·学术报告会论文选

斜映射矩阵和Householder变换的推广

谭 领

在数值线代数中，广泛使用 Householder 变换[1]，其变换阵为

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

其中 u 为 n 维向量，且 $\|u\|=1$，I 为单位阵。通常称这种变换为"平面映射"[2]，说严格一些，应是"关于过坐标原点的平面的正映射"。映射的基准平面是过坐标原点的 $u^T x = 0$，其中 u 为平面法向量。映射方向与平面垂直，也就是 u，故这是一种映射方向与平面法向一致的映射。

如果映射方向 f 与平面法向 u 不一定一致，可称为平面斜映射，本文对关于一般(超)平面(不一定过坐标原点)的斜映射作一矩阵分析，同时也将 Householder 变换加以推广。

客观物体的外形及内部结构常有某种"对称性"，人们往往利用这种对称性成倍地减少计算量。仔细看来，人们通常理解的对称指的均是正对称，其实还有更普遍的关于某方向的斜对称，如图1的飞机模型，P_1 与 Q_1 以及 P_2 与 Q_2 等，均是关于 l_1 的正对称点，整个图形是关于 l_1 的正对称形。但是，P_1 与 P'_1 以及 P_2 与 P'_2 都是沿方向 r_1 的关于 l_2 的斜对称点，P_3 与 P'_3 是沿方向 r_2 关于 l_3 的斜对称点。在本文中，把斜对称叫做斜映射。本文认为，如果在实际计算中充分认识斜映射，必然还可以进一步简化计算。

利用斜映射，也可提出新的数值计算法，如使用 QR[3] 法求阵的特征值时，总是先用相似变换把阵化为 Hessenberg 型，这就可用斜映射的办法，我们将在[4]专讲这个问题。

在数值方法中，不少方法有直观的几何解释，如梯度法，切线法，弦截法等。利用斜映射，可对通用的 Gauss-Jordon 消去法、Gauss 消去法等作出几何解释。特别，求阵的特征值中比较重要的 LU[3]，QR 两个方法，原来似毫不相干，QR 可用正映射作解释。本文指出，可用斜映射去解释 LU，因而将这二方法统一起来。

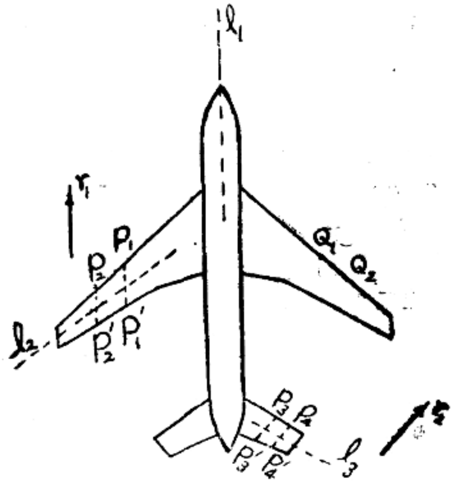

图 1

定义 1　对 $n\times n$ 的阵 S 及 $n\times 1$ 的向量 $\boldsymbol{b}$，如果对任一 $n\times 1$ 的 $\boldsymbol{x}$，当

$$\boldsymbol{y}=S\boldsymbol{x}+\boldsymbol{b} \tag{1}$$

时，恒有

$$\boldsymbol{x}=S\boldsymbol{y}+\boldsymbol{b} \tag{2}$$

则称 S 为互换阵，$\boldsymbol{b}$ 是 S 的相配向量。并称(1)是互换变换。

定理 1　S 是互换阵，$\boldsymbol{b}$ 是 S 的相配向量，其充分必要条件是

$$S^{-1}=S \qquad 和 \qquad S\boldsymbol{b}=-\boldsymbol{b} \tag{3}$$

证　充分性是显然的，下面证必要性。

由于 $\boldsymbol{x}$ 的任意性，就取 $\boldsymbol{x}=\boldsymbol{o}$，则由(1)得 $\boldsymbol{y}=\boldsymbol{b}$，由(2)知 $S\boldsymbol{b}=-\boldsymbol{b}$。

又对任意 $\boldsymbol{x}$，总有 $\boldsymbol{x}=S^2\boldsymbol{x}$，由于 $\boldsymbol{x}$ 的任意性，就知 $S^2=I$，故 $S^{-1}=S$。证完。

单位阵显然是互换阵，$\boldsymbol{b}=\boldsymbol{o}$ 是其相配向量。容易验证，变换

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}=\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}+g\begin{pmatrix} -\sin 2\theta \\ 2\cos^2\theta \end{pmatrix} \tag{4}$$

是互换变换。（g——常数）

对称的互换阵一定是正交阵。一个互换阵 S 如果存在非零的相配向量 $\boldsymbol{b}$ 时，-1 就一定是 S 的特征值，$\boldsymbol{b}$ 即是属于 -1 的特征向量。

定义 2　设 P：$\boldsymbol{a}^T\boldsymbol{x}+d=0$ 是 n 维空间 R_n 的一个平面，$\boldsymbol{r}\neq\boldsymbol{0}$ 是一 n 维向量，且 $\boldsymbol{r}^T\boldsymbol{a}\neq 0$，若 $\boldsymbol{x},\boldsymbol{y}\in R_n$ 且满足

$$\left.\begin{array}{ll} 1^\circ & \boldsymbol{a}^T\left(\dfrac{\boldsymbol{x}+\boldsymbol{y}}{2}\right)+d=0 \\ 2^\circ & \boldsymbol{y}-\boldsymbol{x}=k\boldsymbol{r}\ (k\text{ 为与 }\boldsymbol{x},\boldsymbol{y}\text{ 有关的常数}) \end{array}\right\} \tag{5}$$

则称 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 关于 P 沿方向 $\boldsymbol{r}$ 的(斜)映像。当然，$\boldsymbol{x}$ 也是 $\boldsymbol{y}$ 关于 P 沿方向 $\boldsymbol{r}$ 的映像。特别当 $\boldsymbol{r}=k_1\boldsymbol{a}$（$k_1$——常数）时，称 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 关于 P 的正映像。

若互换变换(1)总使 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 关于平面 P 沿方向 $\boldsymbol{r}$ 的映像，则称这种互换变换是关于平面 P 沿方向 $\boldsymbol{r}$ 的(斜)映射。如果 $\boldsymbol{x},\boldsymbol{y}$ 总是正映像，则称(1)为正映射。

定理 2　设 $\boldsymbol{a}^T=(a_1,a_2,\cdots,a_n)\neq\boldsymbol{0}$，$\boldsymbol{f}^T=(f_1,f_2,\cdots,f_n)\neq\boldsymbol{0}$。满足

$$\boldsymbol{a}^T\boldsymbol{f}=1 \tag{6}$$

作阵及向量

$$S=I-2\boldsymbol{f}\boldsymbol{a}^T \tag{7}$$

$$\boldsymbol{b}=-2d\boldsymbol{f}\ (d\text{——常数}) \tag{8}$$

则 S 是互换阵，$\boldsymbol{b}$ 是 S 的相配向量

证 $S^2=(I-2\boldsymbol{f}\boldsymbol{a}^T)(I-2\boldsymbol{f}\boldsymbol{a}^T)$

$=I-4\boldsymbol{f}\boldsymbol{a}^T+4\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{f}\boldsymbol{a}^T=I$，故 $S^{-1}=S$

$S\boldsymbol{b}=(I-2\boldsymbol{f}\boldsymbol{a}^T)\boldsymbol{b}=\boldsymbol{b}-2\boldsymbol{f}\boldsymbol{a}^T(-2d\boldsymbol{f})$

$=\boldsymbol{b}+4d\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{f}=\boldsymbol{b}-2\boldsymbol{b}=-\boldsymbol{b}$ 证完

定理 3 按定理 2 构造互换变换

$$\boldsymbol{y}=S\boldsymbol{x}+\boldsymbol{b} \qquad (9)$$

当且仅当 $\boldsymbol{x}$ 满足 $\boldsymbol{a}^T\boldsymbol{x}+d=0$ 时，有 $\boldsymbol{x}=S\boldsymbol{x}+\boldsymbol{b}$，即互换变换（9）下的唯一不变面是 $\boldsymbol{a}^T\boldsymbol{x}+d=0$

证 设 $\boldsymbol{x}$ 在该平面上，于是

$$\boldsymbol{y}=S\boldsymbol{x}+\boldsymbol{b}=(I-2\boldsymbol{f}\boldsymbol{a}^T)\boldsymbol{x}+\boldsymbol{b}$$

$$=\boldsymbol{x}-2\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{x}-2d\boldsymbol{f}=\boldsymbol{x}$$

另一方面，若 $\boldsymbol{x}^{(1)}$ 满足 $\boldsymbol{x}^{(1)}=S\boldsymbol{x}^{(1)}+\boldsymbol{b}$，则有 $(I-S)\boldsymbol{x}^{(1)}=\boldsymbol{b}$，也即 $2\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{x}^{(1)}=\boldsymbol{b}$，因此 $2\boldsymbol{a}^T\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{x}^{(1)}=\boldsymbol{a}^T\boldsymbol{b}$，也即 $2\boldsymbol{a}^T\boldsymbol{x}^{(1)}=\boldsymbol{a}^T\boldsymbol{b}=\boldsymbol{a}^T(-2d\boldsymbol{f})=-2d$ 于是

$$\boldsymbol{a}^T\boldsymbol{x}^{(1)}+d=0 \qquad \text{证完。}$$

定理 4 按定理 2 构造互换变换（9），则（9）是关于平面 $\boldsymbol{a}^T\boldsymbol{x}+d=0$ 沿方向 $\boldsymbol{f}$ 的映射。

证 $\frac{1}{2}\boldsymbol{a}^T(\boldsymbol{y}+\boldsymbol{x})+d=\frac{1}{2}\boldsymbol{a}^T[(I-2\boldsymbol{f}\boldsymbol{a}^T)\boldsymbol{x}+\boldsymbol{b}+\boldsymbol{x}]+d$

$$=\frac{1}{2}\boldsymbol{a}^T[2\boldsymbol{x}-2\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{x}+\boldsymbol{b}]+d=0$$

$$\boldsymbol{y}-\boldsymbol{x}=-2\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{x}-2d\boldsymbol{f}=k\boldsymbol{f}$$

其中 $k=-2(\boldsymbol{a}^T\boldsymbol{x}-d)$ 是常数 证完

我们就称 $I-2\boldsymbol{f}\boldsymbol{a}^T$ 为斜映射矩阵。

正映射是斜映射之特殊情形，只需将定理 2 中映射方向 $\boldsymbol{f}$ 取为与平面 P 的法向量 $\boldsymbol{a}$ 平行且满足（6）就行了，即只需取

$$\boldsymbol{f}=\frac{2}{g^2}\boldsymbol{a} \qquad \text{其中} \quad g^2=\sum_{j=1}^{n} a_j^2$$

就有正映射的下述结果

推论： 对 R_n 的平面 P：$\boldsymbol{a}^T\boldsymbol{x}+d=0$，作

$$S = I - \frac{2}{g^2}\boldsymbol{a}\,\boldsymbol{a}^T$$

$$\boldsymbol{b} = -\frac{2d}{g^2}\boldsymbol{a}$$

则 $\boldsymbol{y} = S\boldsymbol{x} + \boldsymbol{b}$ 是关于 P 的正映射。

当 $d = 0$ 时，$\boldsymbol{a}^T\boldsymbol{x} = 0$ 为过原点平面，相应的 S 阵就是人们熟知的 Householder 阵。

下面进一步分析斜映射阵 $I - 2\boldsymbol{f}\boldsymbol{a}^T$ 的特征值及特征向量。

定理 5 若 $a_1 \neq 0$，则 $I - 2\boldsymbol{f}\boldsymbol{a}^T$ 有完全的特征向量系，其特征值及特征向量是：

$$\lambda_1 = -1, \qquad \boldsymbol{x}_1 = \boldsymbol{f} = (f_1, f_2, \cdots, f_n)^T$$

$$\lambda_i = 1, \qquad \boldsymbol{x}_i = a_i\boldsymbol{e}_1 - a_1\boldsymbol{e}_i \qquad (i = 2, 3, \cdots, n)$$

其中 $\boldsymbol{e}_i^T = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$ 即只第 i 分量为 1

证 $(I - 2\boldsymbol{f}\boldsymbol{a}^T)\boldsymbol{f} = -\boldsymbol{f}$，故特征值 $\lambda_1 = -1$ 的特征向量为 $\boldsymbol{f}$，又 $(I - 2\boldsymbol{f}\boldsymbol{a}^T)\boldsymbol{x}_i = \boldsymbol{x}_i - 2\boldsymbol{f}\boldsymbol{a}^T\boldsymbol{x}_i = \boldsymbol{x}_i$ 故 $\lambda_i = 1$ 的特征向量为 $\boldsymbol{x}_i = a_i\boldsymbol{e}_1 - a_1\boldsymbol{e}_i$。诸 $\boldsymbol{x}_i$ 是线性无关，是显然的，它们与 $\boldsymbol{f}$ 也线性无关，因为若 $c\boldsymbol{f} + d\boldsymbol{x}_i = 0$ 则 $c\boldsymbol{a}^T\boldsymbol{f} + \boldsymbol{a}^T d\boldsymbol{x}_i = 0$ 故 $c\boldsymbol{a}^T\boldsymbol{f} = 0$ 即 $c = 0$，于是 $d = 0$。

证完。

下面用斜映射对数值线性计算中若干通用方法作几何解释。为此，从另一角度提出问题，即对于给定的 $\boldsymbol{k}^T = (k_1, k_2, \cdots, k_n)$，能否找到平面及方向，使之 $\boldsymbol{k}$ 之映像位于坐标轴上呢？如果允许斜映射，定理 6 指出，取坐标面就可办到；如果限于正映射，虽然取坐标面已难于解决，但定理 7 指出，取过原点的平面就可办到。前者揭示了解线方程组的 Gauss-Jordon 消去法、Gauss 消去法以及阵的 LU 分解的几何意义，后者正好针对 Householder 法及 QR 分解等。

定理 6 对于 $\boldsymbol{k} = (k_1, k_2, \cdots, k_n)^T$，$(k_i \neq 0)$

取 $\boldsymbol{a} = \frac{1}{2k_i}\boldsymbol{e}_i$，$\boldsymbol{f} = (k_1, k_2, \cdots, 2k_i, k_{i+1}, \cdots, k_n)^T$，则 $\boldsymbol{k}$ 关于坐标面 $x_i = 0$ 沿方向 $\boldsymbol{f}$ 的映像是位于坐标轴上的 $\boldsymbol{q} = -k_i\boldsymbol{e}_i$ （证略）

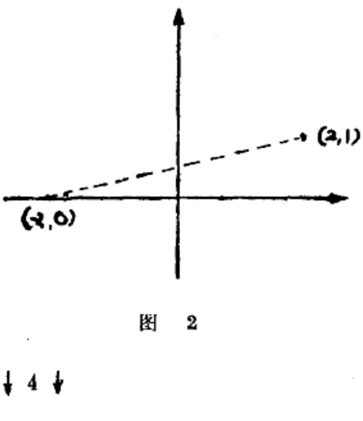

图 2

这个定理在二维平面上的几何直观如图 2。对于点 (2, 1)，可以取坐标轴 Y 轴，使之 (2, 1) 关于 Y 轴的映像是位于 X 轴上的 (-2, 0)，映射方向是 $(2k_1, k_2) = (4, 1)$。

下面对解线方程组的一些方法作几何解释，只着重谈 Gauss-Jordou 消去法。设给定非奇方程组 $B\boldsymbol{x} = \boldsymbol{g}$ 其中

$$B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n), \quad \boldsymbol{b}_j = (b_{1j}, b_{2j}, \cdots b_{nj})^T \quad (1 \leqslant j \leqslant n)$$

$$\boldsymbol{g} = (g_1, g_2, \cdots, g_n)^T,$$

不失一般性，可设 $b_{11}\neq 0$，由定理 6，找坐标面 P_1 及相应方向，使之 $\boldsymbol{b}_1$ 的斜映象为 $\boldsymbol{b}_1^{(1)}=(-b_{11}, 0, \cdots, 0)^T$ 同时，$\boldsymbol{b}_2, \cdots, \boldsymbol{b}_n, \boldsymbol{g}$ 的映像为 $\boldsymbol{b}_2^{(1)}, \cdots, \boldsymbol{b}_n^{(1)}, g^{(1)}$，这即是用阵

$$A_1=I-2\begin{pmatrix}2b_{11}\\ b_{21}\\ \vdots\\ b_{n1}\end{pmatrix}\left(\frac{1}{2b_{11}} 0, \cdots, 0\right)=\begin{pmatrix}-1 & & & & 0\\ -\frac{b_{21}}{b_{11}} & 1 & & & \\ -\frac{b_{31}}{b_{11}} & & 1 & & \\ \vdots & & 0 & \ddots & \\ -\frac{b_{n1}}{b_{11}} & & & & 1\end{pmatrix}$$

去乘诸 $\boldsymbol{b}_i$ 而得 $A_1\boldsymbol{b}_i=\boldsymbol{b}_i^{(1)}$。$A_1\boldsymbol{g}=\boldsymbol{g}^{(1)}$ 再设 $b_{22}^{(1)}\neq 0$，找坐标面 P_2 及相应方向，使 $\boldsymbol{b}_2^{(1)}$ 之斜映像为 $\boldsymbol{b}_2^{(2)}=-b_{22}^{(1)}\boldsymbol{e}_2$ 同时 $\boldsymbol{b}_3^{(1)}, \cdots, \boldsymbol{b}_n^{(1)}, \boldsymbol{g}^{(1)}$ 之斜映像为 $\boldsymbol{b}_3^{(2)}, \cdots, \boldsymbol{b}_n^{(2)}, \boldsymbol{g}^{(2)}$，这即用阵

$$A_2=\begin{pmatrix}1 & -\frac{b_{21}^{(1)}}{b_{22}^{(1)}} & & & 0\\ & -1 & & & \\ & -\frac{b_{32}^{(1)}}{b_{22}^{(1)}} & 1 & & \\ & \vdots & & \ddots & \\ 0 & -\frac{b_{n2}^{(1)}}{b_{22}^{(1)}} & & & 1\end{pmatrix}$$

去乘诸 $b_i^{(1)}$ 而得 $A_2\boldsymbol{b}_i^{(1)}=\boldsymbol{b}_i^{(2)}$，$A_2\boldsymbol{g}^{(1)}=\boldsymbol{g}^{(2)}$，…继续做下去，这看出，Gauss-Jardon 消去过程实质就是一个不断斜映射过程。类似地可解释 Gauss 消去法以及阵的 LU 分解等。对于正映射有

定理 7 对于 $\boldsymbol{k}=(k_1, k_2, \cdots, k_n)^T$，$(k_i\neq 0)$，可取过原点的平面 $\boldsymbol{a}^T\boldsymbol{x}=0$，其中

$$\boldsymbol{a}=(k_1, k_2, \cdots, k_i-q, k_{i+1}, \cdots, k_n)^T=\boldsymbol{k}-\boldsymbol{q}$$

$$\boldsymbol{q}=q\boldsymbol{e}_i, \qquad q^2=\sum_{j=1}^{n} k_j^2$$

使之 $\boldsymbol{k}$ 关于 $\boldsymbol{a}^T\boldsymbol{x}=0$ 的正映像为 $\boldsymbol{q}=q\boldsymbol{e}_i$。

值得指出，定理中的 $q=\pm r=\pm\sqrt{\sum_{j=1}^{n} k_j^2}$，一种符号相应一个平面，这二平面有如下性质

系：在定理 7 的前提下，取平面

$$P_1: \ \boldsymbol{a}_1^T\boldsymbol{x}=0; \qquad P_2: \ \boldsymbol{a}_2^T\boldsymbol{x}=0$$

其中

$$\boldsymbol{a}_1=(k_1, \cdots, k_{i-1}, k_i-r, k_{i+1}, \cdots, k_n)^T, \ \boldsymbol{a}_2=(k_1, \cdots, k_i+r, k_{i+1}, \cdots k_n)^T$$

则有 1° P_1 与 P_2 正交

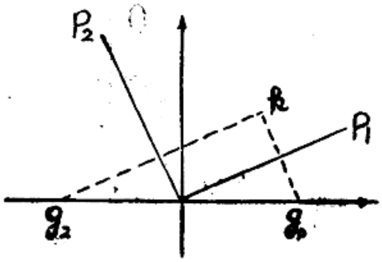

图 3

2° k 关于 P_1, P_2 的正映像均在同一坐标轴上，且关于原点对称（证略）。

这表明，总存在两个相互正交的平面，使之 k 之正映像位于同一坐标轴上。这个事实在二维平面的几何直观如图 3。

人们已经用正映射解释了 QR，现在，又可用斜映射解释 LU，正映射是特殊的斜映射，这就将 LU 与 QR 均从几何上统一起来了。值得指出，事情发展到斜映射，基准平面及反映射方向就有广泛的选择自由了，Gauss-Jordon 消去法是一种选法，在[4]中，又是另一种选法，这是化阵为 Hessenberg 型的一种新的相似变换。

参考文献

[1] Householder, A. S. Bauer. F. L. On certain methods for expanding the Characteristic polynomial. Numer. Math. 1959

[2] Stewart. G. W. Introduction to matrix computation 1973

[3] Wilkinson. J. H. The algebraic eigenvalue problem. Clarendon press. 1965.

[4] 谭 领：化阵为 Hessenberg 型的一种新的相似变换（尚未发表）