第24卷 第1期
1998年1月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 24, No. 1
Janu., 1997

短 文

两种新的有效的非线性系统最小二乘辨识算法

王 晓 韩崇昭 万百五
（西安交通大学系统工程研究所 西安 710049）

摘 要 提出了两种新的有效的最小二乘算法——改进的双对角化最小二乘算法 MBLS-Ⅰ与 MBLS-Ⅱ. 在存在舍入误差的条件下，证明了算法的收敛性. 该算法具有几乎不受舍入误差影响的优点，优于一般常用的最小二乘算法，包括数值性态极佳的 SVD 算法. 同时，基于该算法及 SVD 算法，构造出了一种新的 NARMAX 模型结构与参数辨识的一体化算法. 仿真结果证明了此新算法的优越性.

关键词 非线性系统，系统辨识，双对角化最小二乘.

1 引言

实际中遇到的大多数系统都是非线性的，因而研究非线性系统的建模及辨识具有重要的现实意义. 就此英国的 Billings 等人基于他们自1985年提出的 NARMAX 模型做了大量开创性的工作[1—7]. NARMAX 模型本质上是一类带外生变量的非线性自回归滑动平均模型，且从理论上可逼近任意的离散非线性系统. Billings 等人提出了一些基于直交化的最小二乘辨识算法如 CGS(Classical Gram-Schmidt)算法、MGS(Modified Gram-Schmidt)算法、HT(Householder Transformation)算法及 Givens 算法等用于此一模型的辨识. 这类算法虽有简便易于操作之优点，但有一共同的缺陷，即易受舍入误差影响而导致正交性的损失，并导致差的参数估计精度.

本文提出了两种新的有效的最小二乘辨识算法——改进的双对角化最小二乘算法 MBLS-Ⅰ与 MBLS-Ⅱ，有效地将舍入误差的恶劣影响削减到最低限度，几近于无. 同时，基于本文两个算法及 SVD 算法，构造出了一种新的 NARMAX 模型结构与参数辨识的一体化算法. 给出的仿真结果证明了本文算法的优越性.

2 NARMAX 模型

NARMAX 模型可表述为

收稿日期 1996-01-02

$$y(t)=f(y(t-1),\cdots,y(t-n_y),u(t-1),\cdots,\\ u(t-n_u),\xi(t-1),\cdots,\xi(1-n_\xi))+\xi(t). \tag{1}$$

这里 $y(t)=[y_1(t),y_2(t),\cdots,y_r(t)]^{\mathrm{T}}$, $u(t)=[u_1(t),u_2(t),\cdots,u_p(t)]^{\mathrm{T}}$, $\xi(t)=[\xi_1(t),\xi_2(t),\cdots,\xi_r(t)]^{\mathrm{T}}$ 分别是系统的输出、输入和噪声，而 n_y,n_u,n_ξ 分别是它们的最大延迟；$\{\xi(t)\}$是一个零均值的独立过程；$f(.)$是向量值非线性函数．(1)式可展开为 r 个标量方程的形式

$$y_i(t)=\theta_{i0}+\sum_{j=1}^{n_i}\theta_{ij}x_{ij}(t)+\xi_i(t),\quad i=1,2,\cdots,r.$$

这里方程的右边为 L_i 阶多项式，其中 θ_{i0}为常数项，$x_{ij}(t),i=1,2,\cdots,r,j=1,2,\cdots,n_i$ 是单项式，每一项均由带延迟的输出、输入和噪声构成，而 $n_i=\sum_{j=0}^{L_i}n_{ij},n_{i0}=1,n_{ij}$是 j 阶多项式的项数．由上可清楚地看出，NARMAX 模型的辨识本质上是一个最小二乘问题，可用各种最小二乘算法去解决．

3 算法 MBLS-Ⅰ和 MBLS-Ⅱ及其收敛性证明

MBLS-Ⅰ算法、MBLS-Ⅱ算法是在双对角化算法的基础上提出的．双对角化算法是由 Golub 和 Kahan 于1965年首次引进的[8]，目的是用其计算矩阵的奇异值．该算法在理论上是有限步收敛的迭代算法，但在实际应用中却由于舍入误差的影响一般是发散的．算法可简述如下：

对于给定的一个 $m\times n$ 矩阵 A 和一个 m 维的初始向量 $u_1,\|u_1\|=1$，双对角化算法是一个生成 m 维正交向量序列$\{u_i\}$和 n 维正交向量序列$\{v_i\}$的迭代过程，

$$\alpha_i v_i=A^{\mathrm{T}}u_i-\beta_i v_{i-1},\quad \beta_1 v_0=0, \tag{2}$$

$$\beta_{i+1}u_{i+1}=Av_i-\alpha_i u_i,\quad i=1,2,\cdots. \tag{3}$$

这里 α_i,β_i 是使$\|u_i\|=\|v_i\|=1$的非负参数．假定对 $i=1,2,\cdots,k,\alpha_i$ 和 β_i 均是非零的，则上述迭代过程可执行 k 步．令

$$U\triangleq[u_1,u_2,\cdots,u_k],\quad V\triangleq[v_1,v_2,\cdots,v_k],\quad L\triangleq\begin{bmatrix}\alpha_1 & 0 & \cdots & \cdots & 0\\ \beta_2 & \alpha_2 & 0 & \cdots & \vdots\\ 0 & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \beta_{k-1} & \alpha_{k-1} & 0\\ 0 & \cdots & 0 & \beta_k & \alpha_k\end{bmatrix}. \tag{4}$$

则迭代过程可表示为矩阵形式

$$A^{\mathrm{T}}U=VL^{\mathrm{T}},\quad AV=UL+\beta_{k+1}u_{k+1}e_k^{\mathrm{T}}. \tag{5}$$

这里 e_k 是单位矩阵 I_k 的最后一列．如果 α_k 也是非零的，则迭代过程还可再进行一步．令 $\widetilde{U}\triangleq[U,u_{k+1}],\widetilde{L}\triangleq[L^{\mathrm{T}},\beta_{k+1}e_k]^{\mathrm{T}}$，则迭代过程的矩阵形式可表示为

$$A^{\mathrm{T}}\widetilde{U}=V\widetilde{L}^{\mathrm{T}}+\alpha_{k+1}v_{k+1}e_{k+1}{}^{\mathrm{T}},\quad AV=\widetilde{U}\widetilde{L}. \tag{6}$$

可以证明，$U^{\mathrm{T}}U=V^{\mathrm{T}}V=I_k$，因而向量序列$\{u_i\}$和$\{v_i\}$必是正交的，从而迭代过程必将在某个 k 值，$k\leq\min(m,n)$上终止，算法在不考虑存在舍入误差的情况下是收敛的．迭代过程只可能有两种终态：

$$A^{\mathrm{T}}U = VL^{\mathrm{T}}, AV = UL, \quad U^{\mathrm{T}}U = V^{\mathrm{T}}V = I_k, \tag{7}$$

$$A^{\mathrm{T}}\tilde{U} = V\tilde{L}^{\mathrm{T}}, AV = \tilde{U}\tilde{L}, \tilde{U}^{\mathrm{T}}\tilde{U} = I_{k+1}, \quad V^{\mathrm{T}}V = I_k. \tag{8}$$

如果 $u_1 \in \mathscr{R}(A)$,则迭代终态必为(7)式. 详细证明可参看文[8].

实际计算中,由于舍入误差的不可避免,向量间的正交性很快就会丧失,从而导致双对角化算法的发散. 文[9]基于双对角化算法粗略地给出了一种最小二乘算法,并由双对角化算法理论上的收敛性证明简单地导出算法理论上也是收敛的. 事实上,由于舍入误差的影响,文[9]算法与其作为基础的双对角化算法同样是发散的,并一直被认为不实用. 基于此,本文提出两种改进算法——MBLS-Ⅰ算法与 MBLS-Ⅱ算法,并在存在舍入误差的情况下,给出了算法的收敛性证明.

MBLS 算法详述如下:

对于给定的 $m\times n$ 矩阵 A 及 n 维向量 b, 线性最小二乘问题也可表述为:求 x,使得 $J(x)=\|A^{\mathrm{T}}x-b\|^2$达至极小. 或等价地

求 x 和 r,满足方程 $r+A^{\mathrm{T}}x=b, Ar=0$. (9)

满足方程(9)的解 x 皆被称作最小二乘解, 而其中具有最小范数的解则被称为最小范数解. 由最小二乘理论, 知最小范数解是唯一存在的,且必属于矩阵 A 的值域空间 $\mathscr{R}(A)$.

取 $u_1=Ab/\beta_1, \beta_1=\|Ab\|$. 显然 $u_1\in\mathscr{R}(A)$, 由且式(2),(3)知,对所有的 $i, u_i\in\mathscr{R}(A)$. 假定所有的向量内积计算皆由双精度乘积累加算法得到. 该算法使得在双对角化过程的每一步迭代,向量 u_i, v_i 的单位化计算几乎皆以计算机舍入精度的方式进行. 从而我们有理由假定对所有的 i,均有 $u_i^{\mathrm{T}}u_i=\|u_i\|^2=1, v_i^{\mathrm{T}}v_i=\|v_i\|^2=1$成立. 由于舍入误差的影响,向量之间的正交性将不可避免地遭致破坏并很快丧失,从而 $U^{\mathrm{T}}U\neq I_k$, $V^{\mathrm{T}}V\neq I_k$,而 α_i, β_i 一般将永不可能小到可以忽略的程度. 实际计算中甚至连较小的 α_i, β_i 都很少见,迭代过程将无限进行下去而永无终止.

引理1. 对所有的 $i\geq 2$,我们有 $\alpha_i v_i^{\mathrm{T}}v_{i-1}=\alpha_{i-1}u_i^{\mathrm{T}}u_{i-1}, \beta_i v_i^{\mathrm{T}}v_{i-1}=\beta_{i+1}{u_{i+1}}^{\mathrm{T}}u_i$.

证明. 由(2),(3)式得

$$\beta_i = u_i^{\mathrm{T}}Av_{i-1} - \alpha_{i-1}u_i^{\mathrm{T}}u_{i-1} = {v_{i-1}}^{\mathrm{T}}A^{\mathrm{T}}u_i - \alpha_i v_i^{\mathrm{T}}v_{i-1},$$

$$\alpha_i = u_i^{\mathrm{T}}Av_i - \beta_i {u_{i+1}}^{\mathrm{T}}u_i = v_i^{\mathrm{T}}A^{\mathrm{T}}u_i - \beta_i v_i^{\mathrm{T}}v_{i-1}.$$

从而该引理成立. 引理1得证.

假设迭代过程在第 k 步被强制停止,此时 $\beta_{k+1}\neq 0$,则相应地,进行至第 k 步的迭代过程的矩阵形式为 $A^{\mathrm{T}}U=VL^{\mathrm{T}}, AV=UL+\beta_{k+1}u_{k+1}e_k^{\mathrm{T}}$,这里 L 是前已定义的 $k\times k$ 阶非奇异双对角矩阵. 令 $\beta_1 e_1=LZ, Z^{\mathrm{T}}\doteq[\zeta_1,\zeta_2,\cdots,\zeta_k], W\doteq UL^{-\mathrm{T}}$, 经简单的计算可知 $\zeta_1=\beta_1/\alpha_1$, $\zeta_i=-(\beta_i/\alpha_i)\zeta_{i-1}$,且迭代过程每进行一步,通过解方程 $LW^{\mathrm{T}}=U^{\mathrm{T}}$ 即可计算出 W 的一列. 显然, $A^{\mathrm{T}}W=A^{\mathrm{T}}UL^{-\mathrm{T}}=V$.

引理2. 对所有的 i,均有 $v_i^{\mathrm{T}}b=\zeta_i+v_i^{\mathrm{T}}(\zeta_1 v_1+\cdots+\zeta_{i-1}v_{i-1})$.

证明. 由归纳法证明如下:

1) 由引理1易证,$i=1,2$时引理2成立.

2) 假定引理2对 $i-1$成立,亦即 ${v_{i-1}}^{\mathrm{T}}b=\zeta_{i-1}+{v_{i-1}}^{\mathrm{T}}(\zeta_1 v_1+\cdots+\zeta_{i-2}v_{i-2})$.

下面证明引理2对 i 亦成立.

$$\boldsymbol{v_i^{\mathrm{T}}b = (1/\alpha_i)(A^{\mathrm{T}}u_i - \beta_i v_{i-1})^{\mathrm{T}}b = -(\beta_i/\alpha_i){v_{i-1}}^{\mathrm{T}}b + (1/\alpha_i)u_i^{\mathrm{T}}Ab}$$

$$= \zeta_i - (\beta_i/\alpha_i)v_{i-1}{}^{\mathrm{T}}(\zeta_1 v_1 + \cdots + \zeta_{i-2}v_{i-2}) + (\beta_1/\alpha_i)u_i^{\mathrm{T}}u_1.$$

由 $\alpha_i v_i = A^{\mathrm{T}}u_i - \beta_i v_{i-1}$，易知 $\beta_i v_{i-1} = A^{\mathrm{T}}u_i - \alpha_i v_i$，故有

$$v_i^{\mathrm{T}}b = \zeta_i + v_i^{\mathrm{T}}(\zeta_1 v_1 + \cdots + \zeta_{i-2}v_{i-2}) - (1/\alpha_i)u_i^{\mathrm{T}}(\zeta_1 A v_1 + \cdots + \zeta_{i-2}Av_{i-2} - \beta_1 u_1).$$

由 $\beta_i u_i = Av_{i-1} - \alpha_{i-1}u_{i-1}$，易知 $Av_{i-1} = \beta_i u_i + \alpha_{i-1}u_{i-1}$，故有

$$v_i^{\mathrm{T}}b = \zeta_i + v_i^{\mathrm{T}}(\zeta_1 v_1 + \cdots + \zeta_{i-1}v_{i-1}).$$

故引理2对 i 也是成立的.

由1)，2)及归纳法原理知，引理2对所有的 i 成立. 引理2得证.

令 $\eta_i \doteq v_i^{\mathrm{T}}b - v_i^{\mathrm{T}}(\zeta_1 v_1 + \cdots + \zeta_{i-1}v_{i-1})$，则由引理2，$\zeta_i = \eta_i$. 实际计算中，$\zeta_i$ 与 η_i 之间的误差一般小于10^{-12}. 令 $x_i \doteq WZ_i$，$Z_i^{\mathrm{T}} \doteq [\zeta_1, \zeta_2, \cdots, \zeta_i, 0, \cdots, 0]$. 算法 MBLS-Ⅰ与 MBLS-Ⅱ可给出如下，其中括弧内为 MBLS-Ⅱ算法对应于 MBLS-Ⅰ算法的不同之处.

(i) 初始化：$i \Leftarrow 1$；

$\beta_1 u_1 \Leftarrow Ab$；$\alpha_1 v_1 \Leftarrow A^{\mathrm{T}}u_1$；

$w_1 \Leftarrow u_1/\alpha_1$；

$\zeta_1 \Leftarrow \beta_1/\alpha_1$；$\eta_1 \Leftarrow v_1{}^{\mathrm{T}}b$；[$\varphi_1 = \eta_1$；]

$x_1 \Leftarrow \zeta_1 w_1$；$h_1 \Leftarrow \zeta_1 v_1$；[$x_1 \Leftarrow \eta_1 w_1$；$d_1 \Leftarrow \eta_1 v_1$；]

$J(x_1) \Leftarrow \|A^{\mathrm{T}}x_1 - b\|^2$；

(ii) 迭代步骤：$i \Leftarrow i+1$；

$\beta_i u_i \Leftarrow Av_{i-1} - \alpha_{i-1}u_{i-1}$；

$\alpha_i v_i \Leftarrow A^{\mathrm{T}}u_i - \beta_i v_{i-1}$；

$w_i \Leftarrow (u_i - \beta_i w_{i-1})/\alpha_i$；

$\zeta_i \Leftarrow -(\beta_i/\alpha_i)\zeta_{i-1}$；$\eta_i \Leftarrow v_i^{\mathrm{T}}(b - h_{i-1})$；[$\varphi_i \Leftarrow v_i^{\mathrm{T}}(b - d_{i-1})$；]

$x_i \Leftarrow x_{i-1} + \zeta_i w_i$；$h_i = h_{i-1} + \zeta_i v_i$；[$x_i \Leftarrow x_{i-1} + \eta_i w_i$；$d_i = d_{i-1} + \eta_i v_i$；]

$J(x_i) \Leftarrow \|A^{\mathrm{T}}x_i - b\|^2$；

如果($J(x_i) > J(x_{i-1})$)，

$x_i \Leftarrow x_{i-1} + \eta_i w_i$；$h_i = h_{i-1} + \eta_i v_i$；[$x_i \Leftarrow x_{i-1} + \varphi_i w_i$；$h_i = d_i$；$d_i = d_{i-1} + \varphi_i v_i$；]

$J(x_i) \Leftarrow \|A^{\mathrm{T}}x_i - b\|^2$；

如果($J(x_i) > J(x_{i-1})$)，

$x_i \Leftarrow x_{i-1}$；$h_i = h_{i-1}$；[$x_i \Leftarrow x_{i-1}$；$d_i = d_{i-1}$；]

$J(x_i) \Leftarrow J(x_{i-1})$；

(iii)如果 $J(x_i)$稳定，则迭代停止；否则，转(ii)继续.

定理1. 算法 MBLS-1与算法 MBLS-Ⅱ是收敛的，且所得最小二乘解为最小范数解.

证明. 我们只证明算法 MBLS-Ⅰ的收敛性，算法 MBLS-Ⅱ的收敛性可类似给出.

易知 $J(x) = \langle A^{\mathrm{T}}x - b, A^{\mathrm{T}}x - b\rangle = x^{\mathrm{T}}AA^{\mathrm{T}}x - 2x^{\mathrm{T}}Ab + b^{\mathrm{T}}b$. 假定算法 MBLS-Ⅰ在第 $i-1$步迭代所得近似解为 x_{i-1}，$x_{i-1} = \zeta_1 w_1 + \cdots + \zeta_{i-1}w_{i-1}$，而在第 i 步迭代，近似解取为 $x_i = x_{i-1} + tw_i$，则有

$$\begin{aligned} J(x_i) &= J(x_{i-1} + tw_i) = J(x_{i-1}) + t^2 w_i^{\mathrm{T}}AA^{\mathrm{T}}w_i + 2tw_i^{\mathrm{T}}AA^{\mathrm{T}}x_{i-1} - 2tw_i^{\mathrm{T}}Ab \\ &= J(x_{i-1}) + t^2 v_i^{\mathrm{T}}v_i + 2tv_i^{\mathrm{T}}(\zeta_1 v_1 + \cdots + \zeta_{i-1}v_{i-1}) - 2tv_i^{\mathrm{T}}b \\ &= J(x_{i-1}) + t^2 - 2t[v_i^{\mathrm{T}}b - v_i^{\mathrm{T}}(\zeta_1 v_1 + \cdots + \zeta_{i-1}v_{i-1})] \end{aligned}$$

$$= J(x_{i-1}) + t^2 - 2t\eta_i = J(x_{i-1}) + (t - \eta_i)^2 - \eta_i^2.$$

我们的目的是寻找 x 使得 $J(x)=\|A^{\mathrm{T}}x-b\|^2$ 达至极小．由引理2，选择 $t=\zeta_i$ 可使得迭代过程在迭代的初始阶段满足 $t^2-2t\eta_i<0$ 及 $J(x_i)\leq J(x_{i-1})$；随着迭代过程的继续，$J(x_i)$将变得越来越小，从而在迭代的后期阶段，舍入误差的影响变得越来越大，初始的选择 $t=\zeta_i$ 已不能保证 $J(x_i)\leq J(x_{i-1})$的恒成立．此时算法 MBLS-Ⅰ由迭代步骤(ii)给出的 t 的选择仍保持了 $J(x_i)\leq J(x_{i-1})$的恒成立．从而 $J(x_i)$将随着迭代过程的无限迭代将越来越小，直至稳定在其极小值上（一般小于10^{-25}，在舍入精度内为零）．同时，有 $\zeta_i\to 0$，$\eta_i\to 0$．

由于对所有的 i，$w_i\in\mathscr{R}(A)$，从而对所有的 i 亦有 $x_i\in\mathscr{R}(A)$．由引理1可知，算法所得到的解为最小范数解．定理1得证．

算法 MBLS-Ⅰ每步迭代需存储三个 m 维向量 u_i，w_i，x_i 和两个 n 维向量 v_i，h_i，算法 MBLS-Ⅱ需另外多存储一个 n 维向量 d_i．由于算法的迭代形式，所需存储空间大大地缩小了．此外，算法的迭代公式非常简单，从而每步迭代所需计算量也很小．实际计算中，两种算法皆具有非常好的数值性质，而后者更佳．这是由于随着迭代步骤的无限进行，φ_i 与 η_i 之间的误差值将逼近于零的缘故．但后者需要存储空间及计算量要稍多些．

4　基于算法 MBLS-Ⅰ，MBLS-Ⅱ及 SVD 的 NARMAX 模型辨识算法

假定 NARMAX 模型是单输入单输出的，且常数项为零，则(1)式化为 $y(t)=\sum_{i=1}^{n}\theta_i x_i(t)+\zeta(t)$．基于本文两个算法及 SVD 算法辨识 NARMAX 模型的结构与参数的一体化新算法如下：

第一步．由算法 MBLS-Ⅰ或 MBLS-Ⅱ或 SVD 算法计算设定模型的 n 个参数$\{\theta_i\}$，并计算 $\|\theta_i x_i\| = \left\| \sum_{t=1}^{N}\theta_i x_i(t)\right\|$ 及所得残差序列的均方根值$[err]_1$，这里 N 为数据采样长度．由$\|\theta_i x_i\|$的值从大到小的顺序重排模型项．不失一般性，仍记重排后的模型项为 $x_i(t)$，$i=1,2,\cdots,n$．删除最末一项 $x_n(t)$．

第二步．由算法 MBLS-Ⅰ或 MBLS-Ⅱ或 SVD 算法计算第一步所得简化模型的 $n-1$ 个参数$\{\theta_i\}$，并计算$\|\theta_i x_i\|$及相应残差序列的均方根值$[err]_2$，由$\|\theta_i,x_i\|$的值从大到小的顺序重排模型项．无妨仍记重排后的模型项为 $x_i(t)$，$i=1,2,\cdots,n-1$．删除最末一项 $x_{n-1}(t)$．如果$[err]_2>[err]_k$ 且$[err]_2-[err]_1>\rho$，ρ 为设定的理想阈值，则第一步模型项 $x_n(t)$的删除是错误的，应在随后的待简化模型中加入此项．

之后的模型简化迭代步骤与上述完全类似．迭代将在第 $k(\leq n)$步上终止，当且仅当$[err]_k>[err]_1$且$[err]_2-[err]_1>\rho$，这里 ρ 为设定的理想阈值．过程模型确定之后，以残差序列代替噪声可由与上类似的迭代方法确定噪声模型．之后由新的残差序列开始反复进行残差迭代，直至残差序列近似满足白噪声的性质．

5　仿真例子

例1．众所周知，10阶 Hilbert 矩阵 H_{10}为异常病态的矩阵，其具体形式可参看一般的

数值分析教科书[10]. 给定向量 b 为矩阵 H_{10}的各列之和，则方程 $H_{10}x=b$ 的解为 $x=[1,1,\cdots,1]^{\mathrm{T}}$. 以算法 HT,SVD，MBLS-Ⅰ,MBLS-Ⅱ 所得解分别为

$x_{\mathrm{HT}}=[1.709597,51.415447,\times,\times,\times,\times,\times,\times,\times,0.999969]^{\mathrm{T}}$;

$x_{\mathrm{SVD}}=[1.000000,1.00000,0.999994,1.000052,0.999751,1.000688,0.998867,1.001098,0.999422,1.000127]^{\mathrm{T}}$;

$x_{\mathrm{MBLS\text{-}I}}=[1.000000,1.00000,1.000000,0.999999,1.000003,0.999998,0.999995,1.000011,0.999991,1.000003]^{\mathrm{T}}$;

$x_{\mathrm{MBLS\text{-}II}}=[1.000000,1.00000,1.000000,0.999998,1.000006,0.999994,0.999995,1.000018,0.999985,1.000004]^{\mathrm{T}}$. 这里算法解 $x_{\mathrm{HT}},x_{\mathrm{SVD}}$由 matlab3. 0给出，×号为浮点溢出数据.

例2. 仿真系统给定为[6]

$$x(t)=0.8x(t-1)+0.4\{u(t-1)+u^2(t-1)+u_3(t-1)\},$$

$$y(t)=x(t)+\sigma(t),\sigma(t)=\xi(t)+0.6\xi(t-1).$$

这里 $u(t)$取作零均值方差为1.0的独立均匀分布序列，噪声 $\xi(t)$假定为取值于(−0.3,0.3)的均匀分布的零均值高斯白噪声序列，信噪比为8.24:1. 数据采样长度取作100. 假设系统输入和输出的最大延迟为2，则系统 NARMAX 全模型中的过程模型共包含有33项. 对于噪声模型，假定系统输入和输出的最大延迟为1，而噪声的最大延迟为2，则噪声模型共有6项. 仿真系统所产生的数据被用来拟合整个模型. 由本文辨识算法及 CGS，HT,SVD 辨识算法所得结果如表1所示.

表1 几种辨识算法的比较

模型项	CGS	HT	SVD	MBLS-Ⅰ	MBLS-Ⅱ
$y(t-1)$	0.888439	0.775394	0.800290	0.799975	0.799974
$u(t-1)$	0.406391	0.403927	0.405762	0.402543	0.402543
$u^2(t-1)$	0.402521	0.400017	0.402168	0.401965	0.401966
$u^3(t-1)$	0.586225	0.393700	0.399044	0.399797	0.399798
$\sigma(t-1)$	−0.383179	−0.308808	−0.302236	−0.382732	−0.382732
$\sigma(t-2)$	−0.174788	−0.604240	0.045463	0.062478	0.062478
$y(t-1)\sigma(t-1)$	0.027698	0.043381	−0.604642	−0.663407	−0.663409
$y(t-1)\sigma(t-2)$	−0.014372	0.045038	−0.022785	−0.033554	−0.033558
$u(t-1)\sigma(t-1)$	0.009463	0.018414	0.043382	0.062275	0.062275
$u(t-1)\sigma(t-2)$	−0.085483	−0.022775	0.018414	0.036757	0.036757
残差	0.102648	0.100408	0.100408	0.099139	0.099139

参 考 文 献

[1] Leontaritis I J, Billngs S A. Input-output parametric models for nonlinear systems. Part Ⅰ: deterministic nonlinear systems; Part Ⅱ: Stochastic non-linear systems. *Int. J. Control*, 1985, **41**: 303−344.

[2] Leontaritis I J, Billings S A. Model selection and validation methods for nonlinear systems. *Ibid.*, 1987, **45**: 311−341.

[3] Chen S, Billings S A. Predicton-error estimation algorithm for nonlinear output-affine systems. *Int. J. control*, 1988, **47**: 309−332.

[4] Chen S, Billings S A. Representation of nonlinear systems: the NARMAX model. *Ibid.*, 1989, **49**: 1013−1032.

[5] CHEN S, BILLINGS S A. LUO W. Orthogonal least squares methods and their application to nonlinear system identification. *Int. J. Control*, 1989, **50**: 1873—1896.

[6] KORENBERG M J, BILLINGS S A, LIU Y P, MCILROY P J. Orthogonal parameter estimatino for non-linear stochastic systems. *Int. J. Control*, 1988, **48**: 193—210.

[7] BILLINGS S A, CHEN S, KORENBERG M J. Identification of MIMO non-linear system using a forward-regression orthogonal estimator. *Int. J. Control*, 1989, **49**: 2157—2189.

[8] GOLUB G. KAHAN. Calculating the singular values and pseudo-inverse of a matrix, *SIAM J. Num. Anal.*, 1965, **2**: 205—224.

[9] PAIGE C C. Bidiagonalization of matrices and solution of linear equations, *SIAM J. Num. Anal.*, 1974, **11**: 197—209.

[10] Golub G, Van Loan C F. Matrix Computatons. The Johns Hopkins University Press, 1983.

TWO NEW EFFECTIVE BIDIAGONALIZATION LEAST SQUARES ALGORITHMS FOR NONLINEAR SYSTEM IDENTIFICATION

WANG XIAO HAN CHONGZHAO WAN BAIWU

(*Institute of Systems Engineering, Xi'an Jiaotong University, Xi'an* 710049)

Abstract Two new effective least squares algorithms——the modified bidiagonalization least squares algorithms (MBLS-Ⅰ and MBLS-Ⅱ) are proposed in this paper. Under the condition that round-off errors exist, a convergence proof is given. They are superior to the common used least squares algorithms such as the SVD method for round-off errors have little influence to their convergence. Furthermore, based on the two algorithms and the SVD method, a new integrated algorithm for the NARMAX model's structure and parameters' identification is also proposed here. The simulation results indicate their superiority.

Key words Nonlinear system, system identification, bidiagonalization least squares.