第 14 卷第 5 期
1997 年 10 月
控制理论与应用
CONTROL THEORY AND APPLICATIONS
Vol. 14, No. 5
Oct., 1997

非线性系统辨识的改进双对角化最小二乘算法

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:本文提出了一种新的有效的非线性系统最小二乘辨识算法——改进的双对角化最小二乘(MBLS)算法.在存在舍入误差的条件下,给出了算法的收敛性证明.事实上,算法的收敛性几乎不受舍入误差的影响,算法是大范围数值稳定的.仿真结果说明了新算法的有效性.

关键词:非线性系统;系统辨识;双对角化最小二乘

1 引 言

众所周知,很大一类离散时间非线性系统可由带外生变量的非线性自回归滑动平均模型(NARMAX)来表述[1~5].由 Billings 等提出的 NARMAX 模型其展开式是一种多项式形式,并且关于参数是线性的.在给定模型结构的情况下,系统辨识问题本质上是标准的最小二乘参数估计问题,因而可用各种各样的最小二乘算法来解决.文[3]给出了一些基于正交化的最小二乘辨识算法,如 CGS(Classical Gram-Schmidt),MGS(Modified Gram-Schmidt),HT(Householder Transformation)和 Givens 算法,用于辨识非线性系统的 NARMAX 模型.这些正交估计算法有效地把结构辨识与参数估计结合起来同时进行,计算简单且实用.文[6]利用多层神经元网络并结合直交化算法给出了一种有效的非线性动态系统的辨识算法.该算法的优点是能够对系统做较好的非线性拟合并提供一个普适性的模型,缺点是由于采用了多层神经元网络使得预报误差曲面高度复杂化,给系统分析带来了困难;同时计算量大,代价昂贵.文[7]进一步提出了一类扩展的 NARMAX 模型,其模型项可包括超越函数,但仍关于参数为线性的.文[8~10]中,由 Sontag,Billings 和 Chen,Billings 和 Zhu 提出了另一类 NARMAX 模型子集,称为有理(Rational)NARMAX 模型.该模型为关于参数非线性的模型.文[9,10]基于梯度算法给出了辨识该模型的预报误差迭代算法,但文[11~13]指出了该算法的致命缺陷:计算量大,且在设定模型项数过多时将不可避免地导致实际应用上的存储困难.为克服此弊端,文[11]通过"去分母"运算仍将此模型化成了关于参数线性的形式并给出了类似于 NARMAX 模型的直交估计算法,但这个算法是有偏的;文[12,13]结合全最小二乘算法的一些技巧保证了参数估计的无偏性.所以,文[7~13]中的推广模型本质上与 NARMAX 模型同样,皆可由线性最小二乘算法给予辨识.

但是,由文[3]给出的正交化算法极大程度地依赖于向量间正交性的保持.众所周知,实际计算中舍入误差是不可避免的,从而导致正交性常在迭代几步之后即已遭致严重破坏甚至完全丧失,并导致正交化算法应用上的差的参数估计精度.同时,这些算法的数值稳定性也没有得到证明.因而寻找数值更为稳定的算法用于非线性系统的辨识是必要的.

在本文中,提出了一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(MBLS)算法.在舍入误差存在的条件下,本文给出了算法的收敛性证明.算法是大范围数值稳定的,其收敛性几乎不受舍入误差的影响.与 Billings 提出的直交化算法比较,MBLS 算法

本文于 1995 年 12 月 18 日收到.1996 年 8 月 19 日收到修改稿.

具有其收敛性不依赖于向量间正交性是否保持的优点,且算法给出的最小二乘解恰是范数最小的最小二乘解,并在大多数情况下皆可达到非常高的精度.因而将MBLS算法用于系统的NARMAX模型辨识时,可给出较正交化方法更高精度的参数估计.同时,由于MBLS算法直接基于原方程得到最小二乘解,因而优于基于求解法方程的最小二乘算法,尤其在最小二乘解不唯一的情况下,更是如此.此外,由于算法的迭代形式,所需存储空间不大;而又由于算法的迭代公式特别简单,每步迭代所需的计算量也是非常小的.因此,MBLS算法非常适用于求解工程应用中经常遇到的大型高阶的最小二乘问题.在用于系统的NARMAX模型辨识时,如果模型的非线性阶次不是很高,结合文[13]的模型相关性检验,MBLS算法可以得到成功地应用并获到较高精度的参数估计.仿真结果证明了MBLS算法理论结果的正确性及其本身的优越性.

但是,如果系统的NARMAX全模型的非线性阶次非常高,则所需要的存储空间将会变得非常大,此时MBLS算法也会遇到存储困难.为此,本文另外提出了四种新算法——MGSCGS,HTCGS,GIVCGS,MBLSCGS算法.与Billings的CGS算法比较,这四种新算法数值更为稳定,而MBLSCGS算法是数值稳定性最好的.因而利用这四种新算法并给合文[13]的模型相关性检验,可以得到较CGS算法更高精度的参数估计.仿真结果说明了这四种新算法的优越性.

自然地,上述辨识NARMAX模型的新算法可直接推广至文[7~13]中的模型辨识中去.

2 NARMAX模型

由Billings等人提出的NARMAX模型可表述如下:

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

这里 $y(t)=[y_1(t),y_2(t),\cdots,y_m(t)]^{\mathrm T}$, $u(t)=[u_1(t),u_2(t),\cdots,u_r(t)]^{\mathrm T}$ 及 $\xi(t)=[\xi_1(t),\xi_2(t),\cdots,\xi_m(t)]^{\mathrm T}$ 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,$f(\cdot)$ 是向量值非线性函数.(1)式可分解成 m 个标量方程:

$$\begin{aligned} y_i(t)=f_i(&y_1(t-1),\cdots,y_1(t-n_y),\cdots,y_m(t-1),\cdots,y_m(t-n_y),\\ &u_1(t-1),\cdots,u_1(t-n_u),\cdots,u_r(t-1),\cdots,u_r(t-n_u),\\ &\xi_1(t-1),\cdots,\xi_1(t-n_\xi),\cdots,\xi_m(t-1),\cdots,\xi_m(t-n_\xi))\\ &+\xi_i(t),\qquad i=1,2,\cdots,m. \end{aligned} \tag{2}$$

将 $f_i(\cdot)$ 展开成 L_i 阶的多项式形式,即可得到如下的NARMAX模型

$$y_i(t)=\theta_{i0}+\sum_{j=1}^{n_i}\theta_{ij}x_{ij}(t)+\xi_i(t),\quad i=1,2,\cdots,m. \tag{3}$$

这里

$$n_i=\sum_{j=0}^{L_i}n_{ij},\quad n_{i0}=1,\quad n_{i1}=[m(n_y+n_\xi)+rn_u],$$

$$n_{i2}=\{n_{i1}[m(n_y+n_\xi)+rn_u+1]\}/2=\{n_{i1}(n_{i1}+1)\}/2,$$

$$n_{ij}=\{n_{ij-1}[m(n_y+n_\xi)+rn_u+1]+2n_{ij-2}+\cdots+(j-1)n_{i1}]\}/j,\ j>2.$$

θ_{i0} 是常数项; $x_{ij}(t)$, $i=1,2,\cdots,m$, $j=1,2,\cdots,n_i$ 是单项式,每一项均由带延迟的输出、输入和噪声构成; n_{ij} 是 j 阶多项式的项的个数.

方程(3)可以写成其矩阵形式为:

$$Y_i = X_i\Theta_i + \Xi_i. \tag{4}$$

这里

$$Y_i = [y_i(1), y_i(2), \cdots, y_i(N)]^{\mathrm{T}},$$

$$\Theta_i = [\theta_{i0}, \theta_{i1}, \cdots, \theta_{in_i}]^{\mathrm{T}},$$

$$\Xi_i = [\xi_i(1), \xi_i(2), \cdots, \xi_i(N)]^{\mathrm{T}},$$

$$X_i = \begin{bmatrix} x_{i0}(1) & x_{i1}(1) & \cdots & x_{in_i}(1) \\ x_{i0}(2) & x_{i1}(2) & \cdots & x_{in_i}(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_{i0}(N) & x_{i1}(N) & \cdots & x_{in_i}(N) \end{bmatrix}.$$

这里 N 是数据采样长度，n_i+1 是未知的参数个数. 由上可很清楚地看出，NARMAX 模型的辨识本质上是一个最小二乘问题，因而可用各种最小二乘算法去解决.

3　MBLS 算法及其收敛性

本文的 MBLS 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[17]. 由于舍入误差的影响，该算法在实际计算中发散. 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

$$\alpha_i v_i = A^{\mathrm{T}}u_i - \beta_i v_{i-1}, \quad \beta_1 v_0 = 0, \tag{5}$$

$$\beta_{i+1}u_{i+1} = Av_i - \alpha_i u_i, \quad i = 1, 2, \cdots \tag{6}$$

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

$$U \triangleq [u_1, u_2, \cdots, u_k], \quad V \triangleq [v_1, v_2, \cdots, v_k], \quad L \triangleq \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_k & \alpha_k \end{bmatrix},$$

则迭代过程可表示为矩阵形式：

$$A^{\mathrm{T}}U = VL^{\mathrm{T}}, \quad AV = UL + \beta_{k+1}u_{k+1}e_k^{\mathrm{T}}, \tag{7}$$

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 $\tilde{U} \triangleq [U, u_{k+1}]$，$\tilde{L} \triangleq [L^{\mathrm{T}}, \beta_{k+1}e_k]^{\mathrm{T}}$，则迭代过程的矩阵形式可表示为：

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

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，$k\leqslant \min(m,n)$ 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：

$$A^{\mathrm{T}}U = VL^{\mathrm{T}}, \quad AV = UL, \quad U^{\mathrm{T}}U = V^{\mathrm{T}}V = I_k, \tag{9}$$

$$A^{\mathrm{T}}\tilde{U} = V\tilde{L}^{\mathrm{T}}, \quad AV = \tilde{U}\tilde{L}, \quad \tilde{U}^{\mathrm{T}}\tilde{U} = I_{k+1}, \quad V^{\mathrm{T}}V = I_k. \tag{10}$$

如果 $u_1 \in \mathbb{R}(A)$，则迭代终态必为(9)式. 详细证明可参看文[6].

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[18]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. 在考虑舍入误差的情况下，作者没有

做任何深入的讨论.事实上,由于舍入误差的影响,文[18]算法与其所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,本文提出一种改进算法—— MBLS 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 算法可详述如下:

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以如下方式表述:

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

$$r + A^{\mathrm{T}}x = b, Ar = 0. \tag{11}$$

满足方程(11)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.我们有如下引理:

引理 1 假定 x_0是方程(11)的最小范数解,则 x_0必唯一存在,且 $x_0 \in \mathbb{R}(A)$.

证 易知 $\mathbb{R}^m = \mathbb{R}(A) \bigoplus \mathbb{N}(A^{\mathrm{T}})$.假定方程(11)的解不唯一,则每一个解在空间 $\mathbb{R}^m$中都有唯一的分解,而每一个解在值域空间 $\mathbb{R}(A)$ 中的投影也是方程(11)的解.

假设 x_0^1, x_0^2是方程(11)的两个不同解,$x_0^1 \neq x_0^2$,并在空间 $\mathbb{R}^m$中有唯一分解:$x_0^i = x_0^{0i} + f_i$,$x_0^{0i} \in \mathbb{R}(A)$,$f_i \in \mathbb{N}(A^{\mathrm{T}})$,$i = 1,2$.显然 $x_0^1 - x_0^2 \in \mathbb{N}(A^{\mathrm{T}})$,从而 $x_0^{01} - x_0^{02} \in \mathbb{N}(A^{\mathrm{T}})$,故 $x_0^{01} - x_0^{02} \in \mathbb{R}(A) \bigcap \mathbb{N}(A^{\mathrm{T}}) = \{0\}$,$x_0^{01} = x_0^{02} \triangleq x_0$.

假设 x^*是方程(11)的任意解,则由上述,x^*有唯一分解:$x^* = x_0 + f^*$,$f^* \in \mathbb{N}(A^{\mathrm{T}})$.显然,$\|x^*\|^2 = \|x_0\|^2 + \|f^*\|^2 \geq \|x_0\|^2$,从而引理1成立. 证毕.

取 $u_1 = Ab/\beta_1$,$\beta_1 = \|Ab\|$.显然 $u_1 \in \mathbb{R}(A)$,且由(5),(6),易知对所有的 i,$u_i \in \mathbb{R}(A)$.这里我们假定在迭代过程的每一步,向量 u_i, v_i的单位化计算都是精确进行的,因而对所有的 i,均有 $u_i^{\mathrm{T}}u_i = \|u_i\|^2 = 1$,$v_i^{\mathrm{T}}v_i = \|v_i\|^2 = 1$.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严重的破坏并很快丧失,从而 $U^{\mathrm{T}}U \neq I_k$,$V^{\mathrm{T}}V \neq I_k$,而 α_i, β_i将永远不可能小到可以忽略的程度.事实上,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 2 对所有的 $i \geq 2$,均有 $\alpha_i v_i^{\mathrm{T}} v_{i-1} = \alpha_{i-1} u_i^{\mathrm{T}} u_{i-1}$,$\beta_i v_i^{\mathrm{T}} v_{i-1} = \beta_{i+1} u_{i+1}^{\mathrm{T}} u_i$.

证 由(5)式和(6)式,易知:

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

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

故该引理成立. 证毕.

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可表示为:

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

这里 L 是前已定义的 $k \times k$ 阶非奇异双对角矩阵.令 $\beta_1 e_1 = LZ$,$Z^{\mathrm{T}} \triangleq [\xi_1, \xi_2, \cdots, \xi_k]$,$W \triangleq UL^{-H}$,经简单的计算可知 $\xi_1 = \beta_1/\alpha_1$,$\xi_i = -(\beta_i/\alpha_i)\xi_{i-1}$,且迭代方程每进行一步,通过解方程 $LW^H = U^H$即可计算出 W 的一列.显然,$A^{\mathrm{T}}W = A^{\mathrm{T}}UL^{-\mathrm{T}} = V$.

引理 3 对所有的 $i \geq 2$,均有 $v_i^{\mathrm{T}}b = \xi_i + v_i^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-1}v_{i-1})$.

证 由归纳法证明如下:

a) 由引理 2,易证 $i = 1,2$ 时引理 3 成立.

b) 假定引理 3 对 $i - 1$ 成立,亦即

$$v_{i-1}^{\mathrm{T}}b = \xi_{i-1} + v_{i-1}^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-2}v_{i-2}).$$

下面我们证明引理 3 对 i 亦成立

$$v_i^{\mathrm{T}}b = (1/\alpha_i)(A^{\mathrm{T}}u_i - \beta_i v_{i-1})^{\mathrm{T}}b$$
$$= -(\beta_i/\alpha_i)\, v_{i-1}{}^{\mathrm{T}}b + (1/\alpha_i)\, u_i{}^{\mathrm{T}}Ab$$
$$= \xi_i - (\beta_i/\alpha_i)\, v_{i-1}{}^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-2}v_{i-2}) + (\beta_1/\alpha_i)\, u_i{}^{\mathrm{T}}u_1.$$

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

$$v_i^{\mathrm{T}}b = \xi_i + v_i^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-2}v_{i-2}) - (1/\alpha_i)\, u_i{}^{\mathrm{T}}(\xi_1 Av_1 + \cdots + \xi_{i-2}Av_{i-2} - \beta_1 u_1).$$

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

$$v_i^{\mathrm{T}}b = \xi_i + v_i^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-1}v_{i-1}).$$

故引理 3 对 i 也是成立的.

由 a)，b) 及归纳法原理，知引理 3 对所有的 i 成立. 引理得证.

令 $\eta_i \triangleq v_i^{\mathrm{T}}b - v_i^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-1}v_{i-1})$. 实际计算中，$\xi_i$ 与 η_i 的误差一般小于 10^{-10}. 令 $x_i \triangleq WZ_i, Z_i^{\mathrm{T}} \triangleq [\xi_1, \xi_2, \cdots, \xi_i, 0, 0, \cdots, 0]$，则 $r_i \triangleq b - A^{\mathrm{T}}x_i = b - VZ_i$，因而迭代过程每进行一步，残差 r_i 也可计算出来. 显然 $J(x_i) = \| A^{\mathrm{T}}x_i - b \|^2 = r_i^2$. 不计算每一步的残差 r_i，MBLS 算法可给出如下：

i）初始化：

$i \Leftarrow 1$；

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

$w_1 \Leftarrow u_1/\alpha_1$；

$\xi_1 \Leftarrow \beta_1/\alpha_1$；　$\eta_1 \Leftarrow v_1{}^{\mathrm{T}}b$；

$x \Leftarrow \eta_1 w_1$；

ii）迭代过程：

$i \Leftarrow i + 1$；

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

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

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

$\xi_i \Leftarrow -(\beta_i/\alpha_i)\xi_{i-1}$；　$\eta_i \Leftarrow v_i{}^{\mathrm{T}}b - v_i{}^{\mathrm{T}}(\xi_1 v_1 + \cdots + \xi_{i-1}v_{i-1})$；

iii）如果 $|\eta_i| < \epsilon$，停止；否则

$x \Leftarrow x + \eta_i w_i$；

go to ii)；

这里 ϵ 是一任意小的正数. 基于引理 2，3，我们得到关于 MBLS 算法收敛性的如下定理：

定理　MBLS 算法是收敛的，且收敛所得解恰是具有最小范数的最小二乘解.

证　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 MBLS 算法在迭代的第 $i-1$ 步所得近似解为 x_{i-1}，$x_{i-1} = \eta_1 w_1 + \cdots + \eta_{i-1}w_{i-1}$，而在第 i 步迭代，取 $x_i = x_{i-1} + tw_i$，则由引理 3，易知

$$J(x_i) = J(x_{i-1} + tw_i)$$
$$= J(x_{i-1}) + t^2 w_i^{\mathrm{T}}AA^{\mathrm{T}}w_i + 2tw_i^{\mathrm{T}}AA^{\mathrm{T}}x_{i-1} - 2tw_i^{\mathrm{T}}Ab$$
$$= J(x_{i-1}) + t^2 v_i^{\mathrm{T}}v_i + 2tv_i^{\mathrm{T}}(\eta_1 v_1 + \cdots + \eta_{i-1}v_{i-1}) - 2tv_i^{\mathrm{T}}b$$
$$= J(x_{i-1}) + t^2 - 2t[v_i^{\mathrm{T}}b - v_i^{\mathrm{T}}(\eta_1 v_1 + \cdots + \eta_{i-1}v_{i-1})]$$
$$= J(x_{i-1}) + t_2 - 2t\eta_i$$

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

故 t 应取作 η_i，而算法的第 i 步迭代近似解应取作 $x_i=x_{i-1}+\eta_i w_i$. 相应地，此时 $J(x_i)=J(x_{i-1})-\eta_i^2$. 实际计算中，$\xi_i$ 与 η_i 间有一 10^{-10} 量级的误差，从而 $[v_i^{\mathrm{T}}b-v_i^{\mathrm{T}}(\eta_1 v_1+\cdots+\eta_{i-1}v_{i-1})]$ 并不严格等于 η_i；但是显然地，$t=\eta_i$ 已保证了 $t^2-2t[v_i^{\mathrm{T}}b-v_i^{\mathrm{T}}(\eta_1 v_1+\cdots+\eta_{i-1}v_{i-1})]<0$，从而亦保证了 $J(x_i)<J(x_{i-1})$ 的成立. 所以随着迭代过程的进行，函数 $J(x)$ 的值必将变得越来越小，并最终逼近函数 $J(x)$ 的最小值；同时，必伴随有 $\eta_i^2\to 0, x_i\to x^*$，这里 x^* 是方程(11)的解. 故 MBLS 算法必收敛.

由于对所有的 i，均有 $u_i\in\mathbb{R}(A)$，而 $w_1=u_1/\alpha_1, w_i=(u_i-\beta_i w_{i-1})/\alpha_i$，故 $w_i\in\mathbb{R}(A)$ 对所有的 i 成立. 自然地，对所有的 i，均有 $x_i\in\mathbb{R}(A)$，从而 $x^*\in\mathbb{R}(A)$. 由引理1，$x^*=x_0$.

证毕.

应当指出的是，η_i 也可由 ξ_i 所代替；但据收敛性定理的证明及实际计算经验，η_i 使得 MBLS 算法具有更好的收敛性和最小二乘解精度.

MBLS 算法需要存储三个 m 维向量 u_i, w_i, x_i 和一个 n 维向量 v_i，所需存储空间不大. 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，因而每一步迭代所需计算量也很小.

4 MBLS 算法在 NARMAX 模型辨识中的应用及另外的四种新算法

令 $X_i=A^{\mathrm{T}}$，则 MBLS 算法可直接应用于部分2中已述及的 NARMAX 模型的辨识. 每一子系统的辨识可与其它子系统分开进行. 辨识算法的执行步骤是：首先辨识系统的 NARMAX 全模型，之后根据每一模型项的存在对残差影响的大小简化模型，并由 MBLS 算法再次辨识简化模型，反复迭代即可选出正确的模型. 最后结合文[13]的模型相关性检验判断是否漏选了正确的模型项或误选了不正确的模型项，给出最终的辨识模型. 如果系统的 NARMAX 全模型的非线性阶次不是非常高，MBLS 算法运行良好. 由于其大范围的数值稳定性及不受舍入误差及向量间正交性是否保持的影响的优点，MBLS 算法可给出较 Billings 直交化算法更高精度的参数估计. 仿真结果证明了 MBLS 算法理论结果的正确性及其本身的优越性.

但是，如果系统的 NARMAX 全模型具有太高的非线性阶次，则所需的存储空间将变得非常大，此时 MBLS 算法也会遇到存储困难的问题. 文[3]中给出的 CGS 算法在克服这个困难方面有其独到的长处. 该算法可以边选项边删去不重要的项，不要求存储 NARMAX 全模型的整个系数矩阵. 但正如文[3]所指出的，CGS 算法是数值不稳定的，而 MGS, HT, Givens 算法虽是数值稳定的，但却要求存储 NARMAX 全模型的整个系数矩阵，从而与 MBLS 算法同样，也存在存储困难的问题. 所以本文基于 CGS 算法与 MGS, HT, Givens, MBLS 算法的结合，提出了四种改进的新算法——MGSCGS, HTCGS, GIVCGS, MBLSCGS 算法，用于辨识具有较高非线性阶次的非线性系统. 这些算法的执行步骤是：首先，由 CGS 算法确定系统过程模型的结构，亦即决定过程模型中有哪些项存在；之后则分别由 MGS 或 HT 或 Givens 或 MBLS 算法给出过程模型参数的估计，并决定其噪声特性，对过程模型参数估计进行更好的修正. 最后由文[13]的模型相关性检验判断辨识出的模型是否适当. 与 CGS 算法比较，这四种算法数值更为稳定并运作良好，而 MBLSCGS 算法是数值稳定性最好的. 所以由这四种新算法可给出比 CGS 算法更好的参数估计精度. 仿真结果证明了这四种算法的有效性.

自然地，上述辨识 NARMAX 模型的新算法可直接推广至文[7～13]中的模型辨识中去.

5 仿真研究

例 1 众所周知,Hilbert 矩阵是典型的异常病态的矩阵. 三阶 Hilbert 矩阵及它的逆可给出如下

$$H_3=\begin{bmatrix}1 & 1/2 & 1/3\\ 1/2 & 1/3 & 1/4\\ 1/3 & 1/4 & 1/5\end{bmatrix},\quad H_3^{-1}=\begin{bmatrix}9 & -36 & 30\\ -36 & 192 & -180\\ 30 & 180 & 180\end{bmatrix}.$$

易知 $\|H_3\|_\infty=11/6$, $\|H_3^{-1}\|_\infty=408$, 从而矩阵的条件数 $\mathrm{con}(H_3)_\infty=748$, 显然三阶 Hilbert 矩阵 H_3 已是非常病态的了. 我们解如下方程:

$$H_3x=b.$$

则方程的解 x 必将对舍入误差极为敏感. 取向量 b 分别等于 $[11/6,13/12,47/60]^T$ 和 $[1.0,0.0,0.0]^T$, 则方程的真实解及由算法 CGS,MGS,MBLS 分别得到的近似解可列出如表 1,2 所示. 显然,由 CGS 算法及 MGS 算法所给出的近似解的精度非常差,根本不可接受;但由 MBLS 算法所得近似解仍具有很好的精度.

表 1　本文算法与直交化算法的数值结果比较

$b=[11/6,13/12,47/60]^T$

	real	CGS	MGS	MBLS
x_1	1.0	0.425583	0.427269	1.000001
x_2	1.0	0.989873	0.994406	1.000000
x_3	1.0	1.009777	1.005402	0.999998

表 2　本文算法与直交化算法的数值结果比较

$b=[1.0,0.0,0.0]^T$

	real	CGS	MGS	MBLS
x_1	9.0	−8.120857	−8.121325	9.001142
x_2	−36.0	−36.004086	−36.005356	−36.006153
x_3	30.0	30.003864	30.005074	30.00583

例 2 仿真系统给定为[4]:

$$x(t)=0.8x(t-1)+0.4\{u(t-1)+u^2(t-1)+u^3(t-1)\},$$
$$\sigma(t)=\xi(t)+0.6\xi(t-1),$$
$$y(t)=x(t)+\sigma(t).$$

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 CGS,MGSCGS,MBLSCGS 算法所得的参数估计分别列示于表 3.

显然 MGSCGS,MBLSCGS 算法优于 CGS 算法,而 MBLSCGS 算法给出的参数估计精度最好. HTCGS 和 GIVCGS 算法与 MGSCGS 算法具有相同的数值稳定性,但使用时需要更多的计算.

表 3 本文双对角化辨识算法与直交化辨识算法仿真结果比较

模型项	CGS	MGSCGS	MBLSCGS
$y(t-1)$	0.888439	0.775394	0.800293
$u(t-1)$	0.406391	0.403927	0.405765
$u^2(t-1)$	0.402521	0.400017	0.402167
$u^3(t-1)$	0.586225	0.393700	0.399042
$\sigma(t-1)$	−0.383179	−0.308808	−0.302201
$\sigma(t-2)$	−0.174788	−0.604240	−0.604615
$y(t-1)\sigma(t-1)$	0.027698	0.04338	0.043363
$y(t-1)\sigma(t-2)$	−0.014372	0.045038	0.045449
$u(t-1)\sigma(t-1)$	0.009463	0.018414	0.018430
$u(t-1)\sigma(t-2)$	−0.085483	−0.022775	−0.022790
残差	0.102648	0.100408	0.100408

6 结论

本文提出了一种新的有效的非线性系统辨识算法—— MBLS 算法. 在舍入误差存在的情况下,本文给出了算法的收敛性证明. 算法是大范围数值稳定的,且其收敛性几乎不受舍入误差的影响. 与文[3]中的正交化算法如 CGS,MGS 和 HT 算法相比较,本文算法具有不受因舍入误差的存在而使得正交性遭至破坏乃至完全丧失的破坏性影响的优点,因而可给出更高精度的参数估计. 仿真结果证明了关于 MBLS 算法理论结果的正确性. 为了辨识具有非常高的非线性阶次的复杂非线性系统,本文又另外提出了四种新的算法——MGSCGS,HTCGS,GIVCGS,MBLSCGS 算法. 这些算法都较 CGS 算法数值更为稳定,而 MBLSCGS 算法是数值稳定性最好的. 仿真结果说明了这些新算法的有效性.

参考文献

1 Leontaritis, I. J. and Billings, S. A.. Input-output parametric models for nonlinear systems. Part Ⅰ: deteministic nonlinear systems; Part Ⅱ: Stochastic non-linear systems. Int. J. Control, 1985, 41(2): 303—344; Model selection and validation methods for nonlinear systems. Ibid, 1987, 45(1): 311—341

2 Chen, S. and Billings, S. A.. Prediction-error estimation algorithm for nonlinear output-affine systems. Int. J. Control, 1988, 47(1): 309—332; Representation of nonlinear systems: the NARMAX model. Ibid., 1989, 49(3): 1013—1032

3 Chen, S., Billings, S. A. and Luo, W.. Orthogonal least squares methods and their application to nonlinear system identification. Int. J. Control, 1989, 50(5): 1873—1896

4 Korenberg, M. J., Billings S. A., Liu, Y. P. and Mcilroy P. J.. Orthogonal parameter estimation for non-linear stochastic systems. Int. J. Control, 1988, 48(1): 193—210

5 Billings, S. A., Chen, S. and Korenberg, M. J.. Identification of MIMO non-linear system using a forward-regression orthogonal estimator. Int. J. Control, 1989, 49(6): 2157—2189

6 Chen, S. and Billings, S. A.. Neural networks for nonlinear dynamic system modelling and identification. Int. J. Control, 1992, 56(2): 319—346

7 Billings, S. A. and Chen, S.. Extended model set, global data and threshold model identification of severely nonlinear systems, Int. J. Control, 1989a, 50(5): 1897→1923

8 Sontag, E. D.. Polynomial Response Maps——Lecture Notes in Control and Information Science 13. Berlin: Springer-Verlag, 1979

9 Billings, S. A. and Chen, S.. Identification of nonlinear rational systems using a prediction error estimation algorithm. Int.

J. Systems Sci. ,1989b,20(2):467—494

10 Billings,S. A. and Zhu,Q. M.. Rational model identification using an extended least squares algorithm. Int. J. Control, 1991,54(3):529—546

11 Zhu,Q. M. and Billings,S. A.. Recursive parameter estimator for nonlinear rational models. Journal of systems Engineering,1991,1(1):63—76

12 Zhu,Q. M. and Billings,S. A.. Parameter estimator for stochastic nonlinear rational models. Int. J. Control,1993a,57(1):309—333

13 Billings,S. A. and Zhu,Q. M.. A structure detection algorithm for nonlinear rational models. Int. J. Control,1994a,59(6):1439—1463

14 Billings,S. A. and Zhu,Q. M.. Nonlinear model validation using correlation tests. Int. J. Control,1994,60(6):1107—1120

15 Billings S. A. and Zhu Q. M.. Model validation tests for multivariable nonlinear models including neural netwroks. Int J. Control,1995,62(4):749—766

16 Braess,D.. Nonlinear Approximation Theory. Berlin: Springer-Verlag,1986

17 Zhao,M. W. and Lu,Y. Z.. Parameter identification and convergence analysis based on the least squares method for a class of nonlinear systems. Int. J. Systems Sci. ,1991,22(1):33—48

18 Zhu,Q. M. and Billings,S. A.. Stability of a class of nonlinear systems. Proceedings of Cseeeuk Conference. London, UK,1993b

19 Golub,G. and Kahan. Calculating the singular values and pesudo-inverse of a matrix. SIAM J. Num. Anal. ,1965,2(1): 205—224

20 Paige C. C.. Bidiagonalization of matrices and solution of linear equations. SIAM J. Num. Anal. ,1974,11(1):197—209

The Modified Bidiagonalization Least Squares Algorithm for Nonlinear System Identification

WANG Xiao, HAN Chongzhao and WAN Baiwu

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

Abstract: A new effective least squares algorithm for nonlinear system identification has been proposed in this paper,which is referred to as the Modified Bidiagonalization Least Squares(MBLS)one. Under the condition that round-off errors exist,a convergence proof on it has been given. In fact,it is numerically stable in a large category,and round-off errors have almost no influence to its convergence. The simulation results indicate its superiority.

Key words: nonlinear system; system identification; bidiagonalization least squares

本文作者简介

王 晓 1966 年生. 1988 年于河南师范大学获数学学士学位,1993 年于西安交通大学获应用数学硕士学位,现为西安交通大学系统工程研究所系统工程专业在读博士生. 目前主要研究方向为非线性系统的辨识及频谱分析.

韩崇昭 1943 年生. 教授,博士生导师. 1968 年毕业于西安交通大学,1981 年在中国科学院研究生院获硕士学位,现为西安交通大学系统工程研究所副所长、机械制造系统工程国家重点实验室副主任、陕西省自动化学会副理事长兼秘书长. 著有"泛函分析及其在自动控制中的应用"等四部专著和教材,主要研究方向是工业大系统优化,自适应控制,非线性系统辨识及频谱分析等;国内外发表论文 40 多篇,科研成果多次获奖.

万百五 见本刊 1997 年第 2 期第 265 页.