

文章编号: 1000-8152(2005)02-0307-04

基于支持向量机的直接逆模型辨识

钟伟民, 皮道映, 孙优贤

(浙江大学 工业控制技术国家重点实验室, 现代控制工程研究所, 浙江 杭州 310027)

摘要: 在简单讨论逆模型辨识原理的基础上, 利用支持向量机(SVM)对函数逼近的能力, 提出了基于支持向量机的直接逆模型辨识方法. 分别采用二次核函数以及高斯 RBF 核函数, 利用训练数据对线性和非线性系统进行黑箱辨识. 仿真结果表明, 基于支持向量机的直接逆模型辨识方法在处理线性和非线性对象时, 辨识性能都优于传统的 BP 神经网络, 不仅辨识精度高, 辨识速度快, 而且泛化能力较强.

关键词: 逆模型; 支持向量机(SVM); BP 神经网络; 建模与辨识

中图分类号: TP18 **文献标识码:** A

Support vector machine based direct inverse-model identification

ZHONG Wei-min, PI Dao-ying, SUN You-xian

(National Laboratory of Industrial Control Technology and Institute of Modern Control Engineering,
Zhejiang University, Hangzhou Zhejiang 310027, China)

Abstract: After a simple discussion of the principle of the inverse-model identification, a support vector machines(SVM) based direct inverse-model identification method is developed by using SVM's excellent ability of function approximation. According to the train data, linear and nonlinear systems' black-box identification is performed by using SVM with quadric polynomial and Gaussian RBF kernel respectively. Simulation results show that the performance of SVM based direct inverse-model is better than that of BP neural network in that it has better identification precision, quicker identification speed and stronger generalization ability.

Key words: inverse-model; support vector machine(SVM); BP neural network; modeling and identification

1 引言(Introduction)

逆系统建模与辨识在控制领域中得到了广泛的应用^[1]. 它的主要思想是利用神经网络的逼近能力对系统的逆动态建模, 以使整个系统的输入输出为恒等映射, 从而实现高性能的控制^[2].

基于统计学习理论的支持向量机在结构和功能上与神经网络有类似. 文献[3]指出, 支持向量机可以逼近任意的一类非线性函数, 这为利用支持向量机进行系统辨识和控制提供了理论根据. 同时, 支持向量机运算速度快, 无局部最小点, 不会象神经网络一样陷入局部解. 近年来, 支持向量机在系统辨识及控制中已取得了突破性的进展^[4,5]. 据所掌握的文献来看, 国内外尚无利用支持向量机做逆模型建模和辨识的研究报道.

2 系统的可逆性(Reversibility of system)

对于线性系统, 可逆性问题实质上是一个能控性问题, 只要系统状态能控, 则系统是可逆的^[1].

对于非线性系统, 情况很复杂, 考虑离散单输入单输出系统

$$y(k+1) = f[y(k), \dots, y(k-n), u(k), \dots, u(k-m)]. \quad (1)$$

式中: $y(k) \in \mathbb{R}$, $u(k) \in \mathbb{R}$ 分别是系统的输出和输入, $m \leq n$, f 是一非线性函数.

给定 $[y(k), \dots, y(k-n), u(k-1), \dots, u(k-m)]$, 当任意的 $u'(k) \neq u(k)$ 时, 必有

$$f[y(k), \dots, y(k-n), u'(k), \dots, u(k-m)] \neq f[y(k), \dots, y(k-n), u(k), \dots, u(k-m)], \quad (2)$$

则称系统(1)在点 $[y(k), \dots, y(k-n), u(k-1), \dots, u(k-m)]^T$ 是可逆的.

定理 1^[1] 如果 $f[y(k), \dots, y(k-n), u(k), \dots, u(k-m)]$ 对于 $u(k)$ 是严格单调的, 那么系统在 $[y(k), \dots, y(k-n), u(k-1), \dots, u(k-m)]^T$

是可逆的.如果上述结论对任何时刻的 k 都成立,则系统(1)是可逆系统.

逆系统的建模方法包括直接逆系统建模、正-逆系统建模和逆-逆系统建模等.文中主要讨论直接逆系统建模.

3 支持向量机的函数拟合特性及其直接逆模型辨识 (Support vector regression and direct inverse-model identification)

3.1 支持向量机的函数拟合特性 (Support vector regression)

假设提供的训练样本集为 $(x_i, y_i), i = 1, \dots, n, x \in \mathbb{R}^d, y \in \mathbb{R}$, 求回归函数

$$y = f(x) = (w \cdot x) + b. \quad (3)$$

其中: (\cdot) 是内积, $w \in \mathbb{R}^d, b \in \mathbb{R}$. 则可以得到如下的优化问题

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (4a)$$

$$\text{s. t. } \begin{cases} y_i - (w \cdot x_i) - b \leq \varepsilon + \xi_i, \\ (w \cdot x_i) + b - y_i \leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0. \end{cases} \quad (4b)$$

这里的常数 C 是个惩罚系数,本文采用了 ε 不敏感损失函数

$$|\xi_i^*|_\varepsilon = \begin{cases} |y_i - (w \cdot x_i)|, & |y_i - (w \cdot x_i)| > \varepsilon, \\ 0, & \text{其他,} \end{cases} \quad (5)$$

引入拉格朗日乘子,以及利用映射函数 ϕ 组成满足 Mercer 条件的核函数 $K(x_i \cdot x) = (\phi(x_i) \cdot \phi(x))$, 根据 KKT 条件,可以得到最优函数拟合方程

$$f(x) = \sum_{i=1}^{nsv} a_i K(x_i \cdot x) + b. \quad (6)$$

其中: x_i 为支持向量, nsv 为支持向量的个数, a_i 为对应支持向量的系数, b 为阈值,可由任意两个支持向量求解.

3.2 基于支持向量机的直接逆模型建模 (SVM based direct inverse-model identification)

图1给出了基于支持向量机的直接逆模型建模的框图.

图中, u_d 是输入给未知对象的给定信号, y 为未知对象的输出,以 y 和 u 作为 SVM 网络的输入和输出,以误差 e 来训练 SVM 网络.为了获得良好的逆动力学特性,网络训练所覆盖的范围要比未知系统所需的大一些.

令拟合因子

$$X = [y(k+1), y(k), \dots, y(k-n), u(k-1), \dots, u(k-m)], \quad (7)$$

根据 d 对训练数据 $\{X_i, u_i\} (i = 1, \dots, d), u_i = u(i)$, 通过 SVM 的学习,可以获取相应的输入向量的系数 a_i (这里把系数为零的也放在公式里) 和阈值 b . 根据当前的输入 X , 可以辨识出逆模型的输出

$$u(k) = If[X] = \sum_{i=1}^d a_i (\phi(X_i) \cdot \phi(X)) + b = \sum_{i=1}^{nsv} a_i K(x_i \cdot X) + b. \quad (8)$$

式中 If 代表非线性函数的逆.式(8)中 $u(k)$ 在原始空间中是 X 中各元素的非线性组合,但由于采用了非线性映射函数 ϕ , 在特征空间中,却又是 X 中各元素的线性组合^[4].这正是 SVM 进行非线性函数拟合的基本思想:巧妙利用非线性映射函数把输入空间的数据投影到特征空间中进行线性拟合.

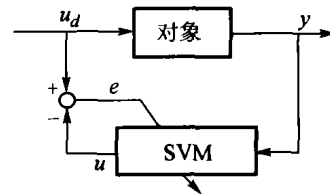


图1 支持向量机直接逆系统建模框图
Fig. 1 Block diagram of SVM based direct inverse-model identification

利用 SVM 建模和辨识的精度主要取决于以下几点:

- 1) 损失函数的选择,这还影响到系统抗噪声能力和鲁棒性,本文采用 ε 不敏感函数;
- 2) C 和 ε 的选取,这是系统拟合的平滑性 (flatness) 和拟合误差之间的折衷,一般靠经验选取.文中,选取 $C = 10000, \varepsilon = 0.001$;
- 3) 核函数的选取,现在用得比较多的有线性、多项式、高斯、B 样条函数等,对于不同的对象,核函数的不同选取,会对系统的拟合速度和精度产生影响.本文分别选取多项式核函数 $K(x_i, x) = ((x \cdot x_i) + 1)^2$ 以及高斯核函数 $K(x_i, x) = \exp\left\{-\frac{\|x - x_i\|^2}{2\sigma^2}\right\}$, 其中 $\sigma = 20$. 多项式核函数结构简单,用得最广泛.高斯核函数 SVM 比 RBF 神经网络最大的优点就在于 SVM 算法自动选择最优的隐节点数目以及隐节点中心,而 RBF 神经网络还要人工选择节点数以及通过聚类等方法确定隐节点中心;
- 4) 式(7)中的拟合因子选取的越多,精度会越

高,但随之拟合速度会变得很慢.可以在仿真过程中逐渐增加因子,在精度和速度之间折衷选择最佳的拟合因子.

4 仿真(Simulation)

本文将利用 SVM 和 BP 神经网络分别对一线性性和非线性系统进行逆模型仿真,以进行性能上的对比.本文采用 SVM MATLAB Toolbox² 作仿真.

4.1 线性系统(Linear system)

设系统可以用下面的差分形式表示^[6]:

$$y(k+1) = 0.3y(k) + 0.4y(k-1) + 1.25u(k) + 0.75u(k-1). \quad (9)$$

假定不知道系统结构,令

$$X = [y(k+1), y(k), y(k-1), y(k-2), u(k-1), u(k-2)].$$

控制信号 $u(k)$ 由斜坡函数和正弦函数组合而成:

$$u(k) = \begin{cases} k/20, & k \leq 20, \\ 1 - (k - 20)/20, & 20 < k \leq 40, \\ \sin(0.1k - 4.1), & k > 40. \end{cases} \quad (10)$$

由于控制信号的输入范围是 -1 到 +1,本文用 -1.5 到 +1.5 区间内的随机信号去训练网络.图 2 和图 3 分别给出了 SVM 根据式(9)对于控制信号(10)的仿真结果,其中: * 点为实际控制信号,实线为模型的输出.

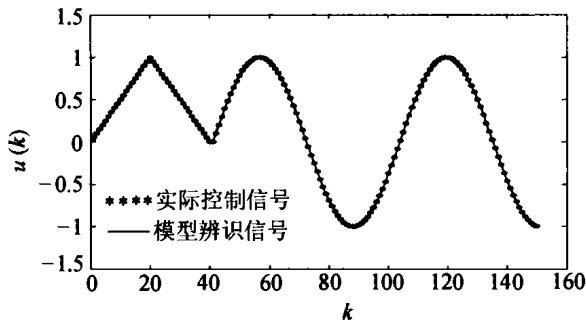


图 2 多项式核 SVM 逆模型仿真曲线
Fig. 2 Simulation results of SVM with quadric kernel based direct inverse-model identification

4.2 非线性系统(Nonlinear system)

设系统可以用下面的方程表示^[6]:

$$y(k+1) = 6y(k)/(1 + y(k)^2) + u(k) + 0.3u(k-1). \quad (11)$$

假定不知道系统结构,令

$$X = [y(k+1), y(k), y(k-1), y(k-2), u(k-1), u(k-2)].$$

控制信号 $u(k)$ 仍然由式(10)提供.

非线性建模中,所有的参数选择和线性系统一样.图 4 和图 5 给出了 SVM 根据式(11)对于控制信号(10)的仿真结果,其中: * 点为实际控制信号,实线为模型的输出.

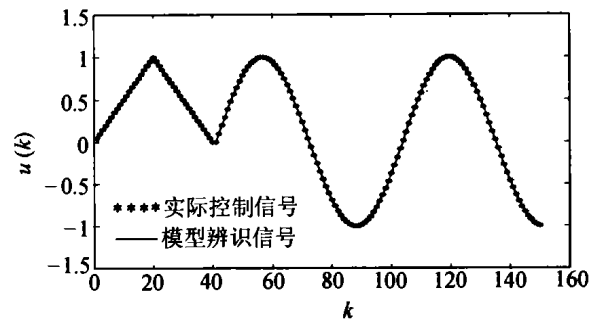


图 3 高斯核 SVM 逆模型仿真曲线
Fig. 3 Simulation results of SVM with Gaussian kernel based direct inverse-model identification

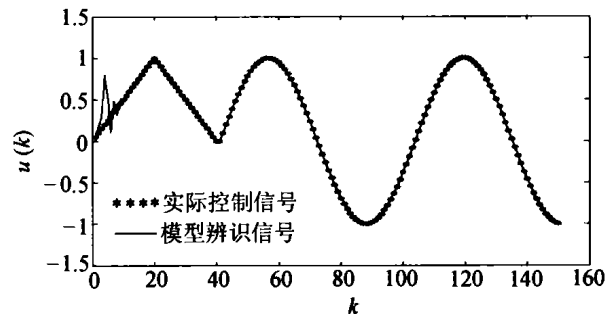


图 4 多项式核 SVM 逆模型仿真曲线
Fig. 4 Simulation results of SVM with quadric kernel based direct inverse-model identification

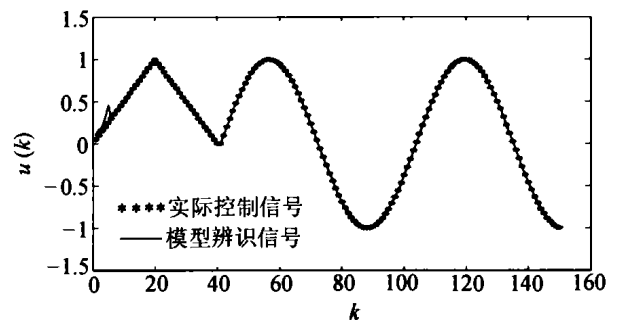


图 5 高斯核 SVM 逆模型仿真曲线
Fig. 5 Simulation results of SVM with Gaussian kernel based direct inverse-model identification

4.3 性能比较(Comparison of performance)

本文对系统(9)和(11)分别建立 BP 神经网络进行仿真,以和 SVM 进行比较.线性系统(9)中采用 3 层 6-8-1 模式,6 个节点的输入层,8 个隐节点,输出为一个,即 $u(k)$. 采用 MATLAB 神经网络工具箱进行仿真^[7],函数选取 logsig 和 purelin,用 Levenberg-Marquardt 算法进行训练.非线性系统(11)中采用 3 层 6-10-1 模式,其余均和线性系统相

同.图6和图7分别为系统(9)和(11)的BP神经网络逆模型辨识仿真曲线.

从给出的仿真曲线可以看出,不管是SVM还是BP神经网络,都能很好地进行逆模型辨识.处理线性系统逆模型辨识时,多项式核函数和高斯核函数SVM的精度都是相当高的.在处理非线性对象时,2阶多项式核函数SVM的精度和BP神经网络的相似,在起始阶段均存在较大超调,但在经过几步调整之后即能很好地跟踪设定曲线.高斯核函数SVM的逆模型辨识效果明显要优于多项式核函数SVM以及BP神经网络的.另外,SVM结构简单,可调参数少,训练相对容易^[3].SVM进行系统逆模型辨识的速度明显优于BP神经网络,尤其是在处理非线性对象的时候.在1.5G的CPU主频下,仿真系统(11),SVM在辨识过程中花费的时间小于1min,而BP神经网络至少需要0.5h以上,尤其在权系数初值设置不当的情况下,更是耗时.

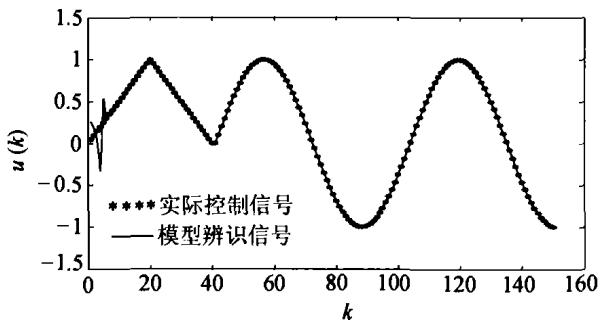


图6 BP神经网络逆模型仿真曲线(线性系统)
Fig. 6 Simulation results of BP NN based direct inverse-model identification(linear system)

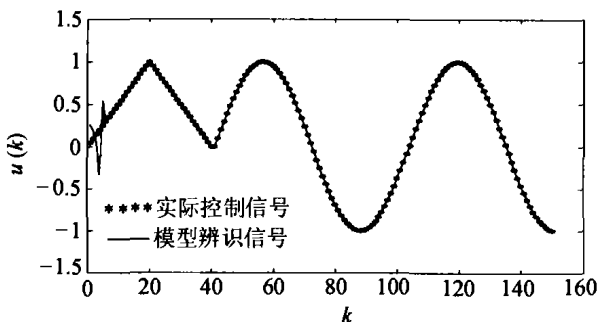


图7 BP神经网络逆模型仿真曲线(非线性系统)
Fig. 7 Simulation results of BP NN based direct inverse-model identification(nonlinear system)

5 结论(Conclusion)

本文提出了基于支持向量机的直接逆模型辨识方法,并分别对一线性和非线性系统进行了仿真.两个例子均参考文献[6],其中非线性的例子,稍加修改,增加了 $u(k-1)$ 项.两个例子中,采用同一个控制律,可以更好地说明模型的泛化能力.通过仿真,

可以得到以下结论:

- 1) 支持向量机能很好地进行线性和非线性系统的逆模型辨识;
- 2) 支持向量机逆模型辨识的精度高于BP神经网络,合理选择核函数及其参数,将影响到支持向量机的辨识精度;
- 3) 支持向量机结构比BP神经网络简单,虽然有损失函数、核函数选择问题,但总体上讲,可调参数要比BP神经网络少得多;
- 4) 支持向量机辨识速度比BP神经网络来得快,支持向量机的训练时间远远少于BP神经网络;
- 5) 由控制律的多样性(斜坡,正弦)可以看出,支持向量机的泛化能力很强.

因此,基于支持向量机的逆模型辨识方案是可行的,完全可以应用到一些具体的工业逆模型控制场合中.

参考文献(References):

- [1] 王永骥,涂健.神经网络控制[M].北京:机械工业出版社,1998.
(WANG Yongji, TU Jian. *Neural Network Control* [M]. Beijing: China Machine Press, 1998.)
- [2] 王树青等.先进控制技术及应用[M].北京:化学工业出版社,2001.
(WANG Shuqing etc. *Advanced Control Technology and Its Application* [M]. Beijing: Chemical Industry Press, 2001.)
- [3] VAPNIK V N. *The Nature of Statistical Learning Theory* [M]. New York: Springer-Verlag, 1995.
- [4] DREZET P M L, HARRISON R F. Support vector machines for system identification [C] // *Proc of UKACC Int Conf on Control '98*. UK: University of Wales Press, 1998: 668 - 692.
- [5] SUYKENS J A K. Nonlinear modeling and support vector machines [C] // *Proc of IEEE Instrumentation and Measurement Technology Conference*. Budapest, Hungary: IEEE Press, 2001: 287 - 294.
- [6] MA Xiaomin. Inverse identification and closed-loop control of dynamic systems using neural networks [J]. *Control Theory & Applications*, 1997, 14(6): 829 - 836.
- [7] 闻新,周露,王丹力,等. MATLAB神经网络应用设计[M].北京:科学出版社,2000.
(WEN Xin, ZHOU Lu, WANG Danli, et al. *Application and Designation of MATLAB Neural Network* [M]. Beijing: Science Press, 2000.)

作者简介:

- 钟伟民 (1976—),男,博士研究生,主要研究方向为非线性控制,软测量以及复杂过程智能建模等,E-mail:wmzhong@iipc.zju.edu.cn;
- 皮道映 (1965—),男,浙江大学教授,博士生导师,研究领域为智能学习、工业过程建模与控制等,E-mail:dypi@iipc.zju.edu.cn;
- 孙优贤 (1940—),男,浙江大学教授,工程院院士,博士生导师,主要从事控制理论与应用、工业自动化方向的研究等.