第 16 卷 第 2 期
1990 年 3 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 16, No. 2
Mar., 1990

矩阵上三角化的递推 Householder 变换公式及其应用

刘整社　文传源　张明廉
（北京航空航天大学）

摘　　要

本文给出利用 Householder 变换对矩阵进行上三角化的按列递推公式. 该公式用于差分、多项式、AR、ARMA、CARMA 等模型的结构和参数辨识时，可大大减小计算量，提高算法的实用性.

关键词——系统辨识，建模，阶次估计，Householder 变换.

一、前　　言

结构辨识是系统辨识理论的一个重要方面. 目前常用的方法有 F-检验[1]，FPE-检验[2]，AIC-准则[2]，CAT-准则[2]等，这些方法都要用到模型参数最小二乘估计的残差平方和. 按照以往的算法，只有等模型参数估计出来之后，才能获得残差平方和，进行结构检验. 但是，模型参数的估计必须以结构已知为前提，这就给实际辨识工作带来了困难.

文献[3—8]分别对差分模型、多项式模型、高次多项式模型、AR 模型、ARMA 模型、CARMA 模型提出了结构和参数同时估计的辨识方法，为模型结构辨识开辟了一条新的途径. 这些方法的核心是利用观测数据构造一个信息矩阵，使该矩阵经 Householder 变换，所形成的上三角矩阵的主对角元素的平方，正好对应于不同结构下，模型参数最小二乘估计的残差平方和，从而在模型参数估计出来之前就可先确定模型结构，同时，由所得到的上三角矩阵回代出模型参数.

构造信息矩阵时，需要给定模型结构参数的范围. 例如，差分模型阶次可能的最大值. 为了避免遗漏，这个范围一般给的较大. 这就加大了信息矩阵的列数. 因为利用现有的 Householder 变换公式使信息矩阵上三角化时，该矩阵的所有列都参与变换，乘法次数正比于信息阵列数的平方，因而运算量较大.

本文给出的 Householder 变换公式，是逐列递推进行的，即对第 i 列变换时，列数大于 i 的各列都不参与运算. 这样，就可以在信息矩阵上三角化的过程中，同时进行结构检验，一得到合适的模型结构就可以停止变换，从而使运算量与模型结构范围的给定值无

本文于 1987 年 11 月 9 日收到.

关.这种算法中，整个辨识过程的运算量仅相当于已知模型结构只做参数估计的运算量.与现有算法相比,运算量大为减小.

二、递 推 公 式

由文献[9]知,矩阵 $\boldsymbol{D}=(d_{ij})_{m\times n}(m\geqslant n)$ 的上三角化可由 Householder 变换公式

$$\begin{cases}\alpha_l=\left[\sum\limits_{i=l}^{m}(d_{il}^{(l)})^2\right]^{\frac{1}{2}},\\ \boldsymbol{u}_l=(0,\cdots,0,d_{ll}^{(l)}+\operatorname{sign}(d_{ll}^{(l)})\alpha_l,d_{l+1l}^{(l)},d_{l+2l}^{(l)},\cdots,d_{ml}^{(l)})^T,\\ \sigma_l=\alpha_l(\alpha_l+|d_{ll}^{(l)}|),\\ \boldsymbol{q}_l^T=\boldsymbol{u}_l^T\boldsymbol{D}^{(l)}/\sigma_l,\\ \boldsymbol{D}^{(l+1)}=\boldsymbol{D}^{(l)}-\boldsymbol{u}_l\boldsymbol{q}_l^T,\\ l=1,2,\cdots,n\end{cases}\tag{2.1}$$

完成. 式中，$\boldsymbol{D}^{(l)}=(d_{ij}^{(l)})_{m\times n}$ 表示 $\boldsymbol{D}$ 经 $(l-1)$ 次变换所得到的矩阵，$\boldsymbol{D}^{(1)}=\boldsymbol{D}$.

若用 $q_l(k)$ 表示 $\boldsymbol{q}_l$ 的第 k 个元素,用 $\boldsymbol{d}_k^{(l)}$ 表示 $\boldsymbol{D}^{(l)}$ 的第 k 列，则当 $k>l$ 时,有

$$\begin{cases}q_l(k)=\boldsymbol{u}_l^T\boldsymbol{d}_k^{(l)}/\sigma_l=\left\{[d_{ll}^{(l)}+\operatorname{sign}(d_{ll}^{(l)})\alpha_l]d_{lk}^{(l)}+\sum\limits_{i=l+1}^{m}d_{il}^{(l)}d_{ik}^{(l)}\right\}\Big/\sigma_l,\\ \boldsymbol{d}_k^{(l+1)}=\boldsymbol{d}_k^{(l)}-q_l(k)\boldsymbol{u}_l.\end{cases}\tag{2.2}$$

(2.2)式表示第 l 次 $(l=1,2,\cdots,k-1)$ 变换时，$\boldsymbol{D}$ 的第 k 列需做的变动. 所以，$\boldsymbol{d}_k^{(k)}$ 可以由公式

$$\begin{cases}\alpha_1=\left[\sum\limits_{i=l}^{m}(d_{il}^{(l)})^2\right]^{\frac{1}{2}},\\ \boldsymbol{u}_l=(0,\cdots,0,d_{ll}^{(l)}+\operatorname{sign}(d_{ll}^{(l)})\alpha_l,d_{l+1l}^{(l)},d_{l+2l}^{(l)},\cdots,d_{ml}^{(l)})^T,\\ \sigma_l=\alpha_l(\alpha_l+|d_{ll}^{(l)}|),\\ q_l(k)=\left\{[d_{ll}^{(l)}+\operatorname{sign}(d_{ll}^{(l)})\alpha_l]d_{lk}^{(l)}+\sum\limits_{i=l+1}^{m}d_{il}^{(l)}d_{ik}^{(l)}\right\}\Big/\sigma_l,\\ \boldsymbol{d}_k^{(l+1)}=\boldsymbol{d}_k^{(l)}-q_l(k)\boldsymbol{u}_l,\\ l=1,2,\cdots,k-1,\end{cases}\tag{2.3}$$

从 $\boldsymbol{d}_k^{(1)}=\boldsymbol{d}_k$ 得到.

可见,只要已知

$$d_{il}^{(l)},i=l,l+1,\cdots,m;\ l=1,2,\cdots,k-1,$$

就可以由 $\boldsymbol{d}_k^{(1)}$ 递推出 $\boldsymbol{d}_k^{(k)}$. 但由于 $d_{il}^{(l)}$, $i=l$, $l+1,\cdots,m$. 在第 l 次变换中已被改变,所以,要从 $\boldsymbol{d}_k^{(1)}$ 递推出 $\boldsymbol{d}_k^{(k)}$,必须保存 $d_{il}^{(l)},i=l,l+1,\cdots,m;l=1,2,\cdots,k-1$. 因为 $d_{il}^{(l+1)}=0,i=l+1,l+2,\cdots,m$,所以,可以在第 l 次变换时,只对 $d_{ll}^{(l)}$ 作变换,而将$d_{il}^{(l)},i=l+1,l+2,\cdots,m$ 保留，在以后用到 $d_{il}^{(i)}$, $i=l+1,l+2,\cdots,m$; $i>l$ 时，用零代替即可. 因此，只要保存 $d_{ll}^{(l)}$, $l=1,2,\cdots,k-1$，就可以由 $\boldsymbol{d}_k^{(1)}$ 递推出 $\boldsymbol{d}_k^{(k)}$.

记

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^T,$$
$$\boldsymbol{p} = (d_{11}^{(1)}, d_{22}^{(2)}, \cdots, d_{nn}^{(n)})^T,$$

则有

$$\begin{cases} p_1 = d_{11}^{(1)} \\ \alpha_1 = \left[\sum_{i=1}^{m} (d_{i1}^{(1)})^2\right]^{\frac{1}{2}} \\ d_{11}^{(2)} = -\mathrm{sign}(p_1)\alpha_1 \\ \left\{\begin{array}{l} \boldsymbol{u}_l = (0, \cdots, 0, p_l + \mathrm{sign}(p_l)\alpha_l, d_{l+1l}^{(l)}, d_{l+2l}^{(l)}, \cdots, d_{ml}^{(l)})^T \\ \sigma_l = \alpha_l(\alpha_l + |p_l|) \\ q_l(k) = \left\{[p_l + \mathrm{sign}(p_l)\alpha_l]\, d_{lk}^{(l)} + \sum_{i=l+1}^{m} d_{il}^{(l)} d_{ik}^{(l)}\right\} \Big/ \sigma_l \\ \boldsymbol{d}_k^{(l+1)} = \boldsymbol{d}_k^{(l)} - q_l(k)\boldsymbol{u}_l \\ l = 1, 2, \cdots, k-1 \end{array}\right. \\ p_k = d_{kk}^{(k)} \\ \alpha_k = \left[\sum_{i=k}^{m} (d_{ik}^{(k)})^2\right]^{\frac{1}{2}} \\ d_{kk}^{(k+1)} = -\mathrm{sign}(p_k)\alpha_k \\ k = 2, 3, \cdots. \end{cases} \tag{2.4}$$

由(2.4)式可以看出，第 k 次变换只改变 $\boldsymbol{D}$ 的第 k 列，其它列均不参与运算，从而实现了对矩阵上三角化的按列递推.

三、应 用 举 例

文献[3—8]中所采用的都是(2.1)式的 Householder 变换公式，因而运算量还是比较大. 如果用本文所给的公式(2.4)代替(2.1)，就可以得到更有效的辨识算法. 我们用(2.4)式对[3—8]中的例子做过验算，结果完全吻合，而且运算量大大减少. 限于篇幅，下面仅给出一个多项式回归的例子.

表 1 磁性材料的 B-H 特性

B(高斯)	0.00	0.10	0.20	0.40	0.60	0.70	0.80	1.00	1.05	1.10
H(安/米)	0.00	0.35	0.60	0.80	0.95	1.00	1.20	1.40	1.50	1.60
B(高斯)	1.15	1.25	1.35	1.44	1.50	1.52	1.60	1.70	1.80	1.90
H(安/米)	1.70	2.00	2.60	3.50	4.50	6.00	13.50	35.00	74.00	135.00

例. 某磁性材料的 B-H 特性测量数据如表 1 所示. 在 95% 的置信带内，由(2.4)式得到的 B-H 特性为 5 次多项式

$$H = a_0 + a_1 B + a_2 B^2 + a_3 B^3 + a_4 B^4 + a_5 B^5,$$

其中 $a_0 = -1.84824, a_1 = 71.14070, a_2 = -361.30420, a_3 = 656.42407, a_4 = -490.26880, a_5 = 128.94095$. 回归值与实测值的比较如图 1 所示. 可见,两者吻合得很好.

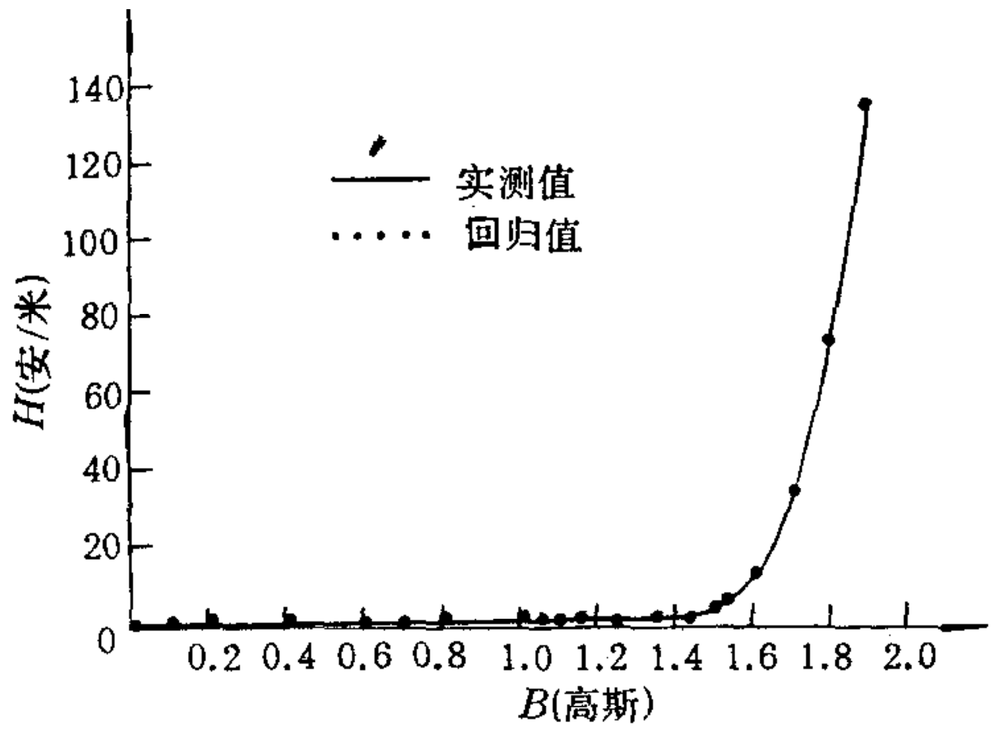

图 1　某磁性材料的磁化曲线

参　考　文　献

[1] 南京大学数学系计算数学专业,概率统计基础和概率统计方法,科学出版社,北京,1979 年,228—233 页.

[2] Haykin, S., Nonlinear Methods of spectral Analysis, Springer-Verlag Berlin Heidelberg, New York, 1979, 41—48.

[3] 黄俊钦,张继志,一种同时辨识模型阶次和参数的方法,仪器仪表学报,第 5 卷,第 4 期,1984 年,339—345 页.

[4] 黄俊钦,刘整社,多项式回归的快速算法,应用数学学报,第 9 卷,第 2 期,1986 年,146—153 页.

[5] 黄俊钦,刘整社,多项式模型方次与参数同时估计的正交化法,航空学报,第 8 卷,第 1 期,58—66 页.

[6] 黄俊钦,余辉里,只估计 AR 参数的 ARMA(p,q) 谱估计的新算法,自动化学报,第 13 卷,第 1 期,1987 年,1—7 页.

[7] 黄俊钦,刘整社,一种同时估计 ARMA 模型阶次与参数的线性算法,仪器仪表学报,第 8 卷,第 1 期,1987 年,90—93 页.

[8] 黄俊钦,刘整社,CARMA 模型结构与参数同时估计的一种线性算法,第五届全国控制理论及其应用学术交流会论文集,1985 年,236—240 页.

[9] 冯康等,数值计算方法,国防工业出版社,北京,1978 年,255—259 页.

A RECURSIVE HOUSEHOLDER TRANSFORMATION FORMULA AND ITS APPLICATIONS

LIU ZHENGSHE　WEN CHUANYUAN　CHANG MNGLIAN

(Beijing University of Acronautics and Astronautics)

ABSTRACT

A recursive Householder formula for transforming high rectangular matrices into upper triangular matrices column by column is proposed. The formula can be used to simultaneously estimate the orders and parameters of polynomial models, difference models, AR models, ARMA models, CARMA models and so on. Since the orders can be estimated in passing by our formula, the computation time for estimating both orders and parameters is only slightly more than that for just estimating the parameters of the model with known orders.

Key words ——System identification; modeling; order estimation; Householder transformation.