

文章编号: 1001-0920(2005)03-0353-03

高维小波网络及其在丙烯腈收率预测中的应用

赵恒平, 俞金寿

(华东理工大学 自动化研究所, 上海 200237)

摘要: 针对基于单尺度小波框架的高维小波网络, 提出一种系统化的设计方法。首先在小波框架内提出一种小波基初始化方法; 然后根据样本的分布特点, 提出一种改进的小波基粗选方法; 最后将自适应投影算法与AIC准则相结合, 对小波基进行精选, 同时完成网络参数的辨识。将该方法应用于丙烯腈收率的预测, 研究结果表明了该方法的有效性。

关键词: 小波框架; 小波网络; 自适应投影算法; AIC准则

中图分类号: TP273 **文献标识码:** A

Large-dimensional wavelet networks and its application in predicting yield of acrylonitrile

ZHAO Heng-ping, YU Jin-shou

(Research Institute of Automation, East China University of Science and Technology, Shanghai 200237, China)

Correspondent: ZHAO Heng-ping, E-mail: zhaohp@163.com

Abstract: To large-dimensional wavelet networks based on single scaling wavelet frames theory, a systematic design method is proposed. First, method for initialization of wavelet basis in wavelet frames is proposed; second, according to the data distribution, a modified method for rough selection of wavelet basis is given; and the last, an adaptive projection algorithm combined with AIC criterion is used to purify the wavelet basis, meanwhile finishing the parameters identification. The proposed method is applied in predicting the yield of acrylonitrile, and the results show the effectiveness of this method.

Key words: wavelet frames; wavelet networks; adaptive projection algorithm; AIC criterion

1 引言

小波网络是近年来神经网络研究中的一个新的分支,它是在小波分析理论的基础上发展起来的^[1~4,6,7]。其优势在于具有多分辨率,这一特性使得小波网络成为复杂非线性系统建模的一个有力工具。在已有的小波网络中,基于框架的小波网络^[3]具有平移和尺度参数不变,只有输出层权值可调,不存在局部极小和学习速度快的特点,因而得到了广泛的应用。在基于框架的小波网络中,小波基的初始化是研究的一个重点,较常见的方法有基于时域的方法^[1]和基于时频集中区域^[2]的方法。前者的缺点是其尺度参数的确定是人为的;后者虽然克服了这一

缺点,但它不适于非均匀分布的不完整样本的建模。二者共同的缺点是在高维情况下网络结构存在“维数爆炸”问题^[1],目前还没有一个令人满意的系统化设计方法。

本文针对基于单尺度小波框架^[3]的高维小波网络,提出了一种系统化的设计方法,并将该方法用于丙烯腈收率的预测。研究结果表明了该方法的有效性。

2 小波网络

单尺度多维小波函数 $\Psi(x) \in L^2(R^q)$ 定义为

$$\begin{cases} \Psi_{m,n}(x) = \varphi_{m,n_1}(x_1)\varphi_{m,n_2}(x_2)\dots\varphi_{m,n_q}(x_q), \\ x = [x_1, x_2, \dots, x_q], m \in Z, n \in Z^q. \end{cases} \quad (1)$$

收稿日期: 2004-04-14; 修回日期: 2004-07-22

作者简介: 赵恒平(1979—),男,山东滕州人,博士生,从事过程智能建模、优化控制的研究;俞金寿(1939—),男,浙江海宁人,教授,博士生导师,从事过程模型化及优化控制、先进控制等研究。

其中: $\varphi_{m,n_i}(x_i)$ 是由母小波函数 $\mathcal{Q}(x) \in L^2(R)$ 按

$$\varphi_{m,n_i}(x) = 2^{m/2} \mathcal{Q}(2^m x - n_i), m, n_i \in Z \quad (2)$$

进行尺度与平移变换产生的; m 和 n_i 分别为尺度参数和平移参数; n 为多维小波函数的平移参数, 且满足

$$n = [n_1, n_2, \dots, n_q] \quad (3)$$

如果在一定条件下 $\{\Psi_{m,n}\}$ 构成了 $L^2(R)$ 中的框架, 那么任意有限支集函数 $f \in L^2(R)$ 可表示为

$$f(x) = \sum_{m,n} \omega_{m,n} \Psi_{m,n}(x). \quad (4)$$

其中: N 代表所选小波基的个数, $\omega_{m,n}$ 为权值

如果将 $\Psi_{m,n}$ 看作隐层传递函数, $\omega_{m,n}$ 看作隐层与输出层的连接权值, 则式(4)可看作一个 3 层神经网络的函数表达式, 称为小波网络

3 小波网络的初始化

不失一般性, 本文只考虑多输入单输出系统. 设系统有 q 个输入变量, 采样数据为 $\{(x_i, y_i)\}_{i=1}^L$ (L 为样本数目), 用矩阵形式表示为

$$\begin{cases} X_{L \times q} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{L1} & x_{L2} & \dots & x_{Lq} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_L \end{bmatrix} = \\ [xx_1, xx_2, \dots, xx_L]^T, \\ Y_{L \times 1} = [y_1, y_2, \dots, y_L]^T. \end{cases} \quad (5)$$

其中: x_{ij} 表示第 j 个变量在第 i 个样本上的观测值; $x_i \in R^q$ 表示第 i 个样本的观测向量; $xx_j \in R^L$ 表示第 j 个变量的观测向量; $y_i \in R$ 表示第 i 个样本输出观测量 ($i = 1, 2, \dots, L; j = 1, 2, \dots, q$).

假设输入变量具有相同的取值范围 $[a, b]$. 本文在文献[4]的基础上, 提出一种系统化的基于单尺度小波框架的小波网络初始化方法, 具体如下:

3.1 尺度参数的初始化

尺度参数 m 可由下式确定:

$$\alpha(b - a) 2^{-m} \leq \beta(b - a), \quad (6)$$

式中 $0 < \alpha < \beta < 1$. 由于 m 必须为整数, 则 m 的集合为

$$\left\{ \left[-\frac{\ln(\beta(b-a))}{\ln 2} \right] + 1, \left[\frac{\ln(\beta(b-a))}{\ln 2} \right] + 2, \dots, \left[-\frac{\ln(\alpha(b-a))}{\ln 2} \right] \right\}. \quad (7)$$

3.2 平移参数的初始化

从式(1)可以看出, 如果认为单尺度多维小波函数的支承域为空间一超立方体^[1], 那么可以先对一维小波函数的平移参数进行初始化, 然后按式(3)得出多维小波函数的平移参数值

对于式(7)所确定的集合中每一元素 m , 一维

小波函数的平移参数 n 由下式确定:

$$a \leq 2^{-m} n \leq b \quad (8)$$

因为 n 必须为整数, 所以 n 的集合为

$$\{[2^m a] + 1, [2^m a] + 2, \dots, [2^m b]\}. \quad (9)$$

此时得到的候选小波基相当于在小波框架内进行正规的截取

4 小波网络的结构辨识

高维情况下, 在小波框架内进行正规截取所得到的候选小波基数量非常大, 不利于小波网络的结构辨识. 对此, 本文提出一种新的小波网络的结构辨识方法: 首先根据样本的分布特点对初始化得到的小波基进行粗选; 然后将自适应投影算法^[5]与 AIC 准则相结合, 对粗选得到的小波基进行精选, 同时完成网络参数的辨识

4.1 小波基的粗选

采样数据 $\{x_i\}_{i=1}^L$ 在输入空间 R^q 的一个有限域 D 内以及 $\Psi_{m,n}$ 的指数具有衰减性, 故当尺度参数为 m 时, 小波基的粗选方法如下:

Step 1: 确定第 j 个输入变量对应小波基的平移参数, 使其满足

$$N_j = \{n_j \mid m: \varphi_{m,n_j}(xx_j)^T \varphi_{m,n_j}(xx_j) > \lambda n\}, \quad j = 1, 2, \dots, q \quad (10)$$

其中: N_j 是当尺度参数值为 m 时所对应平移参数的集合; 训练样本 $xx_j \in R^L$; λ 为指定的正实数, 通过调节 λ 可以控制 N_j 的大小

Step 2: 通过下式遍历所有可能的组合来确定多维小波函数的平移参数:

$$W = \left\{ \Psi_{m,n_1,n_2,\dots,n_q}; \prod_{j=1}^q (\varphi_{m,n_j}(xx_j))^T \times \varphi_{m,n_j}(xx_j) \in n_j \in N_j \right\}, \quad (11)$$

式中 ϵ 为一指定正数

Step 3: 采用聚类方法对 Step 2 中所得的平移参数进行聚类, 将聚类中心作为尺度 m 所对应的平移参数集合. 当平移参数 n 之间的距离小于某一给定临界值时, 则认为它们从样本中提取的信息十分相似^[6].

式(11)利用 $\prod_{j=1}^q (\varphi_{m,n_j}(xx_j))^T \varphi_{m,n_j}(xx_j)$ 的值越大, $\Psi_{m,n_1,n_2,\dots,n_q}$ 包含的数据信息越多^[7]的特点, 对小波基进行粗选, 将粗选得到的小波基记为 $\{\Psi_1, \Psi_2, \dots, \Psi_N\}$. 令

$$\Psi_i = [\Psi_i(x_1), \Psi_i(x_2), \dots, \Psi_i(x_L)], \quad i = 1, 2, \dots, N. \quad (12)$$

于是剩下问题就转化为在满足精度的前提下, 从 $\{\Psi_1, \Psi_2, \dots, \Psi_N\}$ 中选取最佳子集来逼近 Y , 即小波



基的精选

4.2 小波基的精选

对小波基进行精选主要基于以下两个原因: 1) 粗选得到的小波基不仅仍为冗余基, 而且数量巨大; 2) 在满足精度的前提下, 选用更多的小波基往往导致小波网络出现过拟合及泛化能力差的现象

自适应投影算法的基本思想是按最大投影匹配的原则寻找一组“最佳基”, 达到信号自适应分解的目的^[5]。自适应算法的缺点是需要预先指定一个精度值, 如果太低, 则逼近效果不理想; 如果太高, 则将导致系统过于复杂。为了在模型所达到的精度和复杂性之间进行合理的折衷, 本文将 AIC 准则引入自适应投影算法, 作为自适应投影算法的终止条件。

令

$$AIC(M) = N \ln E + 4M, \quad (13)$$

其中: $E = \frac{1}{L} \sum_{k=1}^L [y_k - \hat{y}_k]^2$, y_k 为小波网络输出估计值, M 为当前小波基的个数。小波基的精选算法描述如下:

Step 1: 对 $\{\Psi_i\}_{i=1}^N$ 进行能量归一化, 即

$$h_i = \Psi_i / \|\Psi_i\|, \quad i = 1, 2, \dots, N;$$

Step 2: 选择 $h_{i_0} \in \{h_i\}_{i=1}^N$, 使

$$\|Y, h_{i_0}\| = \max_{i \in \{1, \dots, N\}} \|Y, h_i\|,$$

则 Y 可分解为

$$Y = \langle Y, h_{i_0} \rangle h_{i_0} + R_Y^1,$$

令 $k = 1$, 由式(13) 计算 $AIC(0)$;

Step 3: 选择 $h_{i_k} \in \{h_i\}_{i=1}^N$, 使

$$\|R_Y^k, h_{i_k}\| = \max_{i \in \{1, \dots, N\}} \|R_Y^k, h_i\|,$$

则 R_Y^k 可分解为

$$R_Y^k = \langle R_Y^k, h_{i_k} \rangle h_{i_k} + R_Y^{k+1},$$

由式(13) 计算 $AIC(k)$;

Step 4: 如果 $AIC(k) > AIC(k-1)$, 则结束; 否则, 令 $k = k + 1$, 转 Step 3

若迭代选取的基向量为 $h_{i_0}, h_{i_1}, \dots, h_{i_p}$, 则小波函数隐含层节点的神经元函数为 $\Psi_{i_0}, \Psi_{i_1}, \dots, \Psi_{i_p}$, 网络的输出权值分别为 $\langle R_Y^0, h_{i_0} \rangle / \|h_{i_0}\|, \langle R_Y^1, h_{i_1} \rangle / \|h_{i_1}\|, \dots, \langle R_Y^p, h_{i_p} \rangle / \|h_{i_p}\|$ 。

5 丙烯腈收率预测研究

某化工厂的年产 5 万吨丙烯腈装置, 采用美国标准石油公司 Sohio 的生产工艺, 以 C-41 作催化剂, 用丙烯、氨、空气为原料, 在沸腾床反应器中一次直接氧化制取丙烯腈 (即丙烯氨氧化法)。在丙烯腈装置的生产过程中, 丙烯腈的收率是一个关键指标, 而

及时准确地测量产品中丙烯腈的收率是进行丙烯腈装置先进控制的关键。采用在线分析仪表检测丙烯腈收率, 不仅价格昂贵, 维护保养复杂, 而且有较大的测量滞后, 难以满足生产要求。采用软测量方法对丙烯腈收率进行预测具有重要的意义。

经过对工艺机理进行全面分析及与现场工人的讨论, 发现丙烯腈收率主要受 7 个变量的影响: 反应压力、中段温度、纯丙烯量、空比、氨比、反应线速、触媒量。将现场采集到的 B 反应器数据经过异常数据剔除、平滑处理, 共得到 260 组; 将所有数据归一化到 $[-1, 1]$ 之间, 选择其中的 200 组作为训练数据, 60 组作为测试数据。采用本文提出的方法建立丙烯腈收率软测量模型, 取母小波函数 $\varphi(x)$ 为墨西哥草帽小波, 模型参数 $\alpha = 0.02, \beta = 1, \lambda = 65, \epsilon = 0.5$; 粗选过程中采用最近邻聚类算法^[8], 类与类之间的距离 $r = 1$ 。经过粗选得到 3 341 个小波基, 精选得到 39 个小波基, 模型的泛化曲线如图 1 所示。

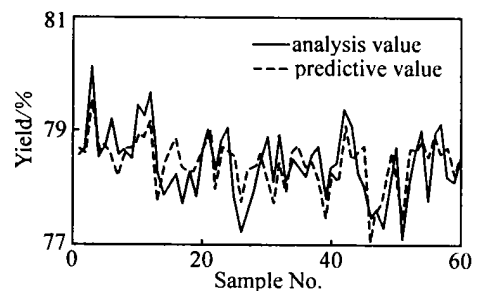


图 1 小波网络模型泛化结果

表 1 为模型仿真结果的对比。对比结果表明, 小波网络在精度上较 PCA-RBF 神经网络有较大的提高, 并具有较强的泛化能力。

表 1 模型仿真结果对比

模 型	训练误差		泛化误差	
	MSE	MAXE	MSE	MAXE
小波网络	0.096 2	1.142 8	0.167 8	0.769 4
PCA-RBF 网络	0.231 3	1.422 3	0.326 7	1.125 1

6 结 语

本文提出一种基于单尺度小波框架的高维小波网络设计方法, 所提出的小波基初始化方法与小波基粗选方法, 在尽可能保留训练数据信息的基础上, 可有效地避免维数爆炸问题; 将 AIC 准则引入自适应投影算法, 对小波基进行精选, 可在模型所达到的精度和复杂性之间进行合理的折衷。将该方法应用于丙烯腈收率的预测, 研究结果表明, 采用本文方法构造的小波网络具有较强的泛化能力, 适于高维非线性系统的建模。

(下转第 360 页)

表3 稳态误差的特征参数

实验次数	补偿算法	均值 $\mu_{\text{ess}}/\mu\text{m}$	标准差 $\sigma_{\text{ess}}/\text{mm}$	峰峰 值/mm
1	整体	593.4	3.9	28.4
	分部	-45.491	3.2	16.2
2	整体	-1.500	3.9	23.4
	分部	381.21	3.1	20.6
3	整体	491.8	4.2	31.3
	分部	18.218	3.6	17.0
4	整体	202.19	5.0	42.6
	分部	112.31	3.5	18.5
5	整体	-126.11	4.0	26.5
	分部	101.27	3.0	17.8

5 结 语

本文提出的摩擦力分部补偿算法对摩擦力的不同分量分别进行补偿,克服了采用FNN对摩擦力整体补偿方法中,因摩擦力非光滑特性引起的较大逼近误差和其他不确定性带来的较大稳态误差,从而使系统具有更高的精密度。

参考文献(References)

- [1] Avinash Taware, Gang Tao, Nilesh Pradhan, et al. Friction compensation for a sandwich dynamic system [J]. *Automatica*, 2003, 39(3): 481-488
- [2] Tan K K, Lee T H, Huang S N, et al. Friction modeling and adaptive compensation using a relay feedback approach [J]. *IEEE Trans on Industrial Electronics*, 2001, 48(1): 169-170
- [3] 郑言海, 庄显义, 强盛, 等. 考虑摩擦力影响精密伺服系统的鲁棒自适应控制[J]. *自动化学报*, 2002, 28(3): 445-449

(Zheng Y H, Zhuang X Y, Qiang S, et al. Robust adaptive control for high-precision servo-system in the presence of friction effects[J]. *Acta Automatica Sinica*, 2002, 28(3): 445-449)

- [4] Huang S N, Tan K K, Lee T H. Adaptive motion control using neural network approximations [J]. *Automatica*, 2002, 38(2): 227-233
- [5] 张娟, 李秀娟. 伺服系统的神经网络摩擦力自适应补偿研究[J]. *计算技术与自动化*, 2002, 21(4): 11-15
(Zhang M, Li X J. Study of the friction compensation in servo systems [J]. *Computing Technology and Automation*, 2002, 21(4): 11-15)
- [6] 蒋志明, 吴伟, 林廷圻. 高斯基函数CMAC神经网络用于克服摩擦力非线性的研究[J]. *机床与液压*, 2000, 28(3): 24-25
(Jiang Z M, Wu W, Lin T Q. Application of CMAC neural network with Gaussian basis functions to overcoming nonlinear friction [J]. *Machine Tool and Hydraulics*, 2000, 28(3): 24-25)
- [7] Semic R, Lewis F L. Neural network approximation of piecewise continuous functions: Application to friction compensation[A]. *Proc of the IEEE Int Symposium on Control[C]*. Istanbul, 1997: 227-232
- [8] 张友旺, 钟向明, 黄元峰. 电液位置伺服系统自适应模糊神经网络控制[J]. *中国机械工程*, 2004, 15(8): 681-684
(Zhang Y W, Zhong X M, Huang Y F. Adaptive fuzzy neural network control for electro-hydraulic position servo system [J]. *China Mechanical Engineering*, 2004, 15(8): 681-684)

(上接第355页)

参考文献(References)

- [1] Zhang Q. Using wavelet network in nonparametric estimation [J]. *IEEE Trans on Neural Networks*, 1997, 8(2): 227-236
- [2] Daubechies I. The wavelet transform, time-frequency localization and signal analysis [J]. *IEEE Trans on Information Theory*, 1990, 36(5): 961-1005
- [3] Kugarajah T, Zhang Q. Multidimensional wavelet frames [J]. *IEEE Trans on Neural Networks*, 1995, 6(6): 1552-1556
- [4] Yacine Qussar, Gérard Dreyfus. Initialization by selection for wavelet network training [J]. *Neurocomputing*, 2000, 34: 131-143
- [5] Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries [J]. *IEEE Trans on Signal*

Processing, 1993, 41(12): 3397-3415

- [6] 吴清, 沈雪勤, 颜威利. 一种多维小波网络的构造性算法[J]. *计算机学报*, 2003, 26(9): 1206-1210
(Wu Q, Shen X Q, Yan W L. A multi-dimensional wavelet network construction algorithm [J]. *Chinese J of Computers*, 2003, 26(9): 1206-1210)
- [7] Daniel W C Ho, Zhang P, Xu J. Fuzzy wavelet networks for function learning [J]. *IEEE Trans on Fuzzy System*, 2001, 9(1): 200-211
- [8] 李治友, 陈才, 曹长修. 一种基于改进的RBF神经网络的铁水脱硫预报模型[J]. *重庆大学学报*, 2003, 26(9): 119-122
(Li Z Y, Chen C, Cao C X. A prediction model for molten iron desulfuration based on an improved RBFNN [J]. *J of Chongqing University*, 2003, 26(9): 119-122)