计算方法及MATLAB实现

郑勋烨　　编著

主审：　高世臣　　褚宝增　　王祖朝　　王翠香

副审：　李少琪　　李明霞　　赵琳琳　　赵俊芳

国防工业出版社

第七章　　矩阵特征值计算

7.2　　豪斯霍尔德反射与吉文斯旋转

7.2.1 豪斯霍尔德初等反射阵

【**定义** 1. **初等反射阵** (Householder Matrix)】

n 维单位列向量 $w=(w_1,w_2,\cdots,w_n)^T\in R^n$, 是满足其欧氏范数为 1 的向量: $\|w\|_2=1$, 或等价地范数平方为 1 :

$$\|w\|_2^2=w^Tw=w_1^2+w_2^2+\cdots+w_n^2=1$$

取 n 阶方阵 $H=I-2ww^T$ ，则它是一个 n 阶 正交对称矩阵， 特别称之为 初等反射阵(Householder Matrix).

即 初等反射阵(Householder Matrix) 形如

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

一般地，对于非零向量 $u \neq 0, H = I - 2\frac{uu^T}{||u||^2}$ 亦为初等反射阵.

【 定理 1. 初等反射阵作为正交对称阵 】初等反射阵 $H = I - 2ww^T$ 是正交对称阵. 即

(1) 初等反射阵是对称阵： $H^T = H$;

(2) 初等反射阵是正交阵： $H^T = H^{-1}$;

【证明】

(1)$H^T = (I - 2ww^T)^T = I - 2ww^T = H;$

(2) 注意到 n 维单位列向量 $w \in R^n$ 满足 $w^T w = 1$，我们有

$H^T H = (I - 2ww^T)^T (I - 2ww^T)$

$= I - 2ww^T - 2ww^T + 4w(w^T w)w^T$

$= I - 4ww^T + 4ww^T = I.$

【 注记. 初等反射阵的几何意义 】 构造单位列向量 w 为法向量且通过原点的超平面：

$$S := \{x \in R^n | w^T x = 0\}$$

对于任意向量 $v \in R^n$ ，我们有线性组合表达

$$v = x + y, \qquad x \in S, y \in S^{\perp}$$

从而我们有

$$Hx = (I - 2ww^T)x = x - 2ww^T x = x - 2w \cdot 0 = x$$

而对于 超平面的正交补空间 中的向量 $y \in S^{\perp}$，因为 $y^T x = 0$，即 $y//w$，不妨设 $y = kw$， k 为比例常数，则有

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

$$= y - 2ww^T \cdot kw = y - 2kw \cdot w^T w$$

$$= y - 2kw = y - 2y = -y.$$

于是对于任意向量 $v \in R^n$ ，我们有

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

而向量 $v' = x - y \in R^n$ 正是向量 $v = x + y \in R^n$ 关于 超平面 S 的 镜面反射 向量，彼此以这张 超平面 S 作为镜子，成为对方的影像.

Householder 初等反射变换的几何意义在此.

【**定理** 2. **初等反射阵作为向量变换阵**】存在初等反射阵 $H = I - 2ww^T$ 使两个等范数的不相等向量等价．即若 $\|x\| = \|y\|, x \neq y$, 则存在初等反射阵 $H = I - 2ww^T$ 使 $Hx = y$.

【**证明**】构造单位列向量 $w = \dfrac{x-y}{||x-y||}$, 则

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

为初等反射阵，从而

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

而注意到 $\|x\| = \|y\|$，即 $x^Tx = y^Ty$，以及作为一个数其转置为自身：　$y^Tx = (y^Tx)^T = x^Ty$，我们有

$||x-y||^2 = (x-y)^T(x-y) = x^Tx + y^Ty - y^Tx - x^Ty$

$= x^Tx + x^Tx - y^Tx - y^Tx = 2(x^Tx - y^Tx).$

故

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

【定理 3. 约化定理】

对非零向量 $x=(x_1,x_2,\cdots,x_n)^T\neq 0$, 存在初等反射阵 $H=I-2ww^T$ 使之与标准单位向量 $e_1=(1,0,\cdots,0)^T$ 等价．即若 $x\neq 0$, 则存在初等反射阵 $H=I-2ww^T$ 使 $Hx=-\sigma e_1$. 其中

$$\begin{cases} H &= I-2ww^T = I-\beta^{-1}uu^T \\ u &= x+\sigma e_1 \\ \sigma &= sgn(x_1)\|x\|_2 \\ \beta &= \dfrac{1}{2}\|u\|^2 = \sigma(x_1+\sigma) \end{cases}$$

【证明】

取数 $\sigma = sgn(x_1)\|x\|_2$，令 $y = -\sigma e_1 = (-\sigma, 0, \cdots, 0)^T$, 则显然 $\|x\| = \|y\| = |\sigma|, x \neq y$, 于是由定理 2 知，存在初等反射阵 $H = I - 2ww^T$ 使 $Hx = y = -\sigma e_1$. 其中单位列向量

$$w = \frac{x-y}{||x-y||} = \frac{x+\sigma e_1}{||x+\sigma e_1||}$$

令

$$u = x - y = x + \sigma e_1 = (x_1 + \sigma, x_2, \cdots, x_n)^T$$

则

$$H = I - 2ww^T = I - 2\frac{uu^T}{||u||^2}$$

为表示简洁可进而引入系数

$$\beta = \frac{1}{2}\|u\|^2 = \frac{1}{2}((x_1+\sigma)^2+x_2^2+\cdots+x_n^2) = \frac{1}{2}(\|x\|^2+\sigma^2+2x_1\sigma)$$
$$= \frac{1}{2}(2\sigma^2 + 2x_1\sigma) = \sigma^2 + x_1\sigma.$$

此时有

$$H = I - 2\frac{uu^T}{\|u\|^2} = I - \beta^{-1}uu^T, u = x + \sigma e_1$$

这里取 $\sigma = sgn(x_1)\|x\|_2$ 的原因在于使 σ 与 x_1 同号，以便计算 $x_1 + \sigma$ 时能够避免有效数字的损失．当然，有时若仅要求等价而不计误差，也可取 $\sigma = -sgn(x_1)\|x\|_2$. 参阅下面的例子。

【 释例. Householder 初等反射阵 】

对于如下对称矩阵:

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

其特征值 $\lambda_1 = 9$ 有对应的非零特征向量（一个单位向量）$x = (\frac{2}{3}, \frac{1}{3}, \frac{2}{3})^T \neq 0$. 利用 Householder 初等反射阵 $H = I - 2ww^T$ 使之与标准单位向量 $e_1 = (1, 0, 0)^T$ 等价，满足 $Hx = e_1$. 并计算相似矩阵 $B = H^T AH$.

解

由 约化定理，存在 3 阶初等反射阵

$$H = I - 2ww^T = I - 2\frac{uu^T}{\|u\|_2^2}$$

使之与 3 维标准单位向量 $e_1 = (1,0,0)^T$ 等价. 注意题设要求 $Hx = e_1$. 意味着取了 $\sigma = -1$, 则 $Hx = -\sigma e_1 = e_1$. 其中

$$\begin{cases} H & = I - \beta^{-1}uu^T \\ u & = x + \sigma e_1 = x - e_1 \\ \beta & = \dfrac{1}{2}\|u\|^2 \end{cases}$$

下面来逐个计算各个参数、向量和矩阵：

$$u = x - e_1 = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}, ||u||^2 = \frac{2}{3}.$$

$$\beta = \frac{1}{2}\|u\|^2 = \frac{1}{3}.$$

$$uu^T = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} (-\frac{1}{3}, \frac{1}{3}, \frac{2}{3}) = \begin{pmatrix} \frac{1}{9} & -\frac{1}{9} & -\frac{2}{9} \\ -\frac{1}{9} & \frac{1}{9} & \frac{2}{9} \\ -\frac{2}{9} & \frac{2}{9} & \frac{4}{9} \end{pmatrix}$$

$$H = I - 2\frac{uu^T}{||u||^2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 3 \begin{pmatrix} \frac{1}{9} & -\frac{1}{9} & -\frac{2}{9} \\ -\frac{1}{9} & \frac{1}{9} & \frac{2}{9} \\ -\frac{2}{9} & \frac{2}{9} & \frac{4}{9} \end{pmatrix}$$

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

从而令 3 阶初等反射正交矩阵取为

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

则

$$HAH=\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}\begin{pmatrix} 2 & 10 & 2 \\ 10 & 5 & -8 \\ 2 & -8 & 11 \end{pmatrix}\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & -9 \end{pmatrix}.$$

于是 Householder 初等反射阵将矩阵正交约化为对角矩阵.

7.2.2 豪斯霍尔德正交相似约化

【 **定义** 2. **海森伯格阵** (Hessenberg Matrix) 】

若矩阵 A 的 腋下对角线 以下的所有元素为 0： 当 $i > j+1$ 时 $b_{ij} = 0$; 则称为 上海森伯格阵 (Hessenberg Matrix). 通常简称为 海森伯格阵 (Hessenberg Matrix). 即

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots & & & b_{1n} \\ b_{21} & b_{22} & b_{23} & \cdots & & & b_{2n} \\ & b_{32} & b_{33} & \ddots & & & b_{3n} \\ & & \ddots & \ddots & \ddots & & \vdots \\ & & & & & & b_{n-1,n} \\ & & & & & b_{n,n-1} & b_{nn} \end{pmatrix}$$

特别地，若矩阵 A 的肩上对角线以上的所有元素也为 0：当 $j > i + 1$ 时 $b_{ij} = 0$; 即矩阵 A 只有主对角线及其腋下和肩上的斜对角线上的元素非零，而其他所有位置的元素均为 0. 则此时的 海森伯格矩阵 (Hessenberg Matrix) 称为 三对角矩阵. 即

$$B = \begin{pmatrix} c_1 & a_1 & & & & & \\ b_1 & c_2 & a_2 & & & & \\ & b_2 & c_3 & a_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots \\ & & & & b_{n-2} & c_{n-1} & a_{n-1} \\ & & & & & b_{n-1} & c_n \end{pmatrix}$$

当 海森伯格阵 A 的 腋下对角线 上的所有元素不为 0： 当 $i = j + 1$ 时 $b_{j+1,j} \neq 0$; 即

$$b_{21} b_{32} \cdots b_{n,n-1} \neq 0$$

则进而称 A 为 不可约 (irreducible) 海森伯格阵. 否则，如果存在某个 腋下对角线 上的元素为 0，则称 A 为 可约 (reducible) 海森伯格阵.

定理 4. 豪斯霍德正交相似 Hessenberg 化定理 (Householder Change)

对任意实矩阵 $A = (a_{ij})_{n\times n} \in R^{n\times n}$ ，存在系列 Householder 初等反射阵 $H = I - 2ww^T$ 的乘积形式的 正交矩阵

$$Q = H_1 H_2 \cdots H_{n-2}, \qquad Q^T = H_{n-2} \cdots H_2 H_1$$

和 海森伯格矩阵 (Hessenberg Matrix)B 使得实矩阵 A 与 海森伯格矩阵 (Hessenberg Matrix) B 正交相似 ： $Q^T AQ = B.$

特别地，对于实对称矩阵 $A=(a_{ij})_{n\times n}\in R^{n\times n}$ ，Hessenberg 标准形成为 对称三对角矩阵. 即

$$B = \begin{pmatrix} c_1 & b_1 & & & & & \\ b_1 & c_2 & b_2 & & & & \\ & b_2 & c_3 & b_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots \\ & & & & b_{n-2} & c_{n-1} & b_{n-1} \\ & & & & & b_{n-1} & c_n \end{pmatrix}$$

【 注记. Hessenberg 标准形的不唯一 】

任意实矩阵 $A=(a_{ij})_{n\times n}$ 正交相似 于 海森伯格矩阵 (Hessenberg Matrix)B ： $Q^TAQ=B$. 称为矩阵的 (上)Hessenberg 化. 海森伯格矩阵 B 称为实矩阵 A 的 (上)Hessenberg 标准形. 但这个 Hessenberg 标准形并非唯一的，相应的 正交矩阵Q 也并非唯一的．但所有的 Hessenberg 标准形是相似的．正所谓“菩萨千面，不离本尊”．我们有下面的定理.

【 **定理** 5. Hessenberg **标准形的相似** 】

对任意实矩阵 $A=(a_{ij})_{n\times n}\in R^{n\times n}$ ，存在系列 Householder 初等反射阵的乘积形式的 正交矩阵Q,U, 以及 海森伯格矩阵 (Hessenberg Matrix)B,G ，使得 $Q^TAQ=B$ ， $U^TAU=G$. 则存在主对角元为 ±1 的 对角矩阵$D=Diag[\pm1,\pm1,\cdots,\pm1]$, 使得 $D^{-1}BD=G$.

【 **证明略** 】可参阅 关治，陆金甫 ，《数值分析基础》，高等教育出版社． P396TH4.1 的证明.

【例 1. Householder 初等反射 Hessenberg 化】

利用 Householder 初等反射阵将如下矩阵正交约化为 Hessenberg 标准形．对于对称矩阵， Hessenberg 标准形成为对称三对角矩阵.

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 1 & 2 \\ 4 & 2 & 1 \end{pmatrix}$$

解

由于首个对角元之下的首列 $n-1=2$ 维向量

$$\alpha_1=(a_{21},a_{31},\cdots,a_{n1})^T=(3,4)^T\neq 0$$

由 约化定理， 存在 $n-1=2$ 阶初等反射阵

$$\tilde{H}_1=I-2w_1w_1^T=I-2\frac{u_1u_1^T}{\|u_1\|_2^2}$$

使之与 $n-1=2$ 维标准单位向量 $e_1=(1,0)^T$ 等价. 即 $\tilde{H}_1\alpha_1=-\sigma_1 e_1$. 其中

$$
\begin{cases}
\tilde{H}_1 & = I - \beta_1^{-1} u_1 u_1^T \\
u_1 & = \alpha_1 + \sigma_1 e_1 \\
\sigma_1 & = sgn(a_{21}) \|\alpha_1\|_2 = sgn(a_{21}) \sqrt{a_{21}^2 + a_{31}^2 + \cdots + a_{n1}^2} \\
\beta_1 & = \dfrac{1}{2} \|u_1\|^2 = \sigma_1 (a_{21} + \sigma_1)
\end{cases}
$$

下面来逐个计算各个参数、向量和矩阵:

$\sigma_1 = sgn(a_{21})\sqrt{a_{21}^2 + a_{31}^2 + \cdots + a_{n1}^2} = sgn(3)\sqrt{3^2 + 4^2} = +\sqrt{25} = 5.$

$u_1 = \alpha_1 + \sigma_1 e_1 = \begin{pmatrix} 3 \\ 4 \end{pmatrix} + 5\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}, ||u_1||^2 = 64 + 16 = 80.$

$\beta_1 = \frac{1}{2}\|u_1\|^2 = \frac{80}{2} = 40.$

$$u_1 u_1^T = \begin{pmatrix} 8 \\ 4 \end{pmatrix} (8, 4) = \begin{pmatrix} 64 & 32 \\ 32 & 16 \end{pmatrix}$$

$$\tilde{H}_1 = I - 2\frac{u_1 u_1^T}{||u||^2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{40} \begin{pmatrix} 64 & 32 \\ 32 & 16 \end{pmatrix}$$

$$= \frac{1}{40} \begin{pmatrix} -24 & -32 \\ -32 & 24 \end{pmatrix}.$$

从而令 3 阶分块对角初等反射正交矩阵取为

$$H_1 = \begin{pmatrix} 1 & \\ & \tilde{H}_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

则

$$H_1AH_1 =$$

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

$$= \begin{pmatrix} 1 & -5 & 0 \\ -5 & \frac{73}{25} & \frac{14}{25} \\ 0 & \frac{14}{25} & -\frac{23}{25} \end{pmatrix}.$$

【注记 1.Householder(豪斯霍尔德) 与矩阵分解】

Alston Householder 早年在美国西北大学和康奈尔大学 (Cornell) 学习数学，研究泛函分析． 1946 年进入美国橡树山 (Oak Ridge) 国家实验室，由于计算机的初兴，迫切要求有效求解线性方程组和矩阵特征值问题，他的兴趣转向数值分析．他发现从矩阵分解的角度分类，许多算法事实上是彼此等价的． 1953 年，他在纽约出版《数值分析原理》(Principles of Numerical Analysis,McGraw-Hill,New York.), 开启了系统使用“范数”作为数值分析工具的先河．