

基于小波和非线性含输入自回归模型的系统辨识算法

石宏理, 蔡远利, 邱祖廉

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 提出了一种结合小波理论和非线性含输入自回归(NARX)模型的系统辨识新算法. 该算法利用小波函数有效的逼近能力避免了应用 NARX 模型系统辨识时确定模型结构的复杂过程, 消除了通常小波网络辨识算法由于输入变量之间可能存在巨大差别而引入的严重失真, 构成了一个通用、有效、不依赖于系统先验信息的非线性辨识框架. 两则数据仿真表明, 对于高度非线性系统, 该算法可使系统估计的均方误差减少 60% 以上.

关键词: 非线性含输入自回归模型; 系统辨识; 小波分析

中图分类号: TP184 **文献标识码:** A **文章编号:** 0253 - 987X(2004)06 - 0611 - 04

System Identification Based on Wavelet and Nonlinear Autoregressive with Exogenous Inputs Model

Shi Hongli, Cai Yuanli, Qiu Zulian

(School of Electronics and Information Engineering, Xi an Jiaotong University, Xi an 710049, China)

Abstract: A new approach to system identification was proposed, which combined wavelet theory and nonlinear autoregressive with exogenous (NARX) model properly. The approach utilized the efficient approximation power of wavelet functions to remove the complicated processes of model structure determination using NARX model in system identification. It avoided potential serious distortion caused by great difference among the input variables in the universal identification algorithm based on wavelet networks and could achieve a more accurate estimation of system. It constructed a universal and efficient framework of nonlinear identification without depending on a priori information. For serious nonlinear systems, two simulation examples showed that the mean of square errors of output estimation caused by the universal wavelet network algorithms could be reduced more than 60% by the proposed approach.

Keywords: nonlinear autoregressive with exogenous model; system identification; wavelet analysis

非线性含输入自回归滑动平均(NARMAX)模型是一种重要的系统辨识工具,它具有辨识精度高、结构简单、收敛速度快等优点. 当模型结构已知时,辨识问题可简化为一个参数估计问题,但其缺点是参数估计精度极大地依赖于模型结构的准确性. 对于非线性系统,当缺乏系统结构的先验信息时,确定准确的 NARMAX 模型是非常困难的. 在实际中, NARMAX 模型只能应用于结构已知或比较简单的辨识问题中^[1,2]. 近年来,小波理论日趋完善,已逐渐成为非常重要的数学工具. 不少学者研究了小波

理论在系统辨识中的应用,提出了很多辨识算法^[3,4]. Coca 等人研究了结合小波理论和 NARMAX 模型的辨识方法^[5],该方法减少了 NARMAX 模型的一些缺点,但仍然没有克服确定模型结构困难和辨识精度严重依赖模型结构准确性的突出缺点,使得辨识过程只能离线完成. 本文充分利用了小波(多维小波)函数对任意局部函数的有效逼近能力和 NARMAX 模型的优点,构造了不依赖于系统结构先验信息同时又可避免确定模型结构过程的辨识算法.

收稿日期: 2003 - 09 - 10. 作者简介: 石宏理(1967~),男,博士生;邱祖廉(联系人),男,教授,博士生导师.

1 NARX 模型

NARMAX 模型是 Leontaritis 等人提出的用于描述非线性离散系统的简洁模型^[1]

$$y(t) = f(y(t-1), \dots, y(t-n_Y), u(t), u(t-1), \dots, u(t-n_U)) + e(t) \tag{1}$$

式中: $y(t-i)$, $i=0,1, \dots, n_Y$ 为系统在 $t-i$ 时刻的输出; $u(t-i)$, $i=0,1, \dots, n_U$ 为系统输入; $e(t)$ 为 0 均值有限方差的白噪声; n_Y, n_U 分别为对应输入和输出的最大延迟; $f(\cdot)$ 为 $(n_Y + n_U + 1)$ 维非线性函数.

由于任何连续函数都可以由其多项式模型以任意精度逼近^[1], 因此将式(1)展开成 k 阶多项式模型时, 有

$$y(t) = 0 + \sum_{m_1=1}^N m_1 x_{m_1}(t) + \sum_{m_1=1}^N \sum_{m_2=1}^N m_1 m_2 x_{m_1}(t) x_{m_2}(t) + \dots + \sum_{m_1=1}^N \dots \sum_{m_k=1}^N m_1 \dots m_k x_{m_1}(t) \dots x_{m_k}(t) + e(t) \tag{2}$$

式中: $N = n_Y + n_U + 1$; $x_k(t) = y(t-k)$, $k=1, \dots, n_Y$; $x_k(t) = u(t+n_Y+1-k)$, $k=n_Y+1, \dots, n_Y+n_U+1$.

可建立式(2)的矩阵形式的线性模型

$$y(t) = P +$$

式中: $P = [1, p_1, \dots, p_{m_1 \dots m_i}, \dots, p_{m_1 \dots m_k}]$, 其中 $p_{m_1 \dots m_i}$ 为回归项, 是若干 x_i 项的乘积; $= [0, 1, \dots, m_1 \dots m_i, \dots, m_1 \dots m_i]^T$ 为待辨识参数; 为模型误差.

由式(2)可知, 未知系统包含非常多的项, 但通常只用少数关键项就足以描述系统的动态特性, 这些项构成了近似的系统模型结构. 当这样的模型结构确定后, 用最小二乘法可以非常有效地估计出参数. 对于未知的非线性系统, 确定模型结构是一个复杂耗时的过程. 通常的方法是逐个计算和比较每一项在描述系统动态特性中的权重, 辨识过程复杂繁琐, 计算量很大, 无法在线完成. 为了实现对未知系统的预测控制, 需要一种不依赖于系统结构先验信息的实时辨识算法, 小波函数渐进最优的逼近能力为解决这一问题提供了新的途径.

2 小波(多维小波)理论

根据小波理论, 任意函数 $f(x) \in L^2(R)$ 可分解为平移和伸缩的小波函数 $\psi(x)$ 的线性和

$$f(x) = \sum_{m,n} c_{m,n} \psi_{m,n}(x) \tag{3}$$

式中: $\psi_{m,n}(x) = 2^{m/2} \psi(2^m x - n)$, $m, n \in Z$ 分别为尺度值和平移量. 平移和伸缩的小波函数可以构成函数基或函数框架, 即存在 $0 < A < B < +\infty$, 对于 $\forall f(x) \in L^2(R)$, 有

$$A \|f\|_2^2 \leq \sum_{m,n} |c_{m,n}|^2 \leq B \|f\|_2^2$$

式中: (\cdot, \cdot) 表示内积. 当 $A = B$ 且 $\psi_{m,n}(x)$ 相互正交时, $\psi_{m,n}(x)$ 构成函数基.

类似于一维小波, 多维小波能够逼近多维函数. 在满足一定的条件下, 平移和伸缩的多维小波函数 $\psi_{l,k}(x)$ 可以构成多维函数框架^[6], 其中 $k \in Z^n$ 为平移量, $l \in Z$ (单尺度伸缩) 或 $l \in Z^n$ (多尺度伸缩) 为尺度值, n 为 x 的维数. 构造多维小波的常用方法有张量乘积法和径向构造法^[7]. 张量乘积法构造的多维小波具有很强的方向性, 逼近精度会高一些, 但运算量随着 x 的维数增加, 在高维情况下较少应用. 径向构造法则与之正好相反.

3 基于多分辨率分解的 NARX 模型

Coca 等人建立了智能的基于小波分析 NARX 模型的辨识方法. 该方法首先构造一个按复杂度递增的多项式模型序列. 模型的确定过程就是在该模型序列中搜索满足要求的最简单模型. 该算法有一些缺点: 建立模型序列需要一定的先验信息, 通常只能根据经验给出中等复杂的模型序列, 并在搜索过程中根据需要随时增加; 算法的运算量极大, 不适合在线辨识. 为了克服该算法的缺点, 构造出具有普适性、不依赖于先验信息、同时又适合在线辨识的模型, 本文充分利用了多维小波序列能够构造多维小波框架的特点, 对系统适当地进行小波分解. 首先将式(1)表示为

$$y(t) = f(y(t-1), \dots, y(t-n_Y), u(t), \dots, u(t-n_U)) + f_Y(y(t-1), \dots, y(t-n_Y)) + f_U(u(t), \dots, u(t-n_U)) + f_{YU}(y(t-1), \dots, y(t-n_Y), u(t), \dots, u(t-n_U)) + e(t) \tag{4}$$

式中: $f(\cdot)$ 是关于 $y(t-i)$ 和 $u(t-i)$ 的线性函数; $f_Y(\cdot), f_U(\cdot)$ 分别是关于 $y(t-i)$ 和 $u(t-i)$ 的非线性函数; $f_{YU}(\cdot)$ 是所有变量的非线性函数. 对式(4)各项分别进行小波(多维小波)分解, 得到变换下的 NARX 模型. 小波的尺度数和平移数取决于模型精度的要求和模型规模之间的折中, 设



$d_i = \{(j, k) \mid 2^{j/2} (2^j x_m - k) \in Z\}$; $j, k \in Z$
 式中: j, k 为单小波的尺度值和平移量; δ 为根据模型精度要求选取的充分小正数,可取为 $\delta/20 \sim \delta/10$,其中 δ 为允许的辨识误差; $x_m = \max(|y(t-1), \dots, y(t-n_Y), y(t-1), \dots, u(t-n_U)|)$. $\{d_i\}$ 构成了一个能够实现对系统充分逼近的尺度值和

平移量集合.对于线性部分,其小波变换的回归量为
 $g_i(t) = g_{j,k}(y(t-1), \dots, y(t-n_Y), u(t), \dots, u(t-n_U)) = [g_{j,k}(y(t-1)), \dots, g_{j,k}(y(t-n_Y)), g_{j,k}(u(t)), \dots, g_{j,k}(u(t-n_U))]^T$ (5)

对于非线性部分(采用径向法构造的多维小波函数),其小波变换的回归量为

$$g_i^y(t) = g_{j,k}(y_n^y(t)); \quad g_i^u(t) = g_{j,k}(y_n^u(t));$$

$$g_i^{yu}(t) = g_{j,k}(y_n^{yu}(t))$$

式中: $y_n^y(t) = [y(t-1), \dots, y(t-n_Y)]^T$,

$$y_n^u(t) = [u(t), \dots, u(t-n_U)]^T$$

$$y_n^{yu}(t) = [y(t-1), \dots, y(t-n_Y), u(t),$$

$$\dots, u(t-n_U)]^T$$

这样,式(4)表示为

$$y(t) = \sum_{i \in I} (l_i g_i^l + \gamma_i g_i^y + u_i g_i^u + \gamma_i^u g_i^{yu})$$
 (6)

式中: I 为 $\{d_i\}$ 的索引集.这是将线性和非线性部分分别进行小波分解的模型.该模型不同于文献[4]中的模型,可避免因输入量和输出量数值之间可能存在巨大差而引入的严重失真.式(6)也可以根据已知的系统结构信息进行简化.

4 仿真算例

例 1 设一个待辨识的非线性系统为

$$y(k) = y(k-1)/10 - x(k)/10 - y(k-1)x(k) + x(k)x(k-1)/3 - y(k-2)x(k-1)/2 - 2y(k-1)y(k-2) + \exp(-u(k-1)/2)/2$$
 (7)

式中:输入信号 $x(k) = 2\sin(0.01k)\exp(-0.001k) - \cos(0.005k)/2$;初始值为 $y(1) = y(2) = -0.5$. $y(k)$ 的观测值 $y_o(k)$ 受到均值为 0、方差为 0.005 的白噪声的干扰,即 $y_o(k) = y(k) + e(k)$. 观察信号如图 1 所示,这是一个包含线性部分和严重非线性部分的系统.

在辨识中,参照式(4)对各项分别进行小波(多维小波)分析.仿真中,选取 $\delta = 0.001$ 构成一个尺度值和

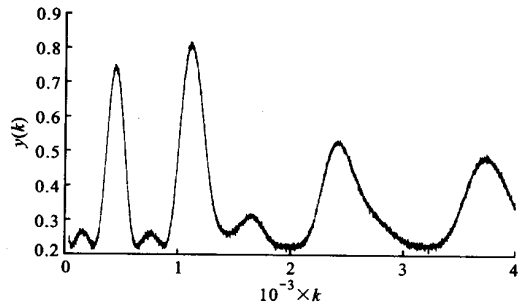


图 1 例 1 中的观测信号

数的傅里叶变换在 $\omega = 0$ 处有 2 阶 0 点,适宜描述函数的快速变化特征.采用径向法构造多维小波函数, $\psi(x) = (n - |x|/2)\exp(-|x|/2)$,其中 n 为 x 的维数.辨识过程中,应用递推最小二乘算法进行参数估计,在 4 000 个数据中选择前 500 个用于系统辨识,即

$$\hat{\theta} = \arg \min_{\theta} \sum_{k=1}^{500} (y(k) - \hat{y}(k))^2$$

其余数据用于评价辨识结果,以 G_{CV} 作为模型评价的判据[7]

$$G_{CV} = \frac{1}{N} \sum_{k=1}^N (\hat{y}(k) - y(k))^2 + \frac{2s^2}{N}$$

式中: \hat{y} 为估计输出; s 是辨识模型中小波函数的个数; N 是采样总数; s^2 为噪声方差.根据辨识结果生成的估计输出信号 $\hat{y}(k)$ 以及 $\hat{y}(k)$ 与实际输出 $y(k)$ 的误差如图 2 所示,其 $G_{CV} = 1.975e-4$.应用文献[4]中的辨识算法时, $G_{CV} = 4.613e-4$.如采用均方误差(MSE)评价,改进算法可以使 MSE 减小 69%.

例 2 设待辨识系统为

$$y(n) = \exp(-u(n-2)u(n-1)u(n)/2)/2$$
 (8)

其中输入信号 $u(n) = \sin(0.05n) + \cos(0.01n)$, $y(n)$ 的观测值 $y_o(n)$ 受到均值为 0、方差为 0.001 的白噪声的干扰.

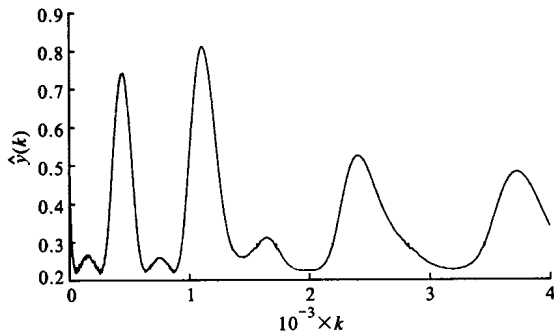
该系统无法应用通常的 NARX 模型进行结构确定,因为虽然 NARX 模型可以包含更一般的回归项[8],例如三角函数项,但在实际中,由于无法确定函数的具体表达式(如三角函数的频率),因而不能建立准确的模型.因此,该系统适宜利用小波变换的 NARX 模型进行辨识.在辨识过程中,参照式(8)对式(6)的模型进行简化,得到小波(多维小波)变换的 NARX 模型

$$y(n) = \sum_{i \in I} (l_i g_i^l(n) + \gamma_i g_i^y(n) + u_i g_i^u(n) + \gamma_i^u g_i^{yu}(n))$$

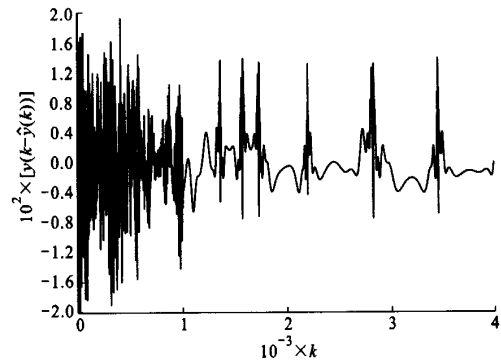
式中: $g_i(n) = [g_{j,k}(u(n), \dots, u(n-n_U))]$ 和 $g_i^{yu}(n) = [g_{j,k}(y(n), \dots, y(n-n_Y), u(n), \dots, u(n-n_U))]$ 分别为对应

于线性部分 $f(\cdot)$ 和非线性部分 $f_U(\cdot)$ 的小波(多维小波)分解. 辨识过程与上例基本类似, 实际输出和根据辨识结果生成的预测信号如图3所示, 其 G_{CV}

$= 0.0265$. 应用文献[4]中的辨识模型时, $G_{CV} = 0.0847$. 采用MSE评价时, 本文的改进算法可以使MSE减小83%.

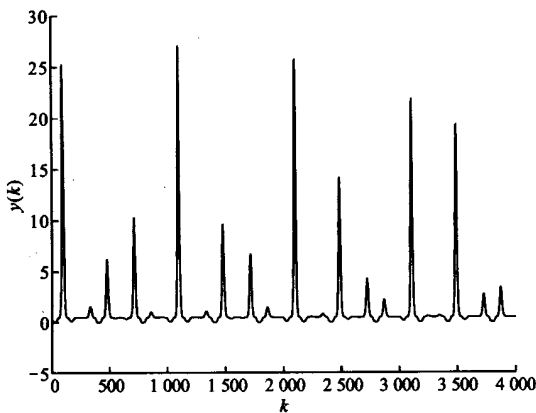


(a) 估计输出信号

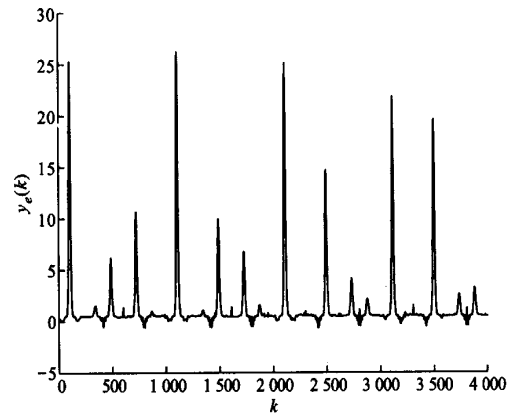


(b) 估计误差信号

图2 例1中的估计信号和估计误差



(a) 观测信号



(b) 估计输出信号

图3 例2中的观测信号和估计信号

5 结论

本文的算法避免了应用NARX模型系统辨识时确定模型结构的复杂过程, 消除了通常小波网络辨识算法可能引入的潜在失真, 构成了一个简洁有效不依赖于系统先验信息的非线性辨识框架. 对于高度非线性系统, 该算法可有效减小估计误差.

参考文献:

- [1] Chen S, Billings S A, Luo W. Orthogonal least squares methods and their application to nonlinear system identification [J]. International Journal of Control, 1989, 50(5): 1 873 ~ 1 896.
- [2] Zhu Q M, Billings S A. Parameter estimation for stochastic nonlinear rational models [J]. International Journal of Control, 1993, 57(2): 309 ~ 333.
- [3] Jun Z, Gilbert G W, Yubo M, et al. Wavelet neural networks for function learning [J]. IEEE Trans Signal

Processing, 1995, 43(6): 1 485 ~ 1 496.

- [4] Zhang Q H. Using wavelet network in nonparametric estimation [J]. IEEE Trans Neural Networks, 1997, 8(2): 227 ~ 236.
- [5] Coca D, Billings S A. Nonlinear system identification using wavelet multiresolution models [J]. International Journal of Control, 2001, 74(18): 1 718 ~ 1 736.
- [6] Tharmarajah K, Zhang Q H. Multidimensional wavelet frames [J]. IEEE Trans Neural Networks, 1995, 6(6): 1 552 ~ 1 556.
- [7] Jonas S, Zhang Q H, Lennart L, et al. Nonlinear black-box modeling system identification: a unified overview [J]. Automatica, 1995, 31(12): 1 691 ~ 1 724.
- [8] Billings S A, Chen S. Extended model set, global data and threshold model identification of severely nonlinear systems [J]. International Journal of Control, 1989, 50(5): 1 897 ~ 1 923.

(编辑 刘 杨)