第19卷 第3期 自动化学报 Vol. 19, No. 3
1993年5月 ACTA AUTOMATICA SINICA May, 1993

随机非线性系统辨识的正交优选算法[1)]

王秀峰 李波
(南开大学计算机与系统科学系,天津 300071) (天津大学管理工程系,天津 300072)

摘 要

本文讨论了基于 NARMAX 模型描述的非线性系统的结构确定和参数估计. 采用正交化手段,使得模型中参数估计相互独立;利用"新息-贡献"准则,并结合其他信息准则优选最有意义的项进入模型. 采用可添可删的双向回归算法,使得模型结构与选项顺序无关,从而能唯一确定模型的最优结构. 仿真例子说明了该算法的有效性.

关键词: 非线性系统,系统辨识,正交化方法,结构确定,参数估计.

一、引 言

随着对实际系统预测和控制要求的提高,在许多情况下,基于线性化的模型已不满足高标准的要求,所以非线性系统的辨识越来越受到重视. NARMAX 模型 (Non-linear Auto Regressive Moving Average model with eXogenous inputs) 为非线性系统提供了一个非常简洁的描述形式,[1,2]它可以描述非常宽的一大类非线性系统并避免了以前描述(如函数级数描述等)的许多弊病,所以基于此种模型的辨识成为八十年代以来研究的焦点. Kortmann 和 Unbehauen[3] 提出了直接选择模型结构的方法. 它的缺点是结构确定和参数估计是分别进行的,这就使得重复计算量非常大. 文献[4]改进了他们的算法,提出了结构和参数同时确定的递推算法,大大减少了计算量,但文章的结果仅适用于确定性系统. 随机非线性系统与确定性非线性系统及随机线性系统都有很大不同. 非线性随机系统由于具有输出、输入和噪声的各种幂次的交互乘积项,即使在非常弱的白噪声情况下,也会使估计产生很大偏差. 所以,非线性随机系统的辨识是很困难的问题. Korenberg 等[5]对于非线性随机系统提出了正交参数估计方法. 基本思想是引入一个辅助正交化模型,使得参数估计相互独立. 然而该算法基于向后回归,开始将所有可能的项全部含于模型中辨识,然后再根据某些准则剔除不重要的项. 但这样做不仅可能导致方程病态,更重要的是该算法所确定的模型结构与项的排列顺序有关(实际上不唯一),所以不能很好地拟合实际系统.

本文提出了双向回归算法. 即从零开始,根据该项所带来的"新息"及对输出的"贡献"逐项添加,同时也随时剔除"贡献"小的项. 不仅避免了参数估计的病态问题,而且大

本文于1991年11月5日收到.
1) 本文得到国家自然科学基金和天津自然科学基金资助.

大减少了计算量．更重要的是，该算法所确定的模型结构不依赖于选项顺序，可唯一得到最优的最终模型，大大提高了拟合精度．在选项过程中，采取边选、边检验、边正交化的方法，并且递推地实现．

二、NARMAX 模 型[1,2]

NARMAX 模型的一般 l 次多项式形式为

$$y(t)=\sum_{k=0}^{n}\theta_k x_k(t)+\varepsilon(t). \tag{1}$$

其中 $x_0(t)=1$,

$$\begin{aligned}
x_k(t)=&\,y(t-n_{y1})\cdots y(t-n_{yp})u(t-d-n_{u1}+1)\cdots u(t-d-n_{uj}+1)\\
&\varepsilon(t-n_{\varepsilon1})\cdots\varepsilon(t-n_{\varepsilon q}),\\
&p\geqslant 0,\ j\geqslant 0,\ q\geqslant 0,\ k=1,2,\cdots,n.\\
&1\leqslant n_{y1}\leqslant N_y,\cdots,\ 1\leqslant n_{yp}\leqslant N_y,\\
&1\leqslant n_{u1}\leqslant N_u,\cdots,\ 1\leqslant n_{uj}\leqslant N_u,\\
&1\leqslant n_{\varepsilon1}\leqslant N_\varepsilon,\cdots,\ 1\leqslant n_{\varepsilon q}\leqslant N_\varepsilon,
\end{aligned} \tag{2}$$

$p=0$, $j=0$, $q=0$ 分别表示项 x_k 不包含 $y(\cdot)$,$u(\cdot)$ 和 $\varepsilon(\cdot)$.

$$\begin{cases}n=\sum_{i=1}^{l}n_i,\\ n_i=[n_{i-1}(N_y+N_u+N_\varepsilon+i-1)]/i,\quad n_0=1.\end{cases} \tag{3}$$

假设取得 N 组数据 $\{y(i),u(i);i=1,\cdots,N\}$，则(1)式的矩阵形式为

$$\boldsymbol{y}=X\cdot\boldsymbol{\theta}+\boldsymbol{\varepsilon}, \tag{4}$$

其中
$$\begin{aligned}
&\boldsymbol{y}=[y(1),y(2),\cdots,y(N)]^T,\quad \boldsymbol{\theta}=[\theta_0,\theta_1,\cdots,\theta_n]^T,\\
&\boldsymbol{\varepsilon}=[\varepsilon(1),\varepsilon(2),\cdots,\varepsilon(N)]^T,\quad X=[\boldsymbol{x}_0,\cdots,\boldsymbol{x}_n],\\
&\boldsymbol{x}_i=[x_i(1),x_i(2),\cdots,x_i(N)]^T.
\end{aligned}$$

三、正交化与“新息-贡献”准则

1. 参数的正交化估计算法[5]

考虑模型(4)．对矩阵 X 的列向量施行 Gram-Schmidt 正交化．

$$\begin{cases}\boldsymbol{v}_0=\boldsymbol{x}_0=1,\\ \boldsymbol{v}_i=\boldsymbol{x}_i-\sum_{j=0}^{i-1}\alpha_{ji}\boldsymbol{v}_j,\ i=1,2,\cdots,n.\end{cases} \tag{5}$$

其中
$$\boldsymbol{v}_i=[v_i(1),v_i(2),\cdots,v_i(N)]^T,$$

$$\alpha_{ij}=\begin{cases}1, & i=j\\ \sum_{t=1}^{N}v_i(t)x_j(t)\Big/\sum_{t=1}^{N}v_i^2(t), & i<j\end{cases} \tag{6}$$

得到(4)式的辅助模型

$$y = Vg + \varepsilon, \tag{7}$$

其中矩阵 $V=[v_0,v_1,\cdots,v_n]$ 是列正交阵. 模型(7)的参数估计为

$$\hat{g} = (V^TV)^{-1}V^Ty,$$

即

$$\hat{g}_i = \sum_{t=1}^{N} v_i(t)y(t) \Big/ \sum_{t=1}^{N} v_i^2(t). \quad i = 0,1,\cdots,n. \tag{8}$$

原模型(4)中参数 $\boldsymbol{\theta}$ 可由下式计算:

$$\hat{\theta}_n = \hat{g}_n$$

$$\hat{\theta}_i = \hat{g}_i - \sum_{j=i}^{n-1} \alpha_{jj+1}\hat{\theta}_{j+1}, \quad i = n-1,\cdots,0. \tag{9}$$

2. "新息-贡献"准则

设 N 维向量 $x_1,x_2,\cdots,x_k$, 相应地正交化向量为 $v_1,v_2,\cdots,v_k$. 记 $R_k=[v_1,v_2,\cdots,v_k]$, 对任意向量 x, 则 x 可由 $v_1,\cdots,v_k$ 线性表出的部分为

$$u = R_k(R_k^TR_k)^{-1}R_k^Tx. \tag{10}$$

x 所带来的"新息"为

$$\beta = x - u \triangleq (\beta(1),\beta(2),\cdots,\beta(N))^T. \tag{11}$$

y 在 β 上的投影平方为

$$Q = \left(\sum_{t=1}^{N} \beta(t)y(t)\right)^2 \Big/ \sum_{t=1}^{N} \beta^2(t). \tag{12}$$

Q 表明了 x 对描述向量 y 的"贡献"大小.

展开(10)式后代入(11)式,并与(5)式比较易见, β 恰好是 x 与 $v_1,\cdots,v_k$ 的正交化向量. 因此,在具体计算时可将正交化与"新息-贡献"准则值的计算统一起来.

四、随机非线性系统辨识算法

算法的基本思路是: 将模型 $y = X\theta + \varepsilon$ 分成两部分,即 $y = F_1[\cdot] + F_2[\cdot]$, 其中 F_1 是确定性部分,即不含噪声($\varepsilon(\cdot)$)的项; F_2 为含有噪声的部分. 首先利用"新息-贡献"准则来选择 F_1 中的项,结合通常意义下的 F 检验和 BIC 准则来确定模型结构,同时估计参数. 在 F_1 的模型确定后,计算误差 $\varepsilon(\cdot)$ 并代替 F_2 中 $\varepsilon(\cdot)$ 项, 重复确定性部分的实施过程,得到 F_2 的估计. 再利用 F_1 和 F_2 计算误差 $\{\varepsilon(\cdot)\}$, 反复进行, 直到误差 $\{\varepsilon(\cdot)\}$ 为白噪声或小于预先给定的阈值为止. 在选择过程中,不仅添加"贡献"大的项,同时判断已选项可否被剔除, 由于采用了边选项边正交化手段,所以在剔除项时不影响其他参数的估计值. 在具体计算过程中, 作者采用了文献[4]中给出的运算技巧, 避免了大量重复计算.

具体计算步骤如下:

1) 给定(1)式中非线性幂次 l, 输出、输入和误差的阶次 N_y, N_u, N_ε 以及数据长度 N; 给出 F 检验的显著性指标 F_α; 计算所有可能项的总数 $m = n + 1$ 及确定性部分总数

n_1 和非确定性部分总数 n_2。构造矩阵 X，其中前 n_1 列为确定性项，后 n_2 列为非确定性项。

2）确定 $F_1[\cdot]$ 的结构。

（a）置 $k=1$；计算 $\boldsymbol{x}_i$ 对输出 $\boldsymbol{y}$ 的贡献

$$Q_i=\left(\sum_{t=1}^{N} y(t)x_i(t)\right)^2\Big/\sum_{t=1}^{N} x_i^2(t),\quad i=0,1,\cdots,n_1.$$

置 $Q_{\max}=\max\limits_i\{Q_i\}$，$\boldsymbol{v}_1=\tilde{\boldsymbol{x}}_1=\boldsymbol{x}_{\max}$。

（b）置 $R_k=[\boldsymbol{v}_1]$，$\Phi_k^{-1}=(R_k^TR_k)^{-1}=1\Big/\sum\limits_{t=1}^{N} v_1^2(t)$。

计算

$$g_1=\sum_{t=1}^{N} v_1(t)y(t)\Big/\sum_{t=1}^{N} v_1^2(t),\quad E_1=\boldsymbol{y}-g_1\cdot\boldsymbol{v}_1.$$

$$S_1^2=\frac{1}{N}E_1^TE_1,\quad BIC_1=N\ln S_1^2+k\cdot\ln N.$$

（c）对剩余的 n_1-k 项 $\boldsymbol{x}_i$ 分别计算

$$\boldsymbol{u}_i=R_k(R_k^TR_k)^{-1}R_k^T\boldsymbol{x}_i,$$

"新息"

$$\boldsymbol{v}_i=\boldsymbol{x}_i-\boldsymbol{u}_i,$$

"贡献"

$$Q_i=\left(\sum_{t=1}^{N} v_i(t)y(t)\right)^2\Big/\sum_{t=1}^{N} v_i^2(t),$$

参数

$$g_i=\sum_{t=1}^{N} v_i(t)y(t)\Big/\sum_{t=1}^{N} v_i^2(t).$$

（d）在以上 n_1-k 项中选出使 Q 取最大值的项，并记相应的向量 $\boldsymbol{v}_{\max}$ 为 $\boldsymbol{v}_{k+1}$，$\boldsymbol{x}_{\max}$ 为 $\tilde{\boldsymbol{x}}_{k+1}$，相应参数 $g_{\max}$ 为 g_{k+1}，置

$$R_{k+1}=[R_k\ \ \boldsymbol{v}_{k+1}],$$

$$\Phi_{k+1}^{-1}=[R_{k+1}^T R_{k+1}]^{-1}=\begin{bmatrix}\Phi_k^{-1} & 0\\ 0 & 1\Big/\sum\limits_{t=1}^{N} v_{k+1}^2(t)\end{bmatrix}.$$

计算

$$E_{k+1}=\boldsymbol{y}-\sum_{i=1}^{k+1} g_i\cdot\boldsymbol{v}_i,\ S_{k+1}^2=\frac{1}{N}E_{k+1}^TE_{k+1},$$

$$BIC_{k+1}=N\ln S_{k+1}^2+(k+1)\ln N,$$

$$F=(S_k^2-S_{k+1}^2)/S_{k+1}^2\cdot(N-k-1).$$

（e）检验信息准则值． 若 $BIC_{k+1}<BIC_k$ 且 $F\geqslant F_\alpha$ 则将 $\boldsymbol{v}_{k+1}$ 选入模型，置

$k=k+1$，继续（f）；否则确定性部分计算结束.

（f）计算已选入模型的各项当其被去掉时的信息准则值. 对于 $i=2$ 至 k 做如下计算：

去掉矩阵 $R_k=[v_1,v_2,\cdots,v_k]$ 第 i 列及其后面所有列，得到矩阵 $R_{k,i}$；去掉矩阵 Φ_k^{-1} 的第 i 列及其后的所有列得到矩阵 $\Phi_{k,i}^{-1}$；计算

$$\tilde{u}_{i+1}=R_{k,i}\Phi_{k,i}^{-1}R_{k,i}^T\tilde{x}_{i+1},\quad \tilde{v}_{i+1}=\tilde{x}_{i+1}-\tilde{u}_{i+1},$$

$$\tilde{g}_{i+1}=\sum_{t=1}^{N}\tilde{v}_{i+1}(t)y(t)\Big/\sum_{t=1}^{N}\tilde{v}_{i+1}^2(t),$$

$$R_{k,i+1}=[R_{k,i}\tilde{v}_{i+1}],\quad \Phi_{k,i+1}^{-1}=\begin{bmatrix}\Phi_{k,i}^{-1} & \\ & 1\Big/\sum_{t=1}^{N}\tilde{v}_{i+1}^2(t)\end{bmatrix}.$$

重复以上步骤则可得到 $\tilde{v}_{i+2},\cdots,\tilde{v}_k$ 和相应参数 $\tilde{g}_{i+2},\cdots,\tilde{g}_k$. 由此计算

$$E_{k,i}=y-\sum_{j=1}^{i-1}g_j\cdot v_j-\sum_{j=i+1}^{k}\tilde{g}_j\cdot\tilde{v}_j,$$

$$S_{k,i}^2=\frac{1}{N}E_{k,i}^TE_{k,i},$$

$$F_i=(S_{k,i}^2-S_k^2)/S_k^2\cdot(N-k),$$

$$BIC_{k,i}=N\ln S_{k,i}^2+(k-1)\ln N.\quad i=2,\cdots,k.$$

（g）置 $F_{\min}=\min\limits_i\{F_i\}$，$BIC_{\min}=\min\limits_i\{BIC_{k,i}\}$. 若 $BIC_{\min}\leqslant BIC_k$ 或 $F_{\min}\leqslant F$，则相应项从模型中剔除. 置 $K=k-1$，$R_k=R_{k,\min}$，$\Phi_k^{-1}=\Phi_{k,\min}^{-1}$，$BIC_k=BIC_{\min}$，$F=F_{\min}$，$v_j=\tilde{v}_{j+1}$，$\tilde{x}_j=\tilde{x}_{j+1}$，$g_j=g_{j+1}$，$j=\min,\min+1,\cdots,k$. 返回第(f)步. 否则转第(c)步.

3）确定 $F_2[\cdot]$ 的结构. 假定确定性部分 $F_1[\cdot]$ 共选了 S 项.

（a）计算误差 $\varepsilon=y-\sum\limits_{i=1}^{S}g_i\cdot v_i$.

（b）若 ε 为零向量或 $\{\varepsilon(k)\}$ 为白噪声，说明此系统为确定性非线性系统，转第 4）步；否则，对矩阵 X 的后 n_2 列中 $x_i(i=n_1+1,\cdots,n_1+n_2)$ 重复第2)步的（c）—（g）步进行选项.

（c）计算误差 $\varepsilon=y-F_1[\cdot]-F_2[\cdot]$. 若 $\{\varepsilon(k)\}$ 为白噪声或参数收敛，则转第 4)步；否则，重复第2)步，（c）—（g）步重新进行非确定部分 F_2 的选项.

4）利用(9)式计算原模型参数 θ 的估计.

注：由于噪声一般远小于信号，所以对非确定性部分选项条件适当放宽. 本文对非确定性部分仅用 F 检验准则而且临界值 F_α 取较小值.

容易证明，上述算法给出的估计是参数的无偏估计.

五、仿 真 例 子

模型 1. $y(t) = 0.5y(t-1) + u(t-1) + \xi(t) - 0.1\xi(t-1) - 0.2\xi(t-2)$.

仿真输入采用伪随机二位式序列，$\xi(t)$ 为 (0,1) 正态白噪声序列，信噪比 S/N 为 9.7:1. 样本长度 $N = 150$. 取 $N_y = N_u = N_\varepsilon = 4$, $l = 1$. 辨识得到模型为

$$\begin{aligned} y(t) = {} & 0.494766y(t-1) + 0.988995u(t-1) + \varepsilon(t) - 0.424761\varepsilon(t-1) \\ & - 0.488296\varepsilon(t-2) - 0.214439(t-3) - 0.082672\varepsilon(t-4). \end{aligned}$$

模型 2. $y(t) = 1.5y(t-1) - 0.7y(t-2) + u(t-1) + 0.5u(t-2) + \xi(t) + 0.6\xi(t-1)$.

输入和噪声同模型 1,信噪比 S/N 为14.4:1,样本长度 $N = 150$. 取 $N_y = N_u = N_\varepsilon = 4$, $l = 1$, 最后得到模型为

$$\begin{aligned} y(t) = {} & 1.525758y(t-1) - 0.694468y(t-2) + 0.987220u(t-1) \\ & + 0.495915u(t-2) + \varepsilon(t) - 0.798839\varepsilon(t-1) - 0.955750\varepsilon(t-2) \\ & - 0.547184\varepsilon(t-3) - 0.369901\varepsilon(t-4). \end{aligned}$$

模型 3. $\begin{cases} x(t) = 0.5 + 0.5x(t-1) + 0.3u(t-1) + 0.3x(t-1)u(t-1) + 0.5u^2(t-1), \\ y(t) = x(t) + \xi(t), \end{cases}$

输入 $u(t)$ 为(−2,2)上的均匀随机序列，$\xi(t)$ 为(0,1)正态白噪声，信噪比 $S/N = 13.4:1$. 样本长度 $N = 100$. 取 $N_y = N_u = N_\varepsilon = 2$, $l = 2$, 最终模型为

$$\begin{aligned} y(t) = {} & 0.728607 + 0.446242y(t-1) + 0.288457u(t-1) + 0.234727y(t-1)u(t-1) \\ & + 0.463478u^2(t-1) + \varepsilon(t) - 0.152786\varepsilon(t-1) - 0.076700y(t-1)\varepsilon(t-1) \\ & - 0.168290u(t-1)\varepsilon(t-1). \end{aligned}$$

模型 4. $\begin{cases} x(t) = 0.2x(t-1) - 0.3x^2(t-1) + 0.2u(t-1) + 0.6u^2(t-1) \\ y(t) = x(t) + \xi(t) \end{cases}$

输入 $u(t)$ 为(−2,2)上的均匀随机序列，$\xi(t)$ 为(0,0.1)正态白噪声,信噪比 $S/N = 10:1$,样本长度 $N = 100$. 取 $N_y = N_u = N_\varepsilon = 2$, $l = 2$, 得到最终模型为

$$\begin{aligned} y(t) = {} & 0.162629y(t-1) - 0.296162y^2(t-1) + 0.190938u(t-1) \\ & + 0.604594u^2(t-1) + \varepsilon(t) + 0.550824y(t-1)\varepsilon(t-1) \\ & + 0.162009u(t-1)\varepsilon(t-1) - 0.159401y(t-2)\varepsilon(t-2) \\ & + 0.830330\varepsilon^2(t-2). \end{aligned}$$

前三个例子的结果均优于文献 [5]. 从仿真例子也可以看出，本文的算法辨识结果不依赖于选项顺序,而且在信噪比较小情况下,也可得到最优模型.

参 考 文 献

[1] Leontaritis, I. J. and Billings, S. A. Input-output Parametric Models for Non-linear Systems. Part I: Deterministic Non-linear Systems. Part II: Stochastic Non-linear Systems, *Int. J. Control*, **41**(1985), (2) 303—344.

[2] Chen, S. and Billings, S. A Representation of Nonlinear System: The NARMAX Model, *Int. J. Control*, **49** (1989), (3), 1013—1032.

[3] Kortmann, M. and Unbehauen, H., Two Algorithms for Model Structure Determination of Nonlinear Dynamic Systems with Applications to Industrial Processes, Preprints 8th IFAC Symp. on Identification and System Parameter Estimation, Beijing, China, (1988), 939—946.

[4] 王秀峰,劳育红,非线性动态系统模型结构确定和参数估计的新算法,自动化学报,18(1992),(4),385—392.
[5] Korenberg, M., S. A. Billings, Y. P. Liu and P. J. Mcilroy, Orthogonal Parameter Estimation Algorithm for Non-linear Stochastic Systems, *Int. J. Control*, 48(1988), (1), 193—210.

ORTHOGONAL OPTIMIZED-CHOICE ALGORITHM FOR STOCHASTIC NON-LINEAR SYSTEMS IDENTIFICATION

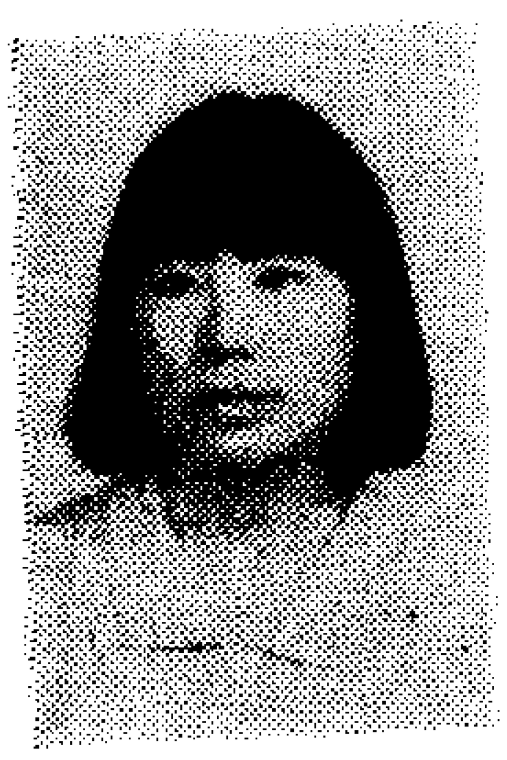
WANG XIUFENG
(Dept. of Computer and System Science, Nankai University)
LI BO
(Dept. of Management ngineering. Tianjin Univversity P. R. China)

ABSTRACT

The structure determination and parameters estimation for stochastic non-linear systems are presented based on the NARMAX model in, this paper. The orthogonal method is applied which allows the parameters estimation independent of each other in the model. It is based on the "Innovation-Contribution" criterion and some other information criteria that we can select the most significant terms into the model: The orthogonal use of the stepwise-regression algorithm mages the model structure independent of the selected term sequence. Several simulation examples are included to demonstrate the affectiveness of the algorithm.

Key words : Non-linear systems; system identification; structure determiniation; parameter estimation; orthogonal method.

王秀峰 1942年2月出生于山东省宁津县. 1960年考入南开大学数学系,1965年毕业并留校任教. 现为南开大学计算机与系统科学系副教授. 现从事控制理论及应用的研究,主要涉及系统辨识、自适应控制、离散事件动态系统、智能控制等领域,发表过30余篇论文.



李波 1967年10月生于山西长治市. 1985年考入南开大学数学系数学专业. 1989年考入本校计算机与系统科学系控制理论与应用专业读硕士. 1992年毕业,现为天津大学管理工程系教师. 主要研究的领域为: 系统辨识、自适应控制、智能管理与控制.