

# 基于支持向量机的 可分离非线性动态系统辨识<sup>1)</sup>

张莉 席裕庚

(上海交通大学自动化研究所 上海 200030)  
(E-mail: zhangli@mail.xidian.edu.cn, ygxi@sjtu.edu.cn)

**摘 要** 针对状态变量和控制变量可分离的非线性动态系统模型, 通过引入两个非线性核函数重新设计了标准支持向量机的回归估计模型, 使之适用于非线性动态系统的辨识. 它包含两个分别关于状态变量和控制变量的非线性函数, 用于辨识可分离变量非线性动态系统中的两个非线性函数. 文中的仿真实验验证了我们算法用于非线性动态系统辨识的有效性.

**关键词** 系统辨识, 支持向量机, 回归估计, 核函数  
**中图分类号** TP18

## Identification of Separable Variable Nonlinear Dynamical System Based on SVMs

ZHANG Li XI Yu-Geng

(Institute of Automation, Shanghai Jiaotong University, Shanghai 200030)  
(E-mail: neuluojia@163.com)

**Abstract** In this paper, for the case that the state variables can be separated from the control variables in a nonlinear dynamic system, an improved regression estimation algorithm based on SVMs is presented and then applied to nonlinear system identification. This kind of learning machines includes two nonlinear functions whose variables are the state and the control ones, respectively, and they are used to identify the two nonlinear functions in the separable-variable nonlinear dynamic system. The simulation results validate the efficiency of our method.

**Key words** System identification, support vector machine (SVM), regression estimation, kernel function

## 1 引言

基于统计学习理论的支持向量机是一种新兴的通用学习工具<sup>[1,2]</sup>. 由于在支持向量机中引入了 Mercer 函数, 可实现强大的非线性逼近功能, 这保证了回归估计支持向量机在非线性系统辨识中的应用<sup>[3]</sup>. 本文针对一类可分离变量的非线性动态系统, 通过引入两个非线性核函数改进了标准支持向量机的回归估计模型, 并将它用于可分离变量的非线性动

1) 上海市科委项目 (04DZ11008)  
Supported by Science Committee of Shanghai (04DZ11008)  
收稿日期 2004-4-16 收修改稿日期 2004-12-28  
Received April 16, 2004; in revised form December 28, 2004

态系统辨识. 由于它继承了支持向量机的优良特性, 具有强大的非线性逼近能力和良好的推广能力, 避免了过拟合和局部极小, 适用于小样本甚至是非持续激励样本. 该方法突破了标准回归估计支持向量机只能将所有变量混在一起估计的限制, 利用了更多的先验知识, 从模型上直接消除了状态变量与输入控制之间的交叉影响. 文中通过多个仿真实验验证了这一方法的有效性, 并与神经网络和标准支持向量机的系统辨识方法作了比较.

## 2 问题描述

本文考虑如下状态变量和控制输入可分离的含噪离散非线性动态系统:

$$y(k+1) = f(\mathbf{y}(k), \mathbf{u}(k)) = g(\mathbf{y}(k)) + h(\mathbf{u}(k)) + e(k+1) \quad (1)$$

其中  $\mathbf{y}(k) = [y(k), y(k-1), \dots, y(k-n+1)]^T \in \mathcal{Y} \subset \mathbb{R}^n$ ,  $y(k)$  是非线性系统的输出;  $\mathbf{u}(k) = [u(k), u(k-1), \dots, u(k-m+1)]^T \in \mathcal{U} \subset \mathbb{R}^m$ ,  $u(k)$  是非线性系统的输入,  $\mathcal{Y}$  和  $\mathcal{U}$  是紧集.  $g(\cdot)$  和  $h(\cdot)$  是未知的非线性函数,  $e(k)$  是系统的干扰. 这里假设  $n$  和  $m$  是系统已知的结构阶数. 为了确保系统可辨识, 假设对于有界的控制输入  $\mathbf{u}(0) \in \mathcal{U}$  和有界的初始状态  $\mathbf{y}(0) \in \mathcal{Y}$ , 系统的输出保持有界; 同时假设非线性函数  $g(\cdot)$  和  $h(\cdot)$  分别在有界紧集  $\mathcal{Y}$  和  $\mathcal{U}$  上连续.

如果有先验知识表明, 待辨识的非线性动态系统其状态变量和输入变量是可分离的, 如 (1) 式. 标准支持向量机无法利用这个先验知识, 虽然辨识仍然可以进行, 但是辨识精度会受到状态变量和输入变量的交叉影响而降低, 要求的样本也将更多, 因为标准支持向量机是将所有输入变量合在一起估计与输出的. 为此我们通过引入两个非线性核函数, 修改了支持向量机的回归估计模型, 并用它对模型 (1) 所表示的非线性动态系统进行辨识.

## 3 改进的回归估计支持向量机

对于模型 (1) 所表示状态变量和输入变量可分离的非线性动态系统, 采用串并结构, 对非线性动态系统 (1) 式进行批处理辨识. 已知辨识数据  $\{(\mathbf{y}(k), \mathbf{u}(k), y(k+1)) | \mathbf{y} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^m, y \in \mathbb{R}, k = 1, \dots, l\}$ , 其中  $l$  是训练样本的个数. 首先对输入样本中状态变量和控制变量分别采用两个 Mercer 核映射将其映射到两个特征空间:  $\mathbf{y} \xrightarrow{\phi} \phi(\mathbf{y}), \mathbf{u} \xrightarrow{\psi} \psi(\mathbf{u})$ , 然后在特征空间中进行线性回归估计

$$\hat{y}(k+1) = \hat{f}(\mathbf{y}(k), \mathbf{u}(k)) = \mathbf{w}_1^T \phi(\mathbf{y}(k)) + \mathbf{w}_2^T \psi(\mathbf{u}(k)) + b \quad (2)$$

引入结构风险, 最小化经验风险与模型容量控制之间的一个好的折衷:

$$\begin{aligned} \min \quad & \frac{1}{2}(\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2) + C \sum_{k=1}^l (\xi(k+1) + \xi^*(k+1)) \\ \text{subject to} \quad & y(k+1) - \hat{f}(\mathbf{y}(k), \mathbf{u}(k)) \leq \varepsilon + \xi(k+1) \\ & \hat{f}(\mathbf{y}(k), \mathbf{u}(k)) - y(k+1) \leq \varepsilon + \xi^*(k+1) \\ & \xi(k+1), \xi^*(k+1) \geq 0, \quad k = 1, \dots, l \end{aligned} \quad (3)$$

其中预先选取的正常数  $C$  为惩罚因子,  $\varepsilon$  为允许误差. 求解上述规划 (3) 的 Wolfe 对偶规划, 我们可得:

$$\max - \frac{1}{2} \sum_{k=1}^l \sum_{j=1}^l (\alpha_k - \alpha_k^*)(\alpha_j - \alpha_j^*) [(\phi(\mathbf{y}(k))\phi(\mathbf{y}(j))) + (\psi(\mathbf{u}(k))\psi(\mathbf{u}(j)))] +$$

$$\begin{aligned} & \sum_{k=1}^l (\alpha_k - \alpha_k^*) y(k+1) - \varepsilon \sum_{k=1}^l (\alpha_k + \alpha_k^*) \\ \text{subject to } & \sum_{k=1}^l (\alpha_k - \alpha_k^*) = 0, \quad \alpha_k, \alpha_k^* \in [0, C], \quad k = 1, \dots, l \end{aligned} \quad (4)$$

由于  $\phi$  和  $\psi$  为 Mercer 核映射, 所以可以分别用两个 Mercer 核函数  $K_\phi(\cdot, \cdot)$  和  $K_\psi(\cdot, \cdot)$  代替 (4) 式中的内积  $(\phi(\mathbf{y}(k))\psi(\mathbf{y}(k)))$  和  $(\psi(\mathbf{u}(k))\psi(\mathbf{u}(k)))$ . 这样改进的回归估计支持向量机的输出为

$$\hat{y}(k+1) = \sum_{j=1}^l (\alpha_j^* - \alpha_j) K_\phi(\mathbf{y}(j), \mathbf{y}(k)) + \sum_{j=1}^l (\alpha_j^* - \alpha_j) K_\psi(\mathbf{u}(j), \mathbf{u}(k)) + b \quad (5)$$

其中  $b$  可以由  $KKT$  条件<sup>[1,2]</sup>求得.

在改进的回归估计支持向量机中核函数  $K_\phi(\cdot, \cdot)$  和  $K_\psi(\cdot, \cdot)$  以及核函数中的参数可以相同, 也可以不同. 这样该方法一方面采用了非线性动态系统模型 (1) 式中状态变量和控制输入可分离的先验知识, 避免了状态变量与控制输入对辨识过程的交叉影响; 另一方面由于引入了两个 Mercer 核函数, 这两个核函数的类型和参数都可以独立选择, 从而能更好地逼近非线性动态系统.

## 4 仿真实验

在以下的实验比较中, 均采用多层前向神经网络、标准回归估计支持向量机和本文提出的方法进行辨识. 辨识误差定义为  $error = \sqrt{\sum_{i=1}^{l_{Test}} (\hat{y}_i - y_i)^2 / l_{Test}}$ , 其中  $\hat{y}$  是  $y$  的估计值,  $l_{Test}$  是指测试集上的样本个数. 本文实验中核参数的选取采用在局部范围内依据交叉验证误差寻优方法<sup>[4]</sup>.

3 个单输入单输出的未知系统见表 1 所示<sup>[5]</sup>, 其实验的结果见表 2. 图 1 和图 2 分别给出了例 1 和例 3 的辨识结果. 辨识误差显示在相应的 (a) 图中, 其中的细实线、虚线和粗实线分别表示了这三种方法得到的误差. 相应的 (b) 图用实线和点线比较了原模型与本文方法得到的辨识模型.

表 1 3 个未知的辨识系统  
Table 1 Three identified systems

例子	干扰 $e(k)$	$u(k)$ 的取值	
		训练	检验
1. $y(k+1) = 0.3y(k) + 0.6y(k-1) + 0.6 \sin(\pi u(k)) + 0.3 \sin(3\pi u(k)) + 0.1 \sin(5\pi u(k)) + e(k+1)$	零均值的、标准差为 0.1 的高斯白噪声	$[-1, 1]$ 均匀随机变量	$u(k) = \sin(2\pi k/250)$
2. $y(k+1) = \frac{y(k)y(k-1)[y(k)+2.5]}{1+y^2(k)+y^2(k-1)} + u(k) + e(k+1)$	零均值的、标准差为 0.1 的高斯白噪声	$[-2, 2]$ 均匀随机变量	$u(k) = \sin(2\pi k/25) + \sin(2\pi k/10)$
3. $y(k+1) = \frac{y(k)}{1+y^2(k)} + u^3(k) + e(k+1)$	零均值的、标准差为 0.1 的高斯白噪声	$[-2, 2]$ 均匀随机变量	$u(k) = \sin(2\pi k/25) + \sin(2\pi k/10)$

表 2 表 1 中 3 个例子的辨识结果  
Table 2 The identification results of three examples in Table 1

算法	参数	辨识误差
例 1	多层前向神经网络 $\mathcal{N}_{3,40,20,1}^3$ , <sup>①</sup> $l = 5000$	0.8286
	标准支持向量机 $C = 100, \varepsilon = 0.01, l = 100$	0.4576
	改进的支持向量机 $C = 100, \varepsilon = 0.01, l = 100$	0.0584
例 2	多层前向神经网络 $\mathcal{N}_{3,40,20,1}^3, l = 5000$	0.2557
	标准支持向量机 $C = 100, \varepsilon = 0.01, l = 100$	0.1494
	改进的支持向量机 $C = 100, \varepsilon = 0.01, l = 100$	0.0959
例 3	多层前向神经网络 $\mathcal{N}_{2,40,20,1}^3, l = 5000$	0.3129
	标准支持向量机 $C = 10000, \varepsilon = 0.01, l = 100$	0.1564
	改进的支持向量机 $C = 1000, \varepsilon = 0.01, l = 100$	0.0557

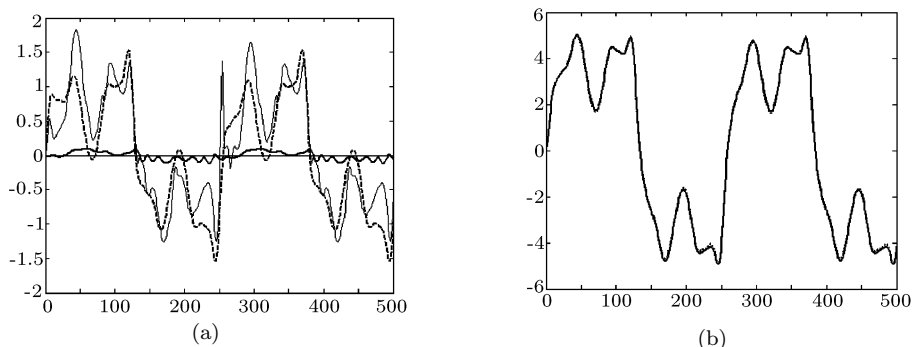


图 1 例 1 的辨识误差图 (a) 和原模型及本文方法得到的辨识模型 (b)

Fig. 1 Example 1: (a) The identification error curves obtained by three methods. (b) Outputs of the original system and the identification model obtained by our method

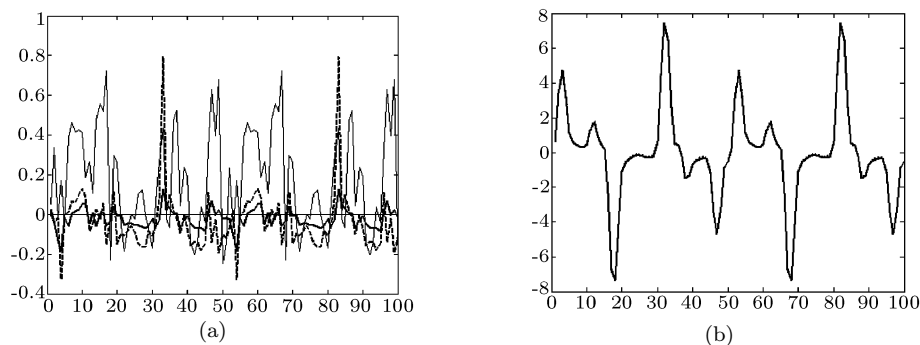


图 2 例 3 的辨识误差图 (a) 和原模型及本文方法得到的辨识模型 (b)

Fig. 2 Example 3: (a) The identification error curves obtained by three methods. (b) Outputs of the original system and the identification model obtained by our method

① 上标表示有三层网络, 下标表示输入层的输入是三维的, 两个隐层的节点个数分别为 40 个和 20 个, 输出层的节点个数为 1.

②  $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2p^2), (p \neq 0 \in \mathbb{R})$

③  $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$

例 4. 我们用同样的方法来辨识多输入多输出系统. 假设未知的多输入多输出辨识系统可以用如下的差分方程来表示<sup>[5]</sup>:

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \end{bmatrix} = \begin{bmatrix} y_1(k)/(1+y_2^2(k)) \\ y_1(k)y_2(k)/(1+y_2^2(k)) \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} + \begin{bmatrix} e_1(k+1) \\ e_2(k+1) \end{bmatrix}$$

其中  $e_1(k)$  和  $e_2(k)$  是零均值的、标准差为 0.1 的高斯白噪声序列. 在训练时,  $u_1(k)$  和  $u_2(k)$  都取  $[-1, 1]$  区间的均匀随机变量; 在检验时,  $u_1(k) = \sin(2\pi k/25)$ ,  $u_2(k) = \cos(2\pi k/25)$ . 实验的结果如表 3 所示, 可以看出对多输入多输出的例子, 我们方法仍然有较好的辨识能力.

表 3 例 4 的辨识结果  
Table 3 The identification result of example 4

算法	参数	辨识误差	
		$y_1$	$y_2$
多层前向神经网络	$\mathcal{N}_{3,50,25,1}^3, l = 5000$	0.7902	0.4251
标准支持向量机	$C = 1000, \varepsilon = 0.01, l = 100$ 核函数为高斯核	0.3062	0.3921
改进的支持向量机	$C = 1000, \varepsilon = 0.01, l = 100$ $K_\phi(\cdot, \cdot)$ 高斯核, $K_\psi(\cdot, \cdot)$ 线性核	0.2476	0.2613

## 5 结束语

本文针对一类可分离变量的非线性动态系统, 通过引入两个非线性核函数改进了标准支持向量机的回归估计模型. 本文方法解决了标准支持向量机只能将所有变量混在一起估计的限制, 利用了更多的先验知识, 从模型上直接消除了状态变量与输入控制之间的交叉影响. 实验结果表明对于模型 (1) 式描述的非线性动态系统, 我们的方法比标准支持向量机具有更好的辨识精度.

## References

- 1 Vapnik V. Statistical Learning Theory. Inc., New York: John Wiley and Sons, 1998
- 2 Drucker H, Burges C, et al. Support vector regression machines. *Advances in Neural Information Processing Systems*, 1997, **9**(9): 155~161
- 3 Drezet P. Support vector machines for system identification. In: *Proceedings of UKACC International Conference on Control'98*, 1998. **1**(1): 688~692
- 4 Kearns M, Ron D. Algorithmic stability and sanity-check bounds for leave-one-out cross validation. In: *Proceedings of the Tenth Conference on Computational Learning Theory*. New York: ACM Press, 1997. 152~162
- 5 Narendra K S, Parthasarathy K. Identification and control of dynamic systems using neural networks, *IEEE Transactions on Neural Networks*, 1990, **1**(1): 4~27

张 莉 2002 年在西安电子科技大学获得博士学位, 现在在上海交通大学博士后流动站工作. 主要研究方向有模式识别、人工神经网络和机器学习.

(ZHANG Li Received her Ph. D. degree in Electronic Engineering from Xidian University in 2002. She is currently a postdoctor at the Institute of Automation of Shanghai Jiaotong University. Her research interests include pattern recognition, neural networks, and machine learning.)

席裕庚 1984 年在德国慕尼黑大学获得博士学位. 现为上海交通大学教授, 博士生导师. 研究兴趣包括预测控制, 大系统理论, 智能机器人系统与技术等.

(XI Yu-Geng Received his Ph. D. degree from Technical University of Munich, Germany in 1984. Now he is a professor at Shanghai Jiaotong University. His research interests include predictive control, control of large scale and complex system, and system and technology of intelligent robot.)