8.3 Householder变换与矩阵的正交分解

一、初等反射阵(Householder变换阵)

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

$$H = I - 2WW^T$$

的n阶方阵称为初等反射阵,或称为*Householder*变换阵.

$$H = \begin{bmatrix} 1-2w_1^2 & -2w_1w_2 & \cdots & -2w_1w_n \\ -2w_2w_1 & 1-2w_2^2 & \cdots & -2w_2w_n \\ \cdots & \cdots & \cdots & \cdots \\ -2w_nw_1 & -2w_nw_2 & \cdots & 1-2w_n^2 \end{bmatrix}$$

例：$W=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}^T \in R^3, \|W\|_2 = 1$

$$H = I - 2WW^T = I - 2\begin{bmatrix}\frac{1}{\sqrt{2}}\\ 0\\ \frac{1}{\sqrt{2}}\end{bmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}$$

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

H阵的性质：

（1）非奇异 $\det(H)=1-2W^TW=-1$

（2）对称正交

$$H=H^T$$

$$HH^T=H^2=(I-2WW^T)(I-2WW^T)$$

$$=I-4WW^T+4WW^TWW^T=I$$

$$H=\begin{bmatrix} 1-2w_1^2 & -2w_1w_2 & \cdots & -2w_1w_n \\ -2w_2w_1 & 1-2w_2^2 & \cdots & -2w_2w_n \\ \cdots & \cdots & \cdots & \cdots \\ -2w_nw_1 & -2w_nw_2 & \cdots & 1-2w_n^2 \end{bmatrix}$$

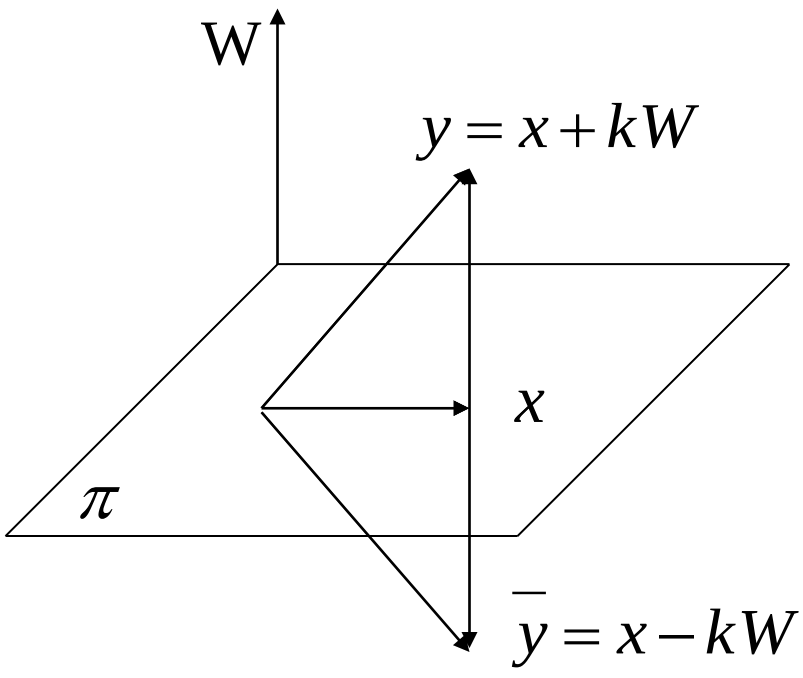
（3）镜映射－几何意义

平面π 方程 $W^T x = 0 \quad \forall x \in \pi$

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

若 $y \notin \pi$ ，$Hy = H(x + kW) = x + k(I - 2WW^T)W$

$$= x + kW - 2kWW^T W = x - kW = \bar{y}$$

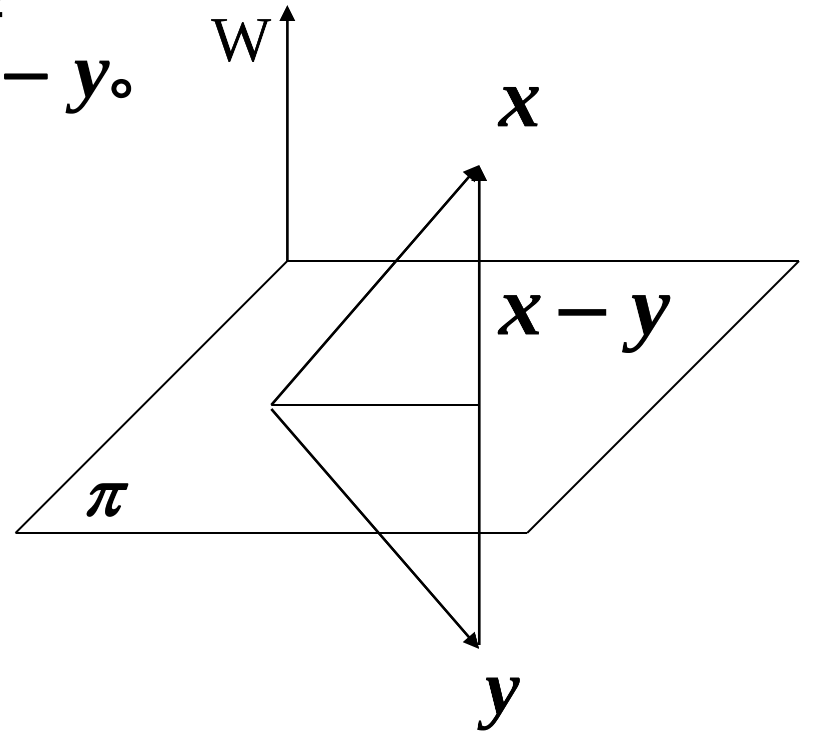


H阵的作用：

定理 设两个不相等的n维向量$x, y \in R^n, x \neq y,$ 但$\|x\|_2 = \|y\|_2$,则存在householder阵

$$H = I - 2\frac{UU^T}{\|U\|_2^2}$$

使$Hx = y$，其中$U = x - y$。



证：若设$W=\dfrac{U}{\|U\|_2}$，则有$\|W\|_2=1$，因此

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

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

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

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

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

代入上式后即得到$Hx=y$

$$\because\ x^Tx=y^Ty,\quad x^Ty=y^Tx$$

1. *Householder*变换可以将给定的向量变为一个与任一个$e_i \in R^n (i = 1, 2, \cdots, n)$同方向的向量。

即：$\forall x = (x_1, x_2, \cdots, x_n)^T \in R^n, x \neq 0,$ 可构造H阵，

使$Hx = y = -\sigma_i e_i = (0, ..., 0, -\sigma_i, 0, \cdots, 0)^T \in R^n$

其中$\sigma_i = sign(x_i)\|x\|_2 = sign(x_i)(\sum_{k=1}^{n} x_k^2)^{\frac{1}{2}},$

$$sign(x_i) = \begin{cases} 1 & x_i \geq 0 \\ -1 & x_i < 0 \end{cases}$$

$$U = x - y = x + \sigma_i e_i = (x_1, \cdots, x_i + \sigma_i, \cdots, x_n)^T,$$

构造初等反射阵

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

有 $Hx = y = -\sigma_i e_i$

其中

$$\rho = \frac{1}{2}U^T U = \frac{1}{2}(x_1^2 + \ldots + (x_i + \sigma_i)^2 + \cdots + x_n^2)$$

$$= \frac{1}{2}(2x_i\sigma_i + 2\sigma_i^2) = \sigma_i(x_i + \sigma_i)$$

例 已知向量$x=(2,0,2,1)^T$,试构造$Householder$阵,使$Hx=Ke_3$,其中$e_3=(0,0,1,0)^T\in R^4$, $K\in R$。

解: $\sigma_3=sign(x_3)\|x\|_2=\sqrt{4+0+4+1}=3$,因$x_3=2>0$,

故取$K=-\sigma_3=-3$ 于是$y=-\sigma_3 e_3=Ke_3=(0,0,-3,0)^T$,

$U=x-y=(2,0,5,1)^T$, $\rho=\sigma_3(\sigma_3+x_3)=3(3+2)=15$

$$\boxed{\rho=\frac{1}{2}U^TU}$$

$$H=I-\frac{1}{\rho}UU^T=\frac{1}{15}\begin{bmatrix}11 & 0 & -10 & -2\\ 0 & 1 & 0 & 0\\ -10 & 0 & -10 & -5\\ -2 & 0 & -5 & 14\end{bmatrix}$$

2. 构造H阵，将向量$x=(x_1,\cdots,x_k,x_{k+1},\cdots,x_n)^T$的后面$n-k$个分量约化为零$(1\le k<n)$。

即：任给定$x=(x_1,x_2,\cdots,x_n)^T\neq 0$，构造$H_k\in R^{n\times n}$，使

$$H_k x=(x_1,x_2,\cdots,x_{k-1},-\sigma_k,0,\cdots,0)^T$$

推导：$\forall x=(x_1,x_2,\cdots,x_n)^T\neq 0$

$$y=(x_1,\cdots,x_{k-1},-\sigma_k,0,\cdots,0)^T$$

$$\sigma_k=sign(x_k)\left(\sum_{i=k}^{n}x_i^2\right)^{\frac{1}{2}},\qquad sign(x_k)=\begin{cases}1 & x_k\ge 0\\ -1 & x_k<0\end{cases}$$

$$U^{(k)} = x - y = (0, \cdots, 0, x_k + \sigma_k, x_{k+1}, \cdots, x_n)^T$$

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

其中 $\rho_k = \frac{1}{2} U^{(k)T} U^{(k)} = \sigma_k (\sigma_k + x_k)$

特别，取k = 1.

$\forall x = (x_1, x_2, \cdots, x_n)^T \in R^n, x \neq 0,$ 可构造H阵，

使$Hx = y = -\sigma_1 e_1 = (-\sigma_1, 0, \cdots, 0)^T \in R^n$

其中 $\sigma_1 = sign(x_1) \|x\|_2 = sign(x_1) (\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}},$

$$sign(x_1) = \begin{cases} 1 & x_1 \geq 0 \\ -1 & x_1 < 0 \end{cases}$$

$$U^{(1)} = x + \sigma_1 e_1 = (\sigma_1 + x_1, x_2, \cdots, x_n)^T,$$

可构造初等反射阵

$$H_1 = I - 2WW^T = I - 2\frac{U_1 U_1^T}{\|U_1\|^2} = I - \frac{1}{\rho} U_1 U_1^T$$

有$H_1 x = y = -\sigma_1 e_1$

其中 $\rho_1 = \frac{1}{2} U_1^T U_1 = \frac{1}{2}((x_1 + \sigma_1)^2 + x_2^2 + \ldots + x_n^2)$

$$= \frac{1}{2}(2x_1\sigma_1 + 2\sigma_1{}^2) = \sigma_1(x_1 + \sigma_1)$$

例：已知向量$x=(2,2,1)^T$，试构造初等反射阵使$y=Hx$最后一个元素为零。

解 $k=2$，构造H_2

$$\sigma_2 = sign(x_2)(x_2^2+x_3^2)^{\frac{1}{2}} = \sqrt{5}$$

$$U^{(2)} = (0, \sigma_2+x_2, x_3)^T = (0, 2+\sqrt{5}, 1)^T$$

$$\rho_2 = \sigma_2(x_2+\sigma_2) = 5+2\sqrt{5}$$

于是 $H_2x = (x_1, -\sigma_2, 0)^T = (2, -\sqrt{5}, 0)^T$

计算 H_2，$$H_2 = I - \frac{1}{\rho_2}U^{(2)}(U^{(2)})^T$$

$$H_2 = \frac{1}{5+2\sqrt{5}} \begin{bmatrix} 5+2\sqrt{5} & 0 & 0 \\ 0 & -(4+2\sqrt{5}) & -(2+\sqrt{5}) \\ 0 & -(2+\sqrt{5}) & (4+2\sqrt{5}) \end{bmatrix}$$

$$H_2 x = (x_1, -\sigma_2, 0)^T = (2, -\sqrt{5}, 0)^T$$

二、矩阵的正交分解

1、正交分解的基本定理

定理 $\forall A \in R^{m\times n}$是列满秩矩阵（$m > n, r(A) = n$），存在分解式$A = QR$，其中$Q \in R^{m\times n}$列法正交矩阵，$R \in R^{n\times n}$非奇异上三角阵。若限定$R$阵对角元符号，则分解式是唯一的。

当$m = n$时，$Q \in R^{n\times n}$正交阵，$R \in R^{n\times n}$非奇异上三角阵。

$$A_{n\times n}=QR=\begin{bmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{bmatrix}\begin{bmatrix} * & \cdots & * \\ & \ddots & \vdots \\ 0 & & * \end{bmatrix}$$

$$A_{m\times n}=QR=\begin{bmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{bmatrix}_{m\times n}\begin{bmatrix} * & \cdots & * \\ & \ddots & \vdots \\ 0 & & * \end{bmatrix}_{n\times n}$$

$$A_{m\times n}=QR=\begin{bmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{bmatrix}_{m\times m}\begin{bmatrix} * & \cdots & * \\ & \ddots & \vdots \\ 0 & & * \\ 0 & \cdots & 0 \end{bmatrix}_{m\times n}$$

称形如

$$H=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n-1} & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n-1} & h_{2n} \\ & h_{32} & h_{33} & \cdots & h_{3n} \\ & & \ddots & \ddots & \vdots \\ & & & h_{nn-1} & h_{nn} \end{bmatrix}$$

的矩阵为上海森堡（Hessenberg) 阵。如果此对角线元$h_{ii-1}(i=2,3,\cdots,n)$全不为零,则称该矩阵为不可约的上Hessenberg矩阵。

讨论用 **Householder** 变换将一般矩阵A相似变换成Hessenberg阵

例：用豪斯霍尔德方法将

$$\boldsymbol{A} = \boldsymbol{A}_1 = \begin{bmatrix} -4 & -3 & -7 \\ 2 & 3 & 2 \\ 4 & 2 & 7 \end{bmatrix}$$

矩阵约化为上**Hessenberg**阵。

解：选取初等反射阵$\mathbf{R}_1$使

$$R_1 c_1 = -\sigma_1 e_1, \text{其中} c_1 = (2,4)^T$$

(1)计算$\boldsymbol{R}_1$: $\alpha = \max(2,4) = 4, \boldsymbol{c}_1 \to \boldsymbol{c}_1' = (0.5,1)^T$

$$\begin{cases} \sigma = \sqrt{1.25} = 1.118034, \\ \boldsymbol{u}_1 = \boldsymbol{c}_1' + \sigma \boldsymbol{e}_1 = (1.618034, 1)^T, \\ \beta_1 = \sigma(\sigma + 0.5) = 1.809017, \\ \sigma_1 = \alpha\sigma = 4.472136, \\ \boldsymbol{R}_1 = \boldsymbol{I} - \beta_1^{-1} \boldsymbol{u}_1 \boldsymbol{u}_1^T. \end{cases}$$

则有

$$\boldsymbol{R}_1 \boldsymbol{c}_1 = -\sigma_1 \boldsymbol{e}_1.$$

（2）约化计算：

令
$$\boldsymbol{U}_1=\begin{bmatrix}1 & 0\\ 0 & \boldsymbol{R}_1\end{bmatrix}$$

则
$$\begin{aligned}\boldsymbol{A}_2&=\boldsymbol{U}_1\boldsymbol{A}\boldsymbol{U}_1\\ &=\begin{bmatrix}-4 & 7.602631 & -0.447214\\ -4.472136 & 7.799999 & -0.400000\\ 0 & -0.399999 & 2.200000\end{bmatrix}\\ &=\boldsymbol{H}\end{aligned}$$