

基于 LS-SVM 的非线性系统盲辨识

朱燕飞¹, 谭洪舟², 章 云¹

(1 广东工业大学 自动化学院, 广东 广州 510090 2 中山大学 信息科学与技术学院, 广东 广州 510275)

摘要: 针对现今在非线性系统盲辨识研究中遇到的困难, 提出了一种基于最小二乘支持向量机 (LS-SVM) 的非线性系统盲辨识方法. 该方法直接对非线性系统输出进行过采样, 运用 LS-SVM 非线性建模技术, 并结合输入的分布特性, 从而完成非线性系统的盲辨识. 介绍了盲系统辨识问题的研究内容及过采样技术原理, 对 LS-SVM 的盲系统辨识机理和算法步骤进行了阐述. 仿真结果表明了该方法在解决非线性系统盲辨识问题上的切实可行性.

关键词: 盲辨识; 支持向量机; 过采样

中图分类号: TP181 文献标识码: A 文章编号: 1007-7162(2007) 02-0076-04

近十几年来, 盲信号处理逐渐成为学术界研究的一个热点. 作为盲信号处理一个重要分支——盲系统辨识, 现已在通信、测量、语音识别、雷达、地震、仿生等各个领域都有广泛的应用^[1-3]. 尽管这些领域的性质各不相同, 但研究中都有一个共同的特点, 即只有输出信号和输入信号的某些相关信息可用来分析和研究未知系统, 这就给系统辨识问题的解决带来了困难.

目前, 国内外处理盲系统辨识的方法主要有: 极大似然方法^[4]、高阶统计量方法^[56]、子空间方法^[7-8]等. 这些方法从不同角度对盲系统辨识问题展开了深入的研究, 特别是对于线性系统, 已基本趋于完善, 但在处理非线性系统的盲辨识问题上却遇到诸多困难. 对于非线性系统盲辨识的研究, Er-Wei Bai^[9-10]提出了基于过采样的 Hammerstein-Wiener 模型, 这种模型将系统分解为线性与非线性两部分的组合, 再利用输出过采样技术解决无输入作用下的系统辨识问题. 这种方法能够处理一定非线性系统的盲辨识; 但在 Hammerstein 或 Wiener 模型中, 盲系统辨识仍然只是针对其线性部分展开, 而非线性部分仍需采用传统的基于输入输出的系统辨识, 实质上仍未实现非线性系统的盲辨识.

针对这一问题, 本文将过采样技术扩展到非线性系统中, 直接对非线性系统输出进行过采样, 运用最小二乘支持向量机 (least squares support vector machines, LS-SVM) 的非线性建模技术, 并结合输入的分布特性, 实现非线性系统的盲辨识.

LS-SVM 是 J. A. K. Suykens^[11]提出的一种建立在统计学习理论基础上的机器学习算法. 它采用最

小二乘线性系统实现非线性系统的学习问题, 在精度要求范围内, 它有优良的学习速度和泛化能力; 而且其结构简单, 算法简练, 在非线系统辨识领域中有着广泛的应用前景.

1 盲辨识问题描述

盲系统辨识是盲信号处理中一个重要分支, 也是系统辨识领域的一个关键的领域. 简单地说, 它是在输出测量数据的分析与处理下, 寻求一种辨识方法或步骤, 使辨识出的模型特性与真实系统特性保持一致. 一般, 系统辨识总同时依赖于系统的输入和输出信号, 而系统盲辨识则是在输入信号未知的前提下, 仅依据输出信号来辨识系统模型. 图 1 给出了一个一般黑箱系统的盲辨识原理框图.

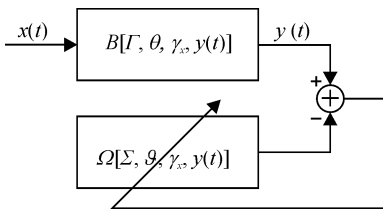


图 1 黑箱系统盲辨识框图

定义 1 黑箱系统 $B[\Gamma, \theta, \gamma_s, y(t)]$ 具有结构 Γ 和参数 θ 且该系统仅输出 $y(t)$ 可测和输入的某些信息 γ_s 已知; 若存在具有结构 Σ 和参数 J 的模型 $\Omega[\Sigma, J, \gamma_s, y(t)]$, 且 $\Sigma \neq \Gamma, J \neq \theta$ 只要 B 与 Ω 具有相同的外部特性, 则说明黑箱系统 B 具有可辨识性.

由定义 1 可知, 对于一个非线性系统 $y(t) =$

$f[x(t)]$, 盲辨识和建模的任务即是从输出测量数据及输入信号的某些特性中寻求一种参数或非参数的建模方法, 使其建立的模型与原系统具有一致的外部特性。

本文依据此定义, 在输出信号过采样技术的基础上, 提出了一种基于 LS-SVM 的非参数的非线性盲辨识方法。

2 过采样技术

一般说来, 仅仅依据输出测量信号进行系统的盲辨识是不可能的, 因为不同的系统, 特别是非线性系统, 由于输入的作用, 可能在采样点 kh 处产生相同的输出, 仅依赖于输出数据无法把握系统正确的模型特性。这时, 对系统的输出进行过采样, 并利用已知的输入分布特性则有可能解决问题。设输入采样周期为 T , 对输出进行过采样, 假定输出采样间隔为 $h = T/n$ (n 值的大小可根据实际情况调整)。则系统输入输出可表示为

$$\begin{aligned} \{y[kh]\} &\Leftrightarrow Y_h(z) = \sum_{k=0}^{\infty} y(kh)z^{-k}, \\ \{y[kT]\} &\Leftrightarrow Y_t(z) = \sum_{k=0}^{\infty} y(kT)z^{-k}, \\ \{y[kT]\} &\Leftrightarrow X_T(z) = \sum_{k=0}^{\infty} x(kT)z^{-k}. \end{aligned}$$

即经如此采样后, 有下面的式子成立:

$$\begin{aligned} x(t) &= 0 \quad t \neq kT; \\ y(kT) &= f\left\{y[(k-1)T], \dots, y[(k-p_y)T], \right. \\ &\quad \left. x[kT], \dots, x[(k-p_x)T]\right\}; \\ y(kh) &= f\left\{y[(k-1)h], \dots, y[(k-p_y)h], \right. \\ &\quad \left. x[kh], \dots, x[(k-p_x)h]\right\}; \end{aligned} \tag{1} \tag{2} \tag{3}$$

也即是说, 输出 y 在过采样后, 其过采样周期 kh 处对应的输入值 x 为零, 只有在原采样周期 kT 处输入才有值。

基于过采样的系统盲辨识即是利用过采样部分输入为零的特点, 只根据系统过采样输出及输入的某些信息 y_x 来辨识系统模型和参数, 即

$$y(kh) = f\left\{y[(k-1)h], \dots, y[(k-p_y)h], y_x\right\}.$$

ErWei Bai^[10]证实, 在可持续激励下, 对于线性系统, 无需了解输入任何信息, 利用系统过采样输出即可完全辨识系统模型和参数; 但对于非线性系统则达不到。本文将条件适当放宽, 在已知输入分布的前提下, 运用过采样机制, 并引入 LS-SVM 非线性建模技术, 从而达到非线性系统的盲辨识目的。

3 LS-SVM 的非线性系统盲辨识

3.1 LS-SVM 建模原理

LS-SVM 的非线性建模思想是将非线性问题经某个非线性变换, 转化成高维特征空间的线性问题。其变换函数: $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$, 为满足 Mercer 条件的核函数。

在高维线性空间中, 设定最优化目标函数为

$$\min_{w, b, e} \left(\frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2 \right) \tag{4}$$

$$s.t. \quad y_i[w^T \phi(x_i) + b] = 1 - e_i \quad (i = 1, 2, \dots, n).$$

构造拉格朗日函数求解, 有

$$\max_{\alpha} \min_{w, b} \left\{ L = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2 - \sum_{i=1}^n \alpha_i \left\{ y_i[w^T \phi(x_i) + b] - 1 + e_i \right\} \right\} \quad (i = 1, 2, \dots, n). \tag{5}$$

依据 Karush-Kuhn-Tucker 最优条件, 得到

$$\frac{\partial L}{\partial w} = 0 \quad \frac{\partial L}{\partial b} = 0 \quad \frac{\partial L}{\partial e_i} = 0 \quad \frac{\partial L}{\partial \alpha_i} = 0$$

求解上述方程, 得到最优解

$$w = \sum_{i=1}^n \alpha_i y_i \phi(x_i), \tag{6}$$

$$\sum_{i=1}^n \alpha_i y_i = 0 \tag{7}$$

$$\alpha_i = \gamma \cdot e_i = 0 \tag{8}$$

$$y_i[w^T \phi(x_i) + b] - 1 + e_i = 0 \tag{9}$$

转化为线性矩阵方程

$$\begin{bmatrix} 0 & Y^T \\ Y & ZZ^T + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{I} \end{bmatrix}, \tag{10}$$

其中: $Z = [\phi(x_1)^T y_1 \quad \phi(x_2)^T y_2 \quad \dots \quad \phi(x_n)^T y_n]^T$, $Y = [y_1 y_2 \dots y_n]^T$, $\tilde{I} = [1 \ 1 \dots 1]^T$, $e = [e_1 \ e_2 \dots e_n]^T$, $\alpha = [\alpha_1 \ \alpha_2 \dots \alpha_n]^T$. 用最小二乘法可解上系统 (10)。

上述建模原理中, 核函数的引入使得将数据隐式表达为特征空间, 并使得在其中训练一个线性学习器成为可能, 从而避免了计算其特征映射问题。然而, 核函数 $K(x_i, x_j)$ 的确定则成为模型特性确立的关键。Mercer 定理^[12]给出了连续对称核函数允许的表达方式, 即

$$K(x_i, x_j) = \sum_{k=1}^{\infty} l_k \phi_k(x_j) \phi_k(x_i) \quad (l_k \geq 0).$$

目前, 常见核函数选取主要有: 线性函数、多项式函数、径向基 (RBF) 函数、神经网络多层感知器函数等等。

假定系统输入 x 的分布 y_x 已知, 可依据输入的分布构造核函数 $K(x_i, x_j)$, 即隐式的确定特征空间的分布, 则可将未知输入的计算过程转化为已知特征空间函数的计算过程, 即

$$\mathbf{x}(t) \Leftrightarrow K(\mathbf{x}_b, \mathbf{x}_j | \mathbf{y}_x) \Leftrightarrow \phi(\mathbf{x}).$$

非线性系统盲辨识则运用构造的特征空间函数 $\phi(\mathbf{x})$ 代替未知的输入 $\mathbf{x}$, 以达到系统盲辨识目的. 这一过程, 同时也将系统的非线性辨识问题转化为高维特征空间的线性辨识问题, 从而使辨识过程更为简化.

3.2 建模步骤

- 1) 确定输入采样周期 T 和输出过采样周期 $h = T/h_c$;
- 2) 对输出进行过采样, 得到: $\{y[kh]\} \Leftrightarrow Y_h(z) = \sum_{k=0}^N y(kh)z^{-k}$, 并定义系统辨识集合: $\{y[(k-1)h], \dots, y[(k-p_y)h]\}$;
- 3) 依据当前输入的分布, 构造核函数 $K(\mathbf{x}_b, \mathbf{x}_j | \mathbf{y}_x)$, 计算特征空间函数 $\phi(\mathbf{x})$;
- 4) 运用 LS-SVM 辨识原理, 进行系统的非参数模型辨识, 求取或修正 LS-SVM 非线性模型, 并计算此时的优化目标结果;
- 5) 累积综合优化目标函数值, 判断其是否达到预定目标, 若达到, 结束计算; 否则返回步骤 2 继续.

4 仿真实例

设非线性输入输出关系为: $y(kT) = 0.15x \times (kT)^2 + 0.8x(kT)^3 + 0.15x(kT - 3T) \cdot x(kT - 5T)$. 已知输入 $x(kT) (k = 1, \dots, 6\,000)$, 在实数空间 $R \in [-5, 5]$ 范围内呈随机分布. 选取 RBF 函数作为核函数的结构, 即 $K(\mathbf{x}_b, \mathbf{x}_j) = \exp(-\|\mathbf{x}_j - \mathbf{x}_i\|^2 / \sigma^2)$, 设定 $\sigma = 0.03, T = 0.1, n = 10$.

运用上述辨识步骤对该过程进行 LS-SVM 非线性系统盲辨识. 图 2 给出了辨识输出和实际输出对比曲线及误差曲线的初始放大部分 ($k = 1, \dots, 120$). 辨识输出最大相对误差不超过 5%, 辨识效果优良.

采样不同的随机输入, 重复进行 10 次辨识过程, 求取每次辨识误差的标准差, 如表 1 所示. 从表 1 的结果可看出, 10 次重复辨识标准差的波动都很小, 辨识结果具有良好的稳定性和可靠性.

5 结论

文章运用过采样机制, 引入 LS-SVM 非线性建模技术, 并结合输入的分布特性, 实现了非线性系统的盲辨识. 可以说在解决直接非线性系统盲辨识问题上做了开拓性的工作. 仿真结果表明, 智能辨识算法的发展将为这一领域的研究带来生机.

这一算法过程属于黑箱系统的盲辨识过程, 其

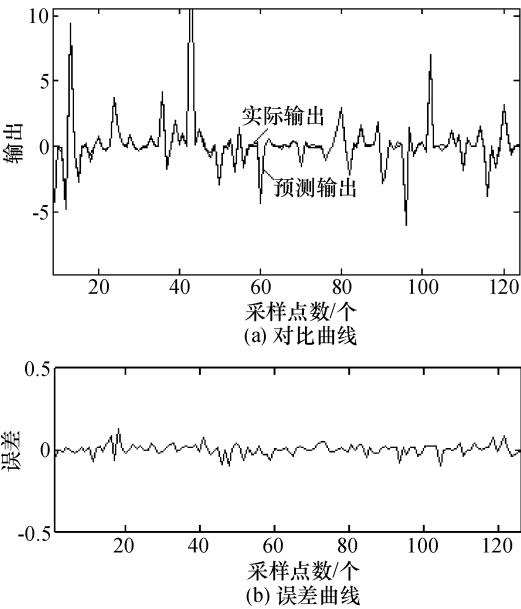


图 2 实际输出与辨识输出对比曲线及误差曲线

表 1 辨识误差的标准差

辨识次数	标准差	辨识次数	标准差
1	0.254	6	0.257
2	0.260	7	0.262
3	0.276	8	0.265
4	0.281	9	0.234
5	0.267	10	0.263

辨识的关键在于核函数的构造. 若输入的分布和范围不知, 则无法有效地进行辨识; 另外, 由于系统的本质非线性, 辨识过程中, P_y, n 等参数的确定都还带有一定的随机性. 这些问题在今后的研究中会进一步深入地探讨.

参考文献:

[1] Proakis J.G. Digital Communications [M]. 3rd Edition, McGraw-Hill, New York, 1995.

[2] Chow T.W.S., Tan H.-Z. HOS-based nonparametric and parametric methodologies for machine fault detection [J]. IEEE Transactions on Industrial Electronics, 2000, 47 (5): 1051-1059.

[3] LUO H., LI Y. The application of blind channel identification techniques to prestack seismic deconvolution [J]. Proceedings of IEEE, 1998, 86(10): 2082-2089.

[4] Abdel-Meraim K., Qiu W. Z., Hua Y. B. Blind system identification [J]. Proceedings of IEEE, 1998, 18(8): 1310-1322.

[5] Tan H.-Z., Chow T.W.S. Blind and total identification of ARMA mode in higher order cumulants domain [J]. IEEE Transactions on Industrial Electronics, 1999, 46(6): 1233-1240.

[6] Tan H.-Z., Chow T.W.S. Blind identification of quadratic

non linear models using neural networks with higher order cumulants[J]. IEEE Transactions on Industrial Electronics, 2000, 47(3): 687-696

[7] Fang Y W, Jiao L C. Blind identification of nonlinear FR-volterra channels[C] // The 5th International Conference on Signal Processing (WCC2000, ICSP2000). Beijing [s n], 2000. 294-297.

[8] Hua Y B, Abed-Meraim K, Wax M. Blind system identification using minimum noise subspace[J]. IEEE Transactions on Signal Processing 1997, 45(3): 770-773.

[9] Bai E W. A blind approach to the Hammerstein-Wiener model identification[J]. Automatica 2002, 38(6): 967-979

[10] Bai E W, Fu M. A blind approach to Hammerstein model identification[J]. IEEE Transactions on Signal Processing 2002, 50(7) : 1610-1619.

[11] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.

[12] Smola A J. Learning with kernels[D]. Germany: Technischen Universität Berlin 1998

Blind Nonlinear System Identification Based on LS-SVM

ZHU Yan-fei¹, TAN Hong-zhou², ZHANG Yun¹

(1. Faculty of Automation, Guangdong University of Technology, Guangzhou 510090, China
2. School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510275, China)

Abstract To solve the nonlinear problems in blind system identification, a novel blind nonlinear system identification approach based on Least-Square Support Vector Machines (LS-SVM) is investigated in this paper. By oversampling the system outputs, more information of the system characteristics can be observed to blindly identify nonlinear systems. The LS-SVM based mathematical approximation provides an adequate modeling of the unknown system given the distribution knowledge of the system inputs. The paper describes the mechanism of the oversampling and LS-SVM, explains the procedure of the algorithm of the blind system identification. Simulation results demonstrate the effectiveness of this approach.

Key words blind identification; nonlinear system; support vector machines; oversampling

加入 CEPS中文电子期刊服务的声明

《广东工业大学学报》自 2007年 6月起,加入 CEPS思博网 – 中文电子期刊服务. 思博网 – 中文电子期刊服务是目前台湾地区最大的期刊全文数据库, 其访问地址为: www.ceps.com.tw. 自此,读者可以通过这一网址检索《广东工业大学学报》自 2007年起各期的全文,在一段时期后,还可以回溯检索若干年前起历年的全文.

此外,由于《广东工业大学学报》被 CEPS收录,故凡向本刊投稿者,均视为其文稿刊登后可供思博网 (CEPS)收录、转载并上网发行;其作者文章著作权使用费与稿酬一次付清,本刊不再另付其他报酬.

请各位继续支持本刊,谢谢!

(广东工业大学学报编辑部)