2=1$的实数，$\theta_1,\theta_2,\theta_3,\theta_4$为实角度。

显然，$\det(U_{ik})=c^2e^{j(\theta_1+\theta_4)}+s^2e^{j(\theta_2+\theta_3)}$

当$\theta_1+\theta_4=\theta_2+\theta_3$时，$\det(U_{ik})=e^{j(\theta_1+\theta_4)}$

当$\theta_1+\theta_4=\theta_2+\theta_3=2n\pi$时，$\det(U_{ik})=1$

2. 性质

（1）$\left[T_{ij}(c,s)\right]^{-1}=\left[T_{ij}(c,s)\right]^T=T_{ij}(c,-s)$，$-s=-\sin(\theta)=\sin(-\theta)$,

旋转θ度再反向旋转度θ

$$\det\left[T_{ij}\left(c,s\right)\right]=1$$

（2）设$x=\begin{bmatrix}\xi_1 & \xi_2 & \cdots & \xi_n\end{bmatrix}^T$，$y=T_{ij}x=\begin{bmatrix}\eta_1 & \eta_2 & \cdots & \eta_n\end{bmatrix}^T$，则有

$$\begin{cases} \eta_i = c\xi_i + s\xi_j \\ \eta_j = -s\xi_i + c\xi_j \\ \eta_k = \xi_k \qquad (k \neq i, j) \end{cases}$$

当$\xi_i^2 + \xi_j^2 \neq 0$时，总可以选$c = \dfrac{\xi_i}{\sqrt{\xi_i^2 + \xi_j^2}}$，$s = \dfrac{\xi_j}{\sqrt{\xi_i^2 + \xi_j^2}}$使

$$\begin{cases} \eta_i = \sqrt{\xi_i^2 + \xi_j^2} \\ \eta_j = 0 \end{cases} \rightarrow T_{ij}x = \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \sqrt{\xi_i^2 + \xi_j^2} & \cdots & 0 & \cdots & \xi_n \end{bmatrix}^T$$

定理 1. 设$x = \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_n \end{bmatrix}^T \neq 0$，则存在有限个 Givens 矩阵的乘积 T，使得$Tx = |x|e_1$

说明：（1）$|x|=\sqrt{\|x\|_2^2}=\sqrt{x^T x}$ (x 为实数时)，$|x|=\sqrt{x^H x}$（x 为复数时）。

（2）$e_1=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\end{bmatrix}^T$

[证明]：$\xi_1 \neq 0$ 的情形

（1） 构造 $T_{12}(c,s): c=\dfrac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}}, s=\dfrac{\xi_2}{\sqrt{\xi_1^2+\xi_2^2}}$

$$T_{12}x=\begin{bmatrix}\sqrt{\xi_1^2+\xi_2^2} & 0 & \xi_3 & \xi_4 & \cdots & \xi_n\end{bmatrix}^T$$

（2） 对 $T_{12}x$ 再考虑

（3） $T_{13}(c,s): c=\frac{\sqrt{\xi_1^2+\xi_2^2}}{\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}}, s=\frac{\xi_3}{\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}}$

$$T_{13}T_{12}x=\begin{bmatrix}\sqrt{\xi_1^2+\xi_2^2+\xi_3^2} & 0 & 0 & \xi_4 & \cdots & \xi_n\end{bmatrix}^T$$

（4） 依此类推，构造

$$T_{1k}(c,s): c=\frac{\sqrt{\xi_1^2+\cdots+\xi_k^2}}{\sqrt{\xi_1^2+\xi_2^2+\cdots+\xi_k^2}}, s=\frac{\xi_k}{\sqrt{\xi_1^2+\xi_2^2+\cdots+\xi_k^2}}$$

(k=2,3,…..n)

$$T_{1k}\left\langle T_{1k-1}\left\{T_{1k-2}\left[\cdots T_{13}\left(T_{12}x\right)\right]\right\}\right\rangle$$

$$=\begin{bmatrix}\sqrt{\xi_1^2+\xi_2^2+\cdots+\xi_k^2} & 0 & 0 & \cdots & 0 & \xi_{k+1} & \cdots & \xi_n\end{bmatrix}^T$$

直至 k=n。令 $T=T_{1n}T_{1n-1}T\cdots T_{12}$，则有

$$Tx=\begin{bmatrix}\sqrt{\xi_1^2+\xi_2^2+\cdots\xi_n^2} & 0 & 0 & \cdots & 0\end{bmatrix}^T=|x|e_1$$

$\xi_1=0$的情形，从第一个不为零的ξ_i开始运用上述方法即可

推论：对于任何非零列向量$x\in R^n$及任何单位列向量$z(|z|=1)$，均存在着有限个 Givens 矩阵的乘积 T，使$Tx=|x|z$。

[证明]：由上述定理，对 x 存在有限个 Givens 矩阵$T_{12}^{(1)},T_{13}^{(1)},\cdots,T_{1n}^{(1)}$的乘积

$$T^{(1)} = T_{1n}^{(1)} T_{1n-1}^{(1)} \cdots T_{13}^{(1)} T_{12}^{(1)}，\text{使} T^{(1)} x = |x| e_1$$

对 z 同理存在有限个 Givens 矩阵 $T_{12}^{(2)}, T_{13}^{(2)}, \cdots, T_{1n}^{(2)}$ 的乘积

$$T^{(2)} = T_{1n}^{(2)} T_{1n-1}^{(2)} \cdots T_{13}^{(2)} T_{12}^{(2)}，\text{使} T^{(2)} z = |z| e_1 = e_1$$

即，

$$T^{(1)} x = |x| T^{(2)} z = T^{(2)} (|x| z) \to \left(T^{(2)}\right)^{-1} T^{(1)} x = |x| z$$

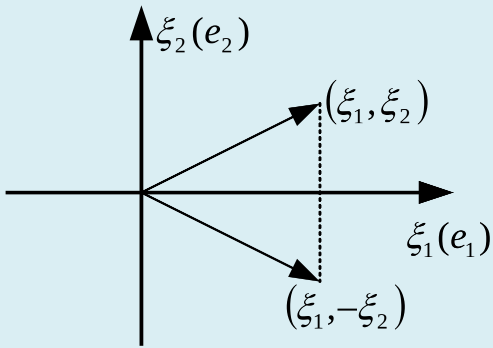
$$\Rightarrow \left(T_{1n}^{(2)} T_{1n-1}^{(2)} \cdots T_{12}^{(2)}\right)^{-1} \left(T_{1n}^{(1)} T_{1n-1}^{(1)} \cdots T_{12}^{(1)}\right) x = |x| z$$

其中

$$\left(T_{1n}^{(2)}T_{1n-1}^{(2)}\cdots T_{12}^{(2)}\right)^{-1}\left(T_{1n}^{(1)}T_{1n-1}^{(1)}\cdots T_{12}^{(1)}\right)$$
$$=\left(T_{12}^{(2)}\right)^{-1}\left(T_{13}^{(2)}\right)^{-1}\cdots\left(T_{1n}^{(2)}\right)^{-1}T_{1n}^{(1)}T_{1n-1}^{(1)}\cdots T_{12}^{(1)}$$
$$=\left(T_{12}^{(2)}\right)^{T}\left(T_{13}^{(2)}\right)^{T}\cdots\left(T_{1n}^{(2)}\right)^{T}T_{1n}^{(1)}T_{1n-1}^{(1)}\cdots T_{12}^{(1)}$$

为有限个 Givens 矩阵的乘积。

二 、Householder 矩阵与 Householder 变换



平面直角坐标系中，将向量$x=\begin{bmatrix}\xi_1 & \xi_2\end{bmatrix}^T$关于$e_1$轴作镜像变换，则得到

$$y=\begin{bmatrix}\xi_1\\-\xi_2\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}\xi_1\\\xi_2\end{bmatrix}=\left(I-2e_2e_2^T\right)x=Hx$$

一般地，可将其推广

1. 定义：设单位列向量$u\in R^n$，称$H=I-2uu^T$为 Householder 矩阵（初等反射矩阵），由 Householder 矩阵所确定的线性变换（$y=Hx$）称为 Householder 变换

2．性质

（1）$H^T=H$（实对称），(2) $H^{-1}=H^T$（正交），(3) $H^2=I$（对合），

(4) $H^{-1}=H$ （自逆），(5) $\det H=-1$

为证明第 5 条，可利用如下引理。

引理：设 $A\in R^{m\times n}, B\in R^{n\times m}$，则 $\det\left(I_m+AB\right)=\det\left(I_n+BA\right)$

[证明]：参考如下的分块矩阵 $\begin{bmatrix} I_n & B \\ -A & I_m \end{bmatrix}$ 的行列式，有

$$\det\begin{bmatrix} I_n & B \\ -A & I_m \end{bmatrix}=\det\begin{bmatrix} I_n & B \\ 0 & I_m+AB \end{bmatrix}=\det\left(I_m+AB\right)=\det\begin{bmatrix} I_n+BA & B \\ 0 & I_m \end{bmatrix}$$

$$=\det\left(I_n+BA\right)$$

故，$\det\left(H\right)=\det\left[I+\left(-2uu^T\right)\right]=1+u^T(-2u)=1-2=-1$

定理 2. 对于任何非零列向量 $x \in R^n$ 及任何单位列向量 $z \in R^n$，存在 Householder 矩阵 H，使得 $Hx = |x| z$。

[证明] 当 $x = |x| z$ 时，选 u 满足 $u^T x = 0$，则 $Hx = \left(I - 2uu^T\right) x = x = |x| z$

当 $x \neq |x| z$ 时，选 $u = \dfrac{x - |x| z}{\left|x - |x| z\right|}$，有

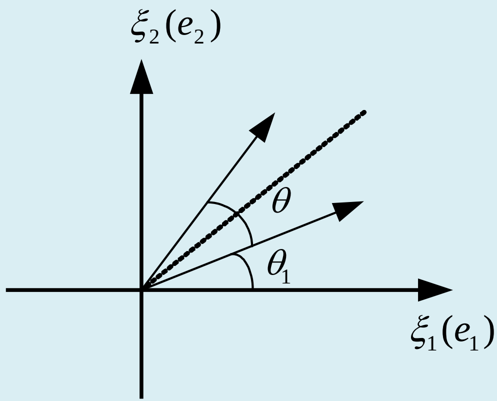
$$\left|x - |x| z\right|^2 = \left(x - |x| z\right)^T \left(x - |x| z\right) = x^T x + |x|^2 z^T z - |x| \left(z^T x + x^T z\right)$$

$$= 2\left(x^T x - |x| z^T x\right) = 2\left(x - |x| z\right)^T x$$

$$Hx = \left(I - 2 \frac{(x - |x| z)}{\left|x - |x| z\right|} \frac{(x^T - |x| z^T)}{\left|x - |x| z\right|} \right) x = x - \left(x - |x| z\right) = |x| z$$

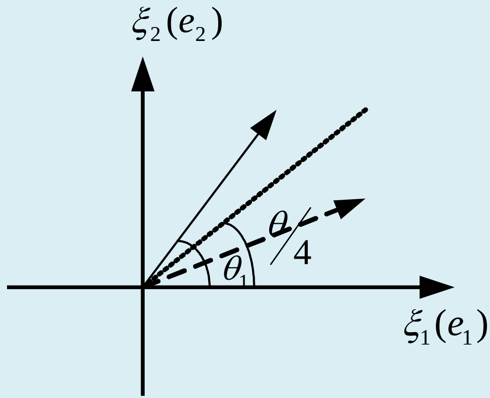
定理 3. 初等旋转矩阵（Givens 矩阵）是两个初等反射矩阵的乘积。

证明参见 p_{202}，较容易。我们这里主要是给出一种几何解释。



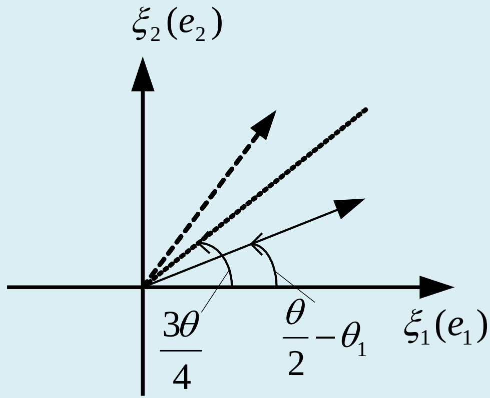
从表明上看，似乎一种反射变换即可代替旋转变换。实际上是不对的，

因为这样的反射变换对应的对称轴沿$\left(\theta_1+\theta/2\right)$方向，与$\theta_1$有关。



实际上，旋转变换可由这样两次反射变换的作用来代替。

首先，关于沿$(\theta/4)$对称轴作反射变换，则原向量沿θ_1方向转至$(-\theta_1+\theta/2)$。



其次，关于沿 $(3\theta/4)$ 对称轴作反射变换，则向量反射至

$$\left(\frac{\theta}{2}-\theta_1\right)+2\left(\frac{\theta}{4}+\theta_1\right)=\theta_1+\theta。$$

正是原向量沿 θ_1 方向转 θ 的结果。

- 旋转变换可用两个反射变换的连续作用来代替，即$T_{ij}=H_vH_u$。但是反射变换却不可能用多个旋转变换的连续作用来代替。这是因为$\det\left(T_{ij}\right)\equiv 1,\det\left(H\right)\equiv -1$。由两个－1 的乘积可得 1，但多个 1 的乘积只能是 1，不是－1。

三、 QR 分解

1. 定义：如果实（复）矩阵 A 可化为正交（酉）矩阵 Q 与实（复）上三角矩阵 R 的乘积，即$A=QR$，则称上式为 A 的 QR 分解。

2. 定理：设 A 为列满秩矩阵，则 A 存在 QR 分解。

两种情况：（1）非奇异方阵

（2）m*n 阶矩阵

3. 求 QR 分解的方法

[方法一]采用 Givens 方法

将 n 阶非奇异矩阵 A 写为

$A=\begin{bmatrix} b^{(1)^T} \\ * \end{bmatrix}^T$ 或者[b$^{(1)}$ *], 则存在有限个 Givens 矩阵的乘积T_1，使得

$$T_1 b^{(1)} = \left| b^{(1)} \right| e_1 \rightarrow T_1 A = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & & & \\ \vdots & & A^{(1)} & \\ 0 & & & \end{bmatrix}$$

$A^{(1)}$ 写成 $A^{(1)} = \begin{bmatrix} b^{(2)^T} \\ * \end{bmatrix}^T$ $\rightarrow$ 存在 T_2，使得

或 $[b^{(2)} \quad *]$

$$T_2 A^{(1)} = \begin{bmatrix} a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & & & \\ \vdots & & A^{(2)} & \\ 0 & & & \end{bmatrix}$$

$\vdots$

$$A^{(n-2)} \quad \text{写成} \; A^{(n-2)} = \begin{bmatrix} b^{(n-1)^T} \\ * \end{bmatrix}^T \rightarrow \text{存在} \; T_{n-1}\text{，使得}$$

$$T_{n-1}A^{(n-2)} = \begin{bmatrix} a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} \\ 0 & a_{nn}^{(n-1)} \end{bmatrix}$$

$$\text{令} \; T = \begin{bmatrix} I_{n-2} & 0 \\ 0 & T_{n-1} \end{bmatrix}\begin{bmatrix} I_{n-3} & 0 \\ 0 & T_{n-2} \end{bmatrix}\cdots\begin{bmatrix} I_2 & 0 \\ 0 & T_3 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & T_2 \end{bmatrix}T_1\text{，则有}$$

$$TA = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{n-1} \end{bmatrix} = R \rightarrow A = T^{-1}R = QR$$

其中，R 为上三角矩阵，Q=T^{-1}为正交矩阵

[方法二]采用 Householde 方法

$A=\begin{bmatrix} b^{(1)} & * \end{bmatrix}$，存在$H_1$，使得

$$H_1 b^{(1)} = \left|b^{(1)}\right| e_1 \rightarrow H_1 A = \left[\begin{array}{c|c} \left|b^{(1)}\right| & * \\ \hline 0 & \\ \vdots & A^{(1)} \\ 0 & \end{array}\right]$$

$A^{(1)}=\begin{bmatrix} b^{(2)} & * \end{bmatrix}$，存在$H_2$，使得

$$H_2A^{(1)}=\begin{bmatrix}\left|b^{(2)}\right| & * \\ 0 & \\ \vdots & A^{(2)} \\ 0 & \end{bmatrix}$$

$$\cdots\cdots$$

经过 n-1 次变换，得

$A^{(n-2)}=\begin{bmatrix}b^{(n-1)} & *\end{bmatrix}$，存在 H_{n-1}，使得 $H_{n-1}A^{(n-2)}=\begin{bmatrix}\left|b^{(n-1)}\right| & * \\ 0 & \\ \vdots & A^{(n-1)} \\ 0 & \end{bmatrix}$

$$\begin{bmatrix} I_{n-2} & 0 \\ 0 & H_{n-1} \end{bmatrix} \begin{bmatrix} I_{n-3} & 0 \\ 0 & H_{n-2} \end{bmatrix} \cdots \begin{bmatrix} I_2 & 0 \\ 0 & H_3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & H_2 \end{bmatrix} H_1 A$$

$$= \begin{bmatrix} b^{(1)} & & & * \\ & b^{(2)} & & \\ & & \ddots & \\ & & & b^{(n-1)} \end{bmatrix} = R$$

令 $S = \begin{bmatrix} I_{n-2} & 0 \\ 0 & H_{n-1} \end{bmatrix} \begin{bmatrix} I_{n-3} & 0 \\ 0 & H_{n-2} \end{bmatrix} \cdots \begin{bmatrix} I_2 & 0 \\ 0 & H_3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & H_2 \end{bmatrix} H_1$

$$H_{l+1} = I_{n-l} - 2uu^T \left(u \in R^{n-1}, u^T u = 1\right)$$

则

$$\begin{bmatrix} I_l & 0 \\ 0 & H_{l+1} \end{bmatrix} = \begin{bmatrix} I_l & 0 \\ 0 & I_{n-l} \end{bmatrix} - 2\begin{bmatrix} 0 & 0 \\ 0 & uu^T \end{bmatrix} = I_n - 2\begin{bmatrix} 0 \\ u \end{bmatrix}\begin{bmatrix} 0^T & u^T \end{bmatrix} = I_n - 2uu^T$$

$$SA = \begin{bmatrix} b^{(1)} & & & * \\ & b^{(2)} & & \\ & & \ddots & \\ & & & b^{(n-1)} \end{bmatrix} = R，\ S^{-1} = Q \text{为正交矩阵}$$

以上两种方法中的前一种方法可推广到复矩阵的情况。

3. Gram-schmidt 正交归一化方法

$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$，各列向量线性无关可进行正交化

$$q_1 = {a_1}/{|a_1|} \qquad 令\ y_1 = a_1$$

$$q_2 = \frac{a_2 + k_{21}q_1}{|a_2 + k_{21}q_1|}, k_{21} = -\langle a_2, q_1 \rangle \qquad 令\ y_2 = a_2 + k_{21}q_1$$

$$\vdots$$

$$q_l = \frac{a_l + \sum_{j=1}^{l-1} k_{lj}q_j}{\left|a_l + \sum_{j=1}^{l-1} k_{lj}q_j\right|}, k_{lj} = -\langle a_l, q_j \rangle \qquad 令\ y_l = a_l + \sum_{j=1}^{l-1} k_{lj}q_j$$

$$q_i^H q_j = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \rightarrow Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}，满足 Q^T Q = I$$

改写：

$$a_1 = q_1 \left| y_1 \right|$$

$$a_2 = q_2 \left| y_2 \right| - k_{21} q_1$$

$$\vdots$$

$$a_l = q_l \left| y_l \right| - \sum_{j=1}^{l-1} k_{lj} q_j$$

$$
\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} \begin{bmatrix} |y_1| & -k_{21} & -k_{31} & \cdots & -k_{n1} \\ & |y_2| & -k_{32} & \cdots & -k_{n2} \\ & & |y_3| & \cdots & \\ & & & \ddots & \vdots \\ & & & & |y_n| \end{bmatrix}
$$

$$
= QR
$$

作业：p219-220，1、7、8