

文章编号: 1001-0920(2003) 03-0378-04

基于正交最小二乘法的小波网络在系统辨识中的应用

李德强, 黄莎白

(中国科学院 沈阳自动化研究所, 辽宁 沈阳 110016)

摘要: 利用单尺度小波框架理论建立初始小波网络。首先, 提出基于ROLS思想的小波网络正交化算法, 实现参数的在线辨识。其次, 提出一种基于OLS思想的小波网络结构辨识算法, 删除那些对网络影响小的小波基, 来解决高维辨识问题。最后, 列举事例证明了所述方法的有效性。

关键词: 小波网络; OLS; ROLS; AIC 准则

中图分类号: TP273 **文献标识码:** A

Orthogonal least squares based wavelet networks in system identification

LI De-qiang, HUANG Sha-bai

(Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China)

Abstract: Single scaling wavelet frame theory is used to build up an initial wavelet network. ROLS method is proposed to orthogonalize the wavelets to realize parametric estimation on-line. Then OLS based structure identification method is given to handle multi-dimensional problems by eliminating those wavelets with least contributions. The results of experience demonstrate the effectiveness of the proposed methods.

Key words: Wavelet networks; Orthogonal least squares; Recursive orthogonal least squares; AIC criterion

1 引言

小波既能逼近低频信息, 又能识别高频信号的多分辨率特性^[1], 使小波网络在系统辨识领域中得到一定应用, 但以前的应用往往局限于低维系统, 在处理高维非线性系统时, 往往出现数据爆炸现象。神经网络虽在处理高维系统上表现出很强的能力, 但在结构辨识和参数辨识方面不是很完善。将神经网络和小波网络有机地结合起来, 取长补短, 既解决了高维问题又能有效地进行结构辨识和参数辨识, 因此对它的研究已引起人们的广泛重视, 并取得了一定成果。S CHEN⁺ 等用 OLS 算法辨识 NARMAX 模型及 RBF 网^[2,3]; Zhang Qinghua 利用 Gram-

Schmidt 算法对小波网络进行正交化^[4]; Gomm JB 用 ROLS 算法选择 RBF 网的中心点^[5]。本文在上述基础上, 提出基于 OLS 思想的小波网络的正交化及结构辨识方法, 较成功地解决了高维小波网络的辨识问题。

2 小波网络的建立

Kugarajah 等提出, 在单尺度下, 直接构造 $L_2(R^d)$ 下的正交小波框架需要 1 个母尺度函数和 $2^d - 1$ 个母小波函数, 若构造 $L_2(R^d)$ 小波框架, 则只需一个母小波即可, 且所构造的小波框架受维数 d 的影响较弱^[6]。

设 $f: R^d \rightarrow R$ 满足下面模型

收稿日期: 2002-02-05; 修回日期: 2002-05-14。

作者简介: 李德强(1971—), 男, 辽宁盘锦人, 博士生, 从事神经网络、系统辨识等研究; 黄莎白(1939—), 女, 研究员, 博士生导师, 从事模式识别与图像处理、系统辨识等研究。

$$y = f(x) + e \quad (1)$$

其中: f 为 $L_2(R^d)$ 上的非线性连续函数; e 为与 x 不相关的白噪声; $\{x_i, y_i\}, i = 1, 2, \dots, N$ 为样本数据。 f 可用单尺度离散小波基的线性组合表示, 即

$$f(x) = \sum_{j, z, k} \theta_{j, z, k} \alpha^{j/d} \psi(\alpha^j x - k\beta) \quad (2)$$

其中: 设 $\psi(x)$ 为母小波函数, α 和 β 为尺度及平移的步长 (一般情况下 $\alpha = 2, \beta = 1$)。

Zhang Qinghua 提出基于样本非均匀分布的情况下, 初始小波网络的建立方法^[4]。

定义 1 $r_j = \max\{x : |\psi_{j,0}(x)| > \epsilon\}$ 为紧支撑半径, ϵ 为小正实数。在尺度 j 下, 小波的支撑半径为 $r_0/2^j$, 栅格的间距为 $1/2^j$ 。构造支撑 x_i 的尺度和平移索引集合

$$I_i = \{(j, k) : x_i^* - k\beta \leq r_j/\alpha^j = r_0\} \quad (3)$$

其中: $x_i^* = \lfloor x_i/\alpha^j \rfloor$ 为尺度索引 j 下, 距 x_i 最近的栅格点。在 I_i 中, 只有支撑 x_i 的小波基被选中, 而对 x_i 来说, 空的小波基则被淘汰。对所有的 I_i 取并集, 所构造的有限小波框架将支撑整个输入空间, 即

$$\Psi = \{\psi_{j,k}(j, k) \quad I_1 \quad I_2 \quad \dots \quad I_N\}$$

用单索引重新表达 $\Psi = [\psi_1, \psi_2, \dots, \psi_L]$ 。至此, 便完成了初始小波网络的构造。

3 正交最小二乘法 OLS

式(1)可表示为

$$Y = \Psi\theta + E \quad (4)$$

$$J(\theta) = \|E\|_2^2 = \|Y - \Psi\theta\|_2^2 \quad (5)$$

其中

$$Y = [y(1), \dots, y(N)]^T$$

$$E = [e(1), \dots, e(N)]^T$$

$$\Psi = [\psi(1), \dots, \psi(N)]^T = [\psi_1, \dots, \psi_L]$$

$$\theta = [\theta_1, \dots, \theta_L]^T$$

3.1 BOLS 算法(一次性正交最小二乘法)

对于式(4), 输出空间是由 ψ_i 张成的, Ψ 中的部分小波基为冗余, 只有一部分是重要的, 也就是说 Ψ 是相关的, 造成 Ψ 是病态, 不能直接用 LS 算法求解, 一般的方法是用正交化算法计算 θ 。

对小波向量矩阵 $\Psi \in R^{N \times L}$ 进行 QR 分解

$$\begin{cases} [\Psi | Y] \xrightarrow{Q^T} \begin{bmatrix} R & | & Y \\ 0 & | & \tilde{Y} \end{bmatrix} \dots L \text{ 行} \\ \dots N - L \text{ 行} \\ Q^T \Psi = \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad Q^T Y = \begin{bmatrix} Y \\ \tilde{Y} \end{bmatrix} \end{cases} \quad (6)$$

其中: $Q \in R^{N \times N}$ 为正交矩阵 $R = [r_1, \dots, r_L] \in R^{N \times L}$

为上三角矩阵。

正交变换对二范数具有保角映射特性, 所以

$$J(\theta) = \|Q^T Y - Q^T \Psi \theta\|_2^2 = \left\| \begin{bmatrix} Y \\ \tilde{Y} \end{bmatrix} - \begin{bmatrix} R \\ 0 \end{bmatrix} \theta \right\|_2^2 = \left\| \begin{bmatrix} Y - R\theta \\ \tilde{Y} \end{bmatrix} \right\|_2^2 \quad (7)$$

令

$$Y = R\theta \quad (8)$$

则

$$J(\theta) = \|\tilde{Y}\|_2^2 \quad (9)$$

是误差的平方和。

对于式(8), 利用 R 为上三角矩阵这一特性, 易求出 θ 。

具体的 QR 分解算法较多, 文献[2] 利用 CGS 和 MGS 算法对 NARMAX 模型进行辨识。

3.2 ROLS 算法(递推正交最小二乘法)

OLS 算法虽然避免了求逆矩阵, 但当 N 很大时, QR 分解计算量非常大, 难于实现在线计算。为此, 本文提出 ROLS 算法来实现 QR 分解。

$$\begin{bmatrix} R(k-1) & | & Y(k-1) \\ \Psi(k) & | & y(k) \end{bmatrix} \xrightarrow{Q^{T(k)}} \begin{bmatrix} R(k) & | & Y(k) \\ 0 & | & \tilde{y}(k) \end{bmatrix} \dots L \text{ 行} \\ \dots 1 \text{ 行} \\ Q^T(k) \begin{bmatrix} R(k-1) \\ \Psi(k) \end{bmatrix} = \begin{bmatrix} R(k) \\ 0 \end{bmatrix}$$

$$Q^T(k) \begin{bmatrix} Y(k-1) \\ y(k) \end{bmatrix} = \begin{bmatrix} Y(k) \\ \tilde{y}(k) \end{bmatrix} \quad (10)$$

当 $Y(k) = R(k)\theta(k)$ 时, 有

$$J(k) = \|\tilde{Y}(k)\|_2^2 = \|\tilde{Y}(k-1) + \tilde{y}(k)\|_2^2 \quad (11)$$

ROLS 算法的实现描述如下:

- 1) 初始化 $k = 0, R(0) = \alpha I$, 其中 α 为一任意小正实数(如 0.01), $Y(0) = 0, \tilde{Y}(0) = 0$;
- 2) 在 $k = 1$ 时, 利用式(10)进行 QR 分解。依次用 $R(k-1)$ 的主对角线元素与最底行对应元素作 Givens 变换, 直到

$$\begin{bmatrix} R(k-1) \\ \Psi(k) \end{bmatrix} \rightarrow \begin{bmatrix} R(k) \\ 0 \end{bmatrix}$$

得到 $R(k), Q^T(k), Y(k)$ 及 $\tilde{y}(k)$, 用式(11)计算 $J(k)$;

- 3) $k = k + 1$ 。若 $k < N$ 时, 转 2); 否则 $Y(N) = R(N)\theta(N)$, 结束。

ROLS 算法的性质概括如下:

- 1) 解决矩阵 Ψ 为病态的问题;
- 2) θ 为唯一确定且误差平方和最小;
- 3) 中间步骤不需计算 $\theta(k)$, 只在最后一步计算 $\theta(N)$;
- 4) 每增加一组新数据, 作 L 次 Givens 变换, 计算量小, 可实现参数的在线辨识。

4 小波网络的结构辨识

本文提出一种基于 OLS 思想的小波网络结构辨识算法, 利用由复杂到简单的方法, 顺序从正交小波网络中删除对网络输出影响最小的小波基, 同时保证网络精度。该结构辨识算法是在上三角矩阵 R 上进行的。假设删除 Ψ 的第 j 列 ψ_j , 对应删除 R 的第 j 列 r_j , j 应满足

$$j = \arg \min \{ r_i \theta^2, i = 1, 2, \dots, L \} \quad (12)$$

删除后 R 和 θ 分别变为

$$R = [r_1, \dots, r_{j-1}, r_{j+1}, \dots, r_L]$$

$$\theta = [\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta]^T$$

对于 Y 和 R , θ 已不能保证 $Y - R\theta$ 最小, 需要重新对 R 正交化

$$[R | Y] \xrightarrow{Q^T} \begin{bmatrix} R_j & Y_j & \dots & L-1 \text{行} \\ 0 & \tilde{y}_j & \dots & 1 \text{行} \end{bmatrix}$$

$$R = Q \begin{bmatrix} R_j \\ 0 \end{bmatrix}, \quad \begin{bmatrix} Y_j \\ \tilde{y}_j \end{bmatrix} = Q^T Y \quad (13)$$

令 $Y_j = R_j \Theta$, 有

$$J_j = \tilde{Y}_j^2 = \tilde{y}_j^2 + \tilde{Y}^2 \quad (14)$$

可见删除 ψ_j 后, 结构虽简单了, 但误差的平方和增加了 \tilde{y}_j^2 。引入 AIC 准则, 取小波网络的结构复杂性和系统精度的折衷点, 当 AIC 函数最小时, 所对应的小波网络作为最合理的。令

$$AIC(\phi) = N \log C(\hat{\theta}_s) + N_s \phi \quad (15)$$

其中: N 为数据长度, $C(\hat{\theta}_s) = \sum Y_j^2 / N$ 代表均方差, N_s 为当前小波基的个数, $\phi = 4$ 。

Leontaritis 等提出式 (15), ϕ 是 chi-squared 分布的一个重要参数, 当 $\phi = 4$ 且 AIC 取最小值时, 模型的置信度为 95.44%^[7]。

小波网络结构辨识算法描述如下:

- 1) $k = 0$, 利用 ROLS 算法初始化 $Y(0), R(0), \theta(0)$, 由式 (15) 计算 $AIC(0)$;
- 2) 选定对输出能量贡献最小的小波基, 并删除

利用式 (13), 从 R 的第 j 列开始到最后一列, 依次用主对角线元素与其下面元素作 Givens 变换, 来计算 Y_{kj} 和 R_{kj} , 由式 (14) 计算 J_{kj} 。令

$$Y(k+1) = Y_{kj}, \quad R(k+1) = R_{kj}$$

$$J(k+1) = J_{kj}$$

由式 (15) 计算 $AIC(k+1)$;

- 3) 若 $AIC(k+1) < AIC(k)$, 计算 $\theta(k+1)$, $k = k+1$, 转到 2), 继续; 否则令 $J^* = J(k), Y^* = Y(k), R^* = R(k)$, 根据 $Y^* = R^*$ 计算 θ^* , 结束。

该算法有以下两大优点:

- 1) 在步骤 2) 中, 只作 $L - k - j$ 次 Givens 变换就能实现正交化, 计算量小;
- 2) 算法中每删除一个小波基都是在上三角矩阵 $R(k)$ 上进行的, 而不需要用 ROLS 算法重新训练网络。

5 试验结果

母小波函数采用墨西哥草帽

$$\psi(x) = (d - x^2)e^{-x^2/2}$$

定义评价指标

$$\delta = \frac{\sum_{i=1}^N |\hat{y}(x_i) - y_i|^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

其中 $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \hat{y}(x_i)$ 为小波网络输出。

例 1 $y = 10\sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5, x = \{x_1, x_2, x_3, x_4, x_5\} = [-10, 10]^5$, 在定义域内均匀选取 10^5 个点作为训练样本, 初始小波网络含 1226 个小波基, 结构辨识后的小波网络由 78 个小波基构成, $\delta = 0.003518$ 。对简化后的小波网络, 再在定义域内取 10^5 个点作为测试样本, $\delta = 0.01147$ 。

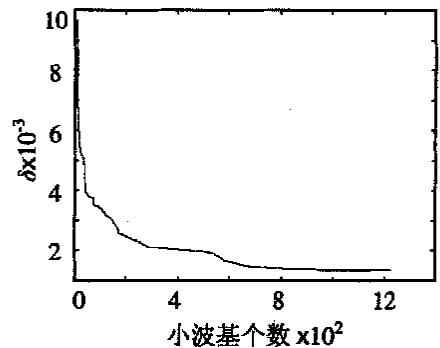


图 1 训练样本的 δ 曲线

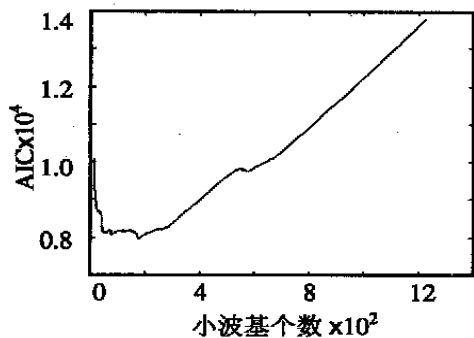


图 2 AIC 曲线

6 结 论

本文利用单尺度小波框架理论,建立了候选小波集合。在此基础上,运用 ROLS 算法进行 QR 分解,实现小波基的正交化,进而能够计算每个小波基对网络输出的贡献,可实现参数的在线辨识。利用从复杂到简单的思想,删除对输出影响小的小波基,引入 AIC 准则,在网络的结构复杂性和精度之间找到一平衡点,实现小波网络的结构辨识。实验结果表明,本文方法计算量小且不用因网络结构的改变而重复训练网络,对高维非线性系统的辨识具有一定的指导意义。

(上接第 377 页)

5 结 论

本文提出了多模型预测函数控制策略,并给出了模型平滑切换方法,有效地解决了对象特性变化过大时,预测函数控制品质变差的问题。火电厂主汽温对象是典型的时变对象,与采用单一模型预测函数控制的主汽温系统相比,采用多模型预测函数控制策略的系统动态品质得到了明显改善。

参考文献(References):

- [1] Kuntze H B, Jacubasch A, Richalet J. On the predictive functional control of an elastic industrial robot[A]. *Proc of 25th CDC*[C]. Athens, 1986. 1877-1881.
- [2] Richalet J, Abu el Ata-Doss S, Arber D, et al. Predictive functional control: Application to fast and accurate robots[A]. *Proc of IFAC 10th World Congress*[C].

参考文献(References):

- [1] Zhang Qinghua, Albert Benveniste. Wavelet networks [J]. *IEEE Trans on Neural Networks*, 1992, 3(6): 889-898.
- [2] Chen S, Billings S A, Luo W. Orthogonal least squares methods and their application to non-linear system identification [J]. *Int J Control*, 1989, 50(5): 1873-1896.
- [3] Chen S, Billings S A. Neural networks for nonlinear dynamic system modeling and identification [J]. *Int J Control*, 1992, 56(2): 319-346.
- [4] Zhang Qinghua. Using wavelet network in nonparametric estimation [J]. *IEEE Trans on Neural Networks*, 1997, 8(2): 227-236.
- [5] Gomm J B. Selecting radial basis function network centers with recursive orthogonal least squares training [J]. *IEEE Trans on Neural Networks*, 2000, 11(2): 306-314.
- [6] Kugarajah K, Zhang Q. Multidimensional wavelets frame [J]. *IEEE Trans on Neural Networks*, 1995, 6(6): 1552-1556.
- [7] Leontaritis I J, Billings S A. Model selection and validation methods for nonlinear systems [J]. *Int J Control*, 1987, 45(1): 311-341.

Munich, 1987. 251-258.

- [3] 范永胜,徐治皋,陈来九. 基于动态机理分析的锅炉过热汽温自适应模糊控制系统研究[J]. *中国电机工程学报*, 1997, 17(1): 23-28.
(Fan Yongsheng, Xu Zhigao, Chen Laijiu. Study of adaptive fuzzy control of boiler superheated steam temperature of based on dynamic mechanism analysis [J]. *Proc of the CSEE*, 1997, 17(1): 23-28.)
- [4] 张泉灵,王树青. 化学反应器温度跟踪预测函数控制的研究及应用[J]. *控制理论与应用*, 2001, 18(4): 559-563.
(Zhang Quanling, Wang Shuqing. Studies and applications of predictive functional control in temperature tracking of chemical reactors [J]. *Control Theory and Applications*, 2001, 18(4): 559-563.)