

文章编号: 1001-0920(2005)05-0537-04

一种模糊支持向量机控制器的研究

袁小芳, 王耀南

(湖南大学 电气与信息工程学院, 湖南 长沙 410082)

摘要: 将支持向量机与模糊逻辑相结合, 设计了一种模糊支持向量机控制器, 并分析了控制器的结构和学习算法。学习过程分为离线学习支持向量机和在线整定模糊比例因子两部分。与模糊神经网络控制器相比, 模糊支持向量机控制器适应小样本学习, 泛化能力强, 解决了过学习、结构设计依赖经验等问题。仿真研究表明, 所设计的控制器具有较优的控制性能。

关键词: 机器学习; 支持向量机; 模糊逻辑; 模糊神经网络

中图分类号: TP18 **文献标识码:** A

On fuzzy support vector machine controller

YUAN Xiao-fang, WANG Yao-nan

(College of Electrical and Information Engineering, Hunan University, Changsha 410082, China Correspondent: YUAN Xiao-fang, Email: yuanxiaof@21cn.com)

Abstract: A fuzzy support vector machine (FSVM) controller is proposed which integrates fuzzy logic and support vector machines (SVM). The structure and learning algorithms of proposed FSVM controller is analysed. And its learning procedure includes two parts: offline training of SVM and online training of fuzzy scale factors. Compared with fuzzy neural networks (FNN) controller, FSVM controller has smaller training-size, better generalization ability, and avoids problems of over-fitting, experiential selection of structure. Simulation results show that FSVM controller has good control capability.

Key words: machine learning; support vector machines; fuzzy logic; fuzzy neural networks

1 引言

模糊神经网络(FNN)^[1]既有模糊推理的结构知识表达能力,又有神经网络的自学习能力,在控制领域得到了广泛应用。但是FNN的结构和类型都是人为设定的,控制性能依赖于专家的经验。同时它在过学习、泛化能力方面也存在不足。

支持向量机(SVM)^[2,3]是Vapnik在统计学习理论上提出的一种新型机器学习方法,目前已应用于模式识别、函数估计、优化控制等多个领域^[4~6]。SVM学习是依据结构风险最小化(SRM)^[3]准则,比基于经验风险最小化(ERM)^[3]准则的神经网络具有更强的理论依据,克服了神经网络过学习、泛化能力不强、依赖经验选择结构和类型等缺点,因

而本文采用SVM代替神经网络来设计控制器。

本文设计了一种模糊支持向量机(FSVM)控制器,并分析了该控制器的结构和学习算法。控制器融合了模糊逻辑和SVM的特点,控制系统输入输出采取模糊化处理,控制决策由SVM完成。控制器的学习分为离线和在线学习两部分,离线学习由序贯最小优化算法(SMO)^[7]训练SVM,在线学习利用BP算法整定模糊比例因子。与模糊神经网络控制器相比,FSVM控制器适应于小样本学习,泛化能力强,结构由二次规划得到而不是人为设定,二次规划过程也不存在过学习问题。FSVM控制器具有优良的控制性能。

收稿日期: 2004-07-05; 修回日期: 2004-09-23

基金项目: 国家自然科学基金项目(60375001)。

作者简介: 袁小芳(1979—),男,湖南安仁人,博士生,从事智能控制、神经网络等研究; 王耀南(1957—),男,云南昆明人,教授,博士生导师,从事智能控制、智能信息处理等研究。

2 支持向量机回归介绍

设给定训练样本为 $\{(x_1, y_1), \dots, (x_n, y_n)\} \subset R^n \times R$, 利用一个非线性映射 Φ 将数据映射到高维特征空间, 在高维特征空间进行线性回归 设 SVM 回归为

$$f(x) = (w \cdot \Phi(x)) + b \quad (1)$$

其中: 矢量 $w \in R^n$, 偏置 $b \in R$. SVM 是在最小化样本误差的同时, 缩小模型泛化误差的上界, 即最小化模型的结构风险, 从而得到其优化问题

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

$$\begin{aligned} \text{s t } & y_i - f(x_i) \leq \epsilon + \xi_i, \\ & f(x_i) - y_i \leq \epsilon + \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0, i = 1, \dots, n. \end{aligned} \quad (3)$$

方程 (2) 第 1 项为最小化结构风险, 第 2 项为最小化样本误差, 常数 C 是在结构风险与样本误差之间作出的折衷, ξ_i 和 ξ_i^* 为松弛变量, ϵ 用于定义 ϵ 线性不敏感损失函数

$$\begin{aligned} L^\epsilon(x, y, f) &= |y - f(x)|_\epsilon = \\ &= \max(0, |y - f(x)| - \epsilon). \end{aligned} \quad (4)$$

方程 (2) 的求解可通过对偶形式的 Lagrange 多项式, 将问题转化为

$$\begin{aligned} \max Q(\alpha^*, \alpha) &= \\ &= \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) - \epsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) - \\ &= \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) K(x_i, x_j). \end{aligned} \quad (5)$$

$$\begin{aligned} \text{s t } & \alpha_i^* = \alpha_i, C > \alpha_i^*, \alpha_i \geq 0, \\ & i = 1, 2, \dots, n. \end{aligned} \quad (6)$$

其中, α_i^* 和 α_i 为 Lagrange 乘子, $K(x_i, x_j)$ 为满足 Mercer 条件^[21] 的核函数, 本文采用 RBF 核函数

$$K(x, x_i) = \exp(-|x - x_i|^2 / \sigma^2).$$

二次规划 (5) 的解中, 由于 SVM 的稀疏性, 只有少量样本的系数 $(\alpha_i^* - \alpha_i)$ 不为零, 系数 $(\alpha_i^* - \alpha_i)$ 不为零所对应的样本 (x_i, y_i) 即为支持向量^[2,3]. 设支持向量个数为 g , 从而由支持向量决定 SVM 回归

$$f(x) = \sum_{i=1}^g (\alpha_i^* - \alpha_i) K(x, x_i) + b \quad (7)$$

其中

$$\begin{aligned} b &= \\ &= \frac{1}{2} \left\{ \min(y_i - \sum_{i=1}^g (\alpha_i^* - \alpha_i) K(x, x_i)) + \right. \\ & \left. \max(y_i - \sum_{i=1}^g (\alpha_i^* - \alpha_i) K(x, x_i)) \right\} \end{aligned} \quad (8)$$

3 FSVM 控制系统结构

FSVM 控制系统结构如图 1 所示 控制器为 2 个输入、1 个输出的网络结构, 控制器中间部分为 SVM, 外围部分为模糊化部分, 对变量空间进行模糊处理 与模糊神经网络不同, FSVM 的结构取决于二次规划方程 (5), 中间节点个数即为二次规划结果中的支持向量个数

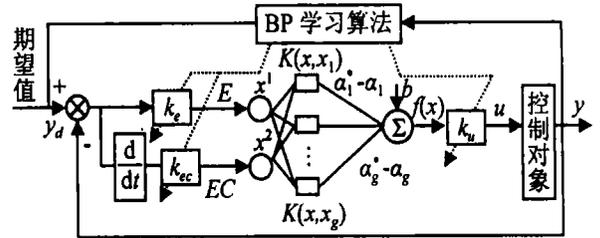


图 1 FSVM 的控制系统

利用 e 和 ec 表示控制误差和误差变化率, u 表示实际控制量, 并建立模糊集论域 $\{NB, NM, NS, Z, PS, PM, PB\}$, 对应论域的 7 个等级, 即 $\{-3, -2, -1, 0, 1, 2, 3\}$, 模糊化 e, ec 和 u , 对应的模糊变量为 E, EC 和 U , 比例因子 k_e, k_{ec} 和 k_u 根据不同的被控对象进行在线整定

控制决策过程如下:

1) 输入层 对变量 e 和 ec 进行模糊化: $E = k_e \cdot e, EC = k_{ec} \cdot ec$, 以此作为 FSVM 控制器的输入 $x = (x^1, x^2) = (E, EC)$.

2) 隐含层 二维输入 x 与支持向量进行核运算

$$K(x, x_i) = \exp(-|x - x_i|^2 / \sigma^2).$$

3) 输出层 SVM 回归

$$f(x) = \sum_{i=1}^g (\alpha_i^* - \alpha_i) K(x, x_i) + b,$$

再乘上比例因子 k_u , 即 $u = k_u \cdot f(x)$, u 为作用于被控对象的实际控制量

4 FSVM 学习算法

4.1 离线学习训练支持向量机

离线学习用于训练 SVM, SVM 的训练转换为二次规划方程 (5), 从而避免了神经网络过学习、局部极小点等缺陷 由 $\beta_i = \alpha_i^* - \alpha_i$ 且 $\alpha_i \alpha_i^* = 0$, 方程 (5) 等价于

$$\begin{aligned} \max Q(\beta) &= \\ &= \sum_{i=1}^n y_i \beta_i - \epsilon \sum_{i=1}^n |\beta_i| - \frac{1}{2} \sum_{i,j=1}^n \beta_i \beta_j K(x_i, x_j), \end{aligned} \quad (9)$$

$$\begin{aligned} \text{s t } & \sum_{i=1}^n \beta_i = 0, C > \beta_i \geq -C, \\ & i = 1, 2, \dots, n. \end{aligned} \quad (10)$$

这里采用序贯最小优化算法(SMO) 求解方程 (9). 序贯最小优化是将问题分解成一系列最小子集

的二次优化, 每次迭代只优化其中 2 个乘子 β_1 和 β_2 。选取 β_1 和 β_2 , 采用启发式方式, 计算 n 个样本的 Karush-Kuhn-Tucker (KKT) [8] 条件, 找到最违反 KKT 条件样本所对应的乘子 β_i 作为拟优化的乘子, β_1 和 β_2 取 $\max |E_1 - E_i|$ 样本所对应的乘子 β_i 。训练误差为

$$E_i = f(X_i) - y_i \quad (11)$$

由约束条件(10)可知, β_1 和 β_2 应在一条直线上, 即

$$\beta_1^{new} + \beta_2^{new} = \text{常数} = \beta_1^{old} + \beta_2^{old} \quad (12)$$

不失一般性, 首先计算 β_2^{new} , 而约束条件(10)为 β_2^{new} 确定了其约束空间, 即

$$U \leq \beta_2^{new} \leq V \quad (13)$$

这里

$$\begin{cases} U = \max(-C_U^2, \beta_1^{old} + \beta_2^{old} - C_V^1), \\ V = \min(C_V^2, \beta_1^{old} + \beta_2^{old} - C_U^1). \end{cases} \quad (14)$$

当 $\beta_i^{old} = 0$ 时, $C_U^i = 0, C_V^i = C$; 当 $\beta_i^{old} < 0$ 时, $C_U^i = -C, C_V^i = 0$ 设

$$k = K(X_1, X_1) + K(X_2, X_2) - 2K(X_1, X_2), \quad (15)$$

则有

$$\beta_2^{new, clipped} = \beta_2^{old} + \frac{(E_1 - E_2) - \epsilon(\text{sign } \beta_2 - \text{sign } \beta_1)}{k} \quad (16)$$

通过剪辑, 使得 β_2^{new} 满足约束方程(13), 即

$$\beta_2^{new} = \begin{cases} U, & \beta_2^{new, clipped} < U; \\ \beta_2^{new, clipped}, & \beta_2^{new, clipped} \in [L, H]; \\ V, & \beta_2^{new, clipped} > V. \end{cases} \quad (17)$$

这样, 便计算出了 β_2^{new} , 并可由方程(12)求得

$$\beta_1^{new} = \beta_1^{old} + \beta_2^{old} - \beta_2^{new} \quad (18)$$

此时, 序贯最小优化算法的一次迭代得到了满足约束条件的新的 β_1 和 β_2 。下一次迭代按照 KKT 条件重新寻取乘子, 优化计算, 如此循环迭代。当所有的乘子 β_i 均满足 KKT 条件时, 便得到了方程(9)和(5)的解, 从而确定了 SVM 的节点个数和参数取值

4.2 在线学习整定比例因子

对于不同的控制对象, 控制器比例因子 k_e, k_{ec} 和 k_u 需进行相应的整定, 本文采用 BP 算法在线整定比例因子。设 n 个样本, 参考输入 y_d , 系统输出 y , 采样周期 T , η 表示学习率, 记

$$e(t) = y_d(t) - y(t),$$

$$\omega = \{\omega, \omega, \omega\} = \{k_e, k_{ec}, k_u\}.$$

设定性能指标函数为

$$J(t) = \frac{1}{2} \sum_{p=1}^n e_p^2(t) = \frac{1}{2} \sum_{p=1}^n (y_{dp} - y_p)^2, \quad (19)$$

则参数修改表达式为

$$\begin{aligned} \omega(t+1) &= \omega(t) + \Delta\omega(t) = \\ &= \omega(t) - \eta \frac{\partial J(t)}{\partial \omega(t)}. \end{aligned} \quad (20)$$

对于 ω , 有

$$\frac{\partial J(t)}{\partial \omega(t)} = \frac{\partial J(t)}{\partial y(t)} \frac{\partial y(t)}{\partial u(t)} \frac{\partial u(t)}{\partial \omega} = -e(t) \frac{\partial y(t)}{\partial u(t)} f'(x, t). \quad (21)$$

对于 ω , 有

$$\begin{aligned} \frac{\partial J(t)}{\partial \omega(t)} &= \frac{\partial J(t)}{\partial y(t)} \frac{\partial y(t)}{\partial u(t)} \frac{\partial u(t)}{\partial \omega} \frac{\partial \omega}{\partial \omega} = \\ &= -e(t) \frac{\partial y(t)}{\partial u(t)} \omega(t) \left(\sum_{i=1}^g -((E(t) - x_i^1)^2 + (EC(t) - x_i^2)^2) / \sigma^2 \right) \alpha f'(x_i) (-2(EC(t) - x_i^2 / \sigma^2)) ((e(t-1) - e(t-2)) / T). \end{aligned} \quad (22)$$

对于 ω , 有

$$\begin{aligned} \frac{\partial J(t)}{\partial \omega(t)} &= \frac{\partial J(t)}{\partial y(t)} \frac{\partial y(t)}{\partial u(t)} \frac{\partial u(t)}{\partial \omega} \frac{\partial \omega}{\partial \omega} = \\ &= -e(t) \frac{\partial y(t)}{\partial u(t)} \omega(t) \left(\sum_{i=1}^g -((E(t) - x_i^1)^2 + (EC(t) - x_i^2)^2) / \sigma^2 \right) \alpha f'(x_i) \times (-2(E(t) - x_i^1) / \sigma^2) e(t-1). \end{aligned} \quad (23)$$

在式(21) ~ (23)中, 若被控对象模型已知, 则直接计算 $\partial y(t) / \partial u(t)$ 。若对象模型未知, 则由下面的方法近似求得:

$$\begin{aligned} \text{当 } u(t) = u(t-1) \text{ 时} \\ \frac{\partial y(t)}{\partial u(t)} = \frac{y(t) - y(t-1)}{u(t) - u(t-1)}, \end{aligned}$$

当 $u(t) \neq u(t-1)$ 时, 则取分母为 0.001T。

这样, 对于不同的被控对象, 由 BP 算法在线整定比例因子 k_e, k_{ec} 和 k_u , 提高了控制系统适应性和控制性能

5 仿真研究

例 1 (线性被控对象) 控制对象的传递函数为

$$\frac{s+1.9}{s^2+3s+2},$$

采样周期为 0.02 s, 图 2 为 FSVM 与模糊神经网络 PD 控制器的阶跃响应的对比。图中曲线, 和 分别表示 FSVM 控制器、模糊神经网络和 PD 控制器的响应曲线。仿真结果表明, FSVM 控制器比模糊神经网络 PD 控制器具有更优的控制性能, 其响应速度、超调量、过渡时间等性能指标均满足要求。

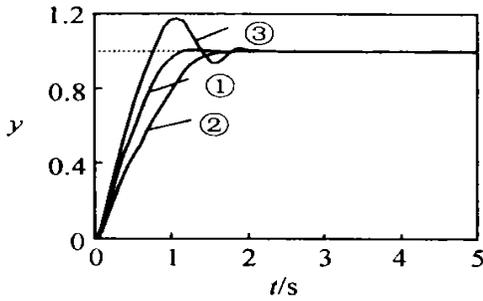


图2 线性系统的阶跃响应

例2(非线性系统被控对象) 采样周期为

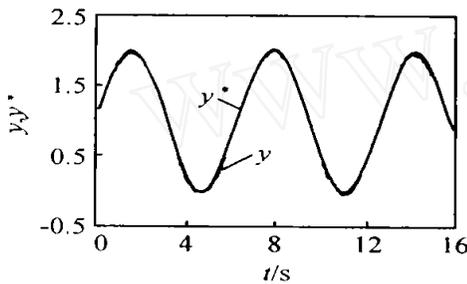
0.01 s, 控制对象为一个非线性系统

$$y(k+1) =$$

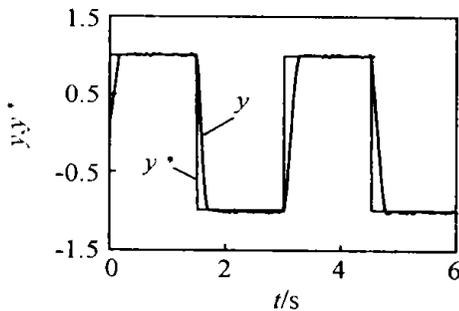
$$\frac{y(k)y(k-1)(0.4y(k)+1)}{1+y^2(k)+y^2(k-1)} + 0.4y(k) +$$

$$0.2y(k-1) + 0.8u(k) + 0.2u(k-1).$$

图3为跟踪响应仿真结果 其中图3(a)为正弦函数跟踪响应, 参考输入 $y^* = 1 + \sin t$, 系统输出为 y ;



(a) 正弦函数跟踪响应



(b) 正弦函数跟踪响应

图3 非线性系统的跟踪响应仿真

图3(b)为方波函数跟踪响应, 参考输入 y^* 为方波函数, 系统输出为 y .

6 结论

支持向量机比神经网络具有更强的理论依据和泛化能力, 它克服了模糊神经网络依赖经验设计结构、过学习、泛化能力不强等缺陷。本文的模糊支持向量机控制器, 结合了模糊控制和支持向量机的优点, 既具有模糊逻辑不依赖被控对象模型、鲁棒性强的特点, 又具有支持向量机小样本学习、泛化能力强等优点。对于线性、非线性系统, 所设计的控制器控制性能均较优。

参考文献(References)

[1] 王耀南. 智能控制系统[M]. 长沙: 湖南大学出版社, 1996

[2] Nello Cristianini, John Shawe-Taylor. An introduction to support vector machines and other kernel-based learning methods[M]. Cambridge: Cambridge University Press, 2000

[3] Vapnik V N. An overview of statistical learning theory [J]. IEEE Trans on Neural Networks, 1999, 10(5): 988-999

[4] Christopher J C Burges. A tutorial on support vector machines for pattern recognition [J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167.

[5] Chan W C, Chan C W, Cheung K C, et al. On the modelling of nonlinear dynamic system using support vector neural networks[J]. Engineering Applications of Artificial Intelligence, 2001, 14(2): 105-113

[6] Suykens J A K, Vandewalle J, Moor B De. Optimal control by least squares support machines[J]. Neural Networks, 2001, 14(1): 23-35

[7] Platt J. Sequential minimal optimization: A fast algorithm for training support vector machine [R]. Washington: Microsoft Research, 1998

[8] LeCun Y A, Jackel L D, Bottou L, et al. Learning algorithms for classification: A comparison on handwritten digit recognition [A]. Neural Networks: The Statistical Mechanics Perspective [C]. Singapore, 1995: 261-276

(上接第536页)

[12] 冯蔚东, 陈剑. 虚拟企业中核心能力的定性定量识别 [J]. 系统工程理论与实践, 2002, 16(5): 48-54
(Fen W D, Chen J. Identifying the core competence for virtual enterprises: A qualitative framework and a quantitative method [J]. Systems Engineering — Theory and Practice, 2002, 22(5): 48-54)

[13] 王毅, 陈劲. 企业核心能力高标定位研究 [J]. 管理工程学报, 2002, 16(4): 22-27.
(Wang Y, Chen J. Benchmarking core competence in China's firms [J]. J of Industrial Engineering/Engineering Management, 2002, 16(4): 22-27.)