

LM 滚动优化的 SVMR 预测控制研究

王定成^{1,2} 宋良图²¹(南京信息工程大学 计算机科学技术系 南京 210024)²(中国科学院合肥智能机械研究所 合肥 230031)

摘 要 支持向量机回归建模方法简单,有完备的理论基础,所反映的直接是系统的非线性特征,因而与神经网络等非线性回归方法相比具有独特的优点.因此,本文提出的非线性预测控制结构采用支持向量机回归建立非线性系统的预测模型,采用 LM (Levenberg-Marquardt)非线性最小二乘优化方法作为非线性滚动优化方法.实验证明了这种预测控制具有较好的非线性控制效果,而且能够较好地抵消不可测干扰.

关键词 支持向量机回归,预测控制,非线性

中图分类号 TP273

Research on SVMR Predictive Control Based on LM Rolling Optimization

WANG Ding-Cheng^{1,2}, SONG Liang-Tu²¹(Department of Computer Science and Technology, Nanjing University of
Information Science and Technology, Nanjing 210024)²(The Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031)

ABSTRACT

Supporting by mathematics theory, Support Vector Machines Regression (SVMR) has a simple structure and nonlinear modeling properties, therefore, SVMR shows better performance compared with other nonlinear regression, such as neural networks. As a result, the predictive control, which models the nonlinear system process using SVMR and optimizes using Levenberg Marquardt (LM), is presented. The simulation experiment shows that the control algorithm has a good result for nonlinear system and rejection capacities for unknown disturbance.

Key Words Support Vector Machines Regression, Predictive Control, Nonlinear

1 引 言

预测控制的机理是预测模型、滚动优化和反馈校正.应用模型进行预测是预测控制的基本特征,模

型是否能够反映系统的特征直接关系到控制的精度和稳定性^[1,2],因此研究高精度、有效、简单的信息预测模型是目前非线性预测控制要解决的重要问题.

收稿日期:2003-09-18;修回日期:2004-06-30

作者简介 王定成,男,1967年生,博士,副研究员,主要研究方向为机器学习、智能控制理论与应用等. E-mail: dcwang@iim.ac.cn. 宋良图,男,1963年生,副研究员,博士研究生,主要研究方向为模式识别、信息农业等理论与应用.

神经网络具有较好处理非线性的特性,因而使神经网络刚刚进入控制就出现了基于神经网络的预测控制^[3].但基于经验风险最小化的神经网络的局部极小点、过学习以及结构和类型选择过分依赖经验等固有缺陷严重降低了其应用效果.基于结构风险最小化的 SVMR 直接反映的是非线性模型的特征,克服了传统的非线性建模的近似线性或类似线性系统的设计方法所引起的模型精确度差的缺点^[4].不仅如此,基于结构风险最小化的支持向量机回归较成功地克服了其他学习理论局部极小点、过学习等缺陷^[5].因而,采用支持向量机回归理论设计的预测控制能简化建模手段、提高预测精度、改善系统的性能.

2 支持向量机回归的原理

2.1 支持向量机

支持向量机最初是用来解决模式识别的问题.在模式识别中,为了发现具有好的推广能力的决策规则,所选择的训练数据的一些子集,称作支持向量.最佳的支持向量分离等效于所有数据的分离.

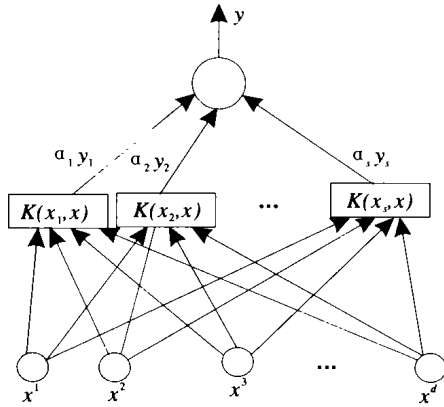


图 1 支持向量机结构示意图
Fig.1 The chart of structure of SVM

支持向量机是从线性可分情况下的最优分类面发展而来的.其基本思想可参见文献[5-8].支持向量机形式上类似一个神经网络,输出是中间节点的线性组合,每个中间节点对应一个支持向量.其结构示意图如图 1 所示.图中的 \$x^1, x^2, \dots, x^d\$ 为输入样本向量; \$K(x_1, x), \dots, K(x_s, x)\$ 为支持向量机的核; \$(x_1, y_1), (x_2, y_2), \dots, (x_s, y_s)\$ 为支持向量; \$y\$ 为支持向量机的输出.

2.2 支持向量机回归原理

支持向量机回归的基本思想是通过一个非线性

映射 \$\Phi\$ 将数据 \$x\$ 映射到高维特征空间 \$F\$,并在这个空间进行线性回归.如式(1)所示:

$$f(x) = (w \cdot \Phi(x)) + b \quad (\Phi: R^n \rightarrow F, w \in F), \quad (1)$$

其中 \$b\$ 是阈值.这样,在高维特征空间的线性回归就对应于低维输入空间的非线性回归,免去了在高维空间 \$w\$ 和 \$\Phi(x)\$ 点积的计算.由于 \$\Phi\$ 是固定不变的,因此影响 \$w\$ 的有经验风险的总和 \$R_{emp}\$ 以及使其在高维空间平坦的 \$\|w\|^2\$,则有

$$R(w) = R_{emp} + \lambda \|w\|^2 = \sum_{i=1}^l e(f(x_i) - y_i) + \lambda \|w\|^2, \quad (2)$$

其中, \$l\$ 表示样本的数目, \$e(\cdot)\$ 是损失函数, \$\lambda\$ 是调整的常数.最小化 \$R(w)\$ 即得用数据点表示的 \$w\$,即

$$w = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \Phi(x_i), \quad (3)$$

其中 \$\alpha_i, \alpha_i^*\$ 是最小化 \$R(w)\$ 的解.考虑方程(1)和(3), \$f(x)\$ 可表示为

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) (\Phi(x_i) \cdot \Phi(x)) + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b, \quad (4)$$

其中 \$K(x_i, x) = (\Phi(x_i) \cdot \Phi(x))\$ 称为核函数. \$K(x_i, x)\$ 是满足 Mercer 条件的任何对称的核函数对应于特征空间的点积^[8].核函数的种类很多,其中有^[6,9]多项式函数 \$K(x_i, x) = [(x, x_i) + 1]^q\$, RBF 函数 \$K(x_i, x) = \exp\{-|x - x_i|^2 / 2\sigma^2\}\$, Sigmoid 函数 \$K(x_i, x) = \tanh(v(x \cdot x_i) + c)\$.

根据文献[5],[6]和[9],式(2)中的损失函数 \$e(\cdot)\$ 有

(1) 线性 \$\epsilon\$ 不敏感损失函数

$$e(f(x) - y) = \max(0, |f(x) - y| - \epsilon). \quad (5)$$

(2) 二次损失函数

$$e(f(x) - y) = (f(x) - y)^2. \quad (6)$$

(3) Huber 损失函数

$$e(f(x) - y) = \begin{cases} \epsilon |f(x) - y| - \epsilon^2 / 2, & \text{若 } |f(x) - y| > \epsilon \\ \frac{1}{2} |f(x) - y|^2, & \text{其他} \end{cases} \quad (7)$$

在约束条件

$$\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0, \alpha_i, \alpha_i^* \in [0, C]$$

下求最小化 \$R(w)\$,即得(4)式中的 \$\alpha_i - \alpha_i^*\$.

(4)式中的 \$b\$ 取在边界上的一点就可以计算,但为了稳定性的目的,采用边界点上的平均值来计算^[9],即

$$b = \text{average}_k \{ \delta_k + y_k - \sum_{i=1}^l (\alpha_i - \alpha_i^* k(x_i, x_k)) \}, \quad (8)$$

其中 δ_k 为预测误差. 对于 ϵ 不敏感损失函数来说, $\delta_k = \epsilon \text{sign}(\alpha_k - \alpha_k^*)$; 对于 Huber 损失函数来说, $\delta_k = (1/C)(\alpha_k - \alpha_k^*)$.

3 预测控制的原理

传统的预测控制如图 2 所示, 图中 $y_r(k+1)$ 为参考设定值, 最小化 $J(k)$ 为优化的目标函数, $u(k)$ 为控制器输出, $v(k)$ 为干扰, $y_M(k+i)$ 为模型输出, $y_p(k+i)$ 为预测输出, $y(k+1)$ 为系统的输出. 模型输出一般都建立在预测模型、滚动优化和反馈校正的原理之上.

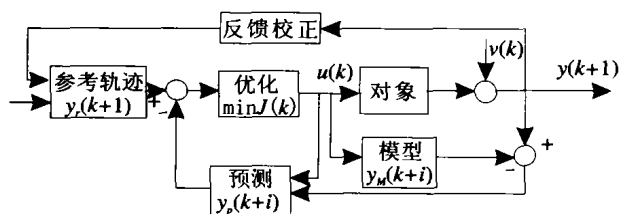


图 2 预测控制的基本结构图

Fig. 2 The structure of model predictive control

预测模型的功能是根据对象的历史信息和未来输入预测其未来输出, 这里只强调模型的功能而不强调其结构形式. 滚动优化是指, 优化不是一次离线进行而是反复在线进行的. 反馈校正可以保持预测模型不变的基础上, 对未来的误差作出预测并加以补偿.

4 支持向量机预测控制

支持向量机预测控制采用最新的支持向量机回归的方法建立非线性过程模型来进行模型预测, 采用非线性最小二乘优化算法 Levenberg - Marquardt(LM) 来进行滚动优化.

4.1 SVM 回归进行模型预测

考虑一个单输入单输出非线性模型

$$y(k+1) = f(y(k), \dots, y(k-n), u(k), \dots, u(k-m)), \quad y \in R, u \in R, m \leq n, \quad (9)$$

其中, u 和 y 分别代表对象的控制输入和系统输出. 设有连续的控制输入

$$u(k-m), u(k-m+1), \dots, u(k)$$

和输出

$$y(k-n), y(k-n+1), \dots, y(k),$$

其中, 当 $k-m+r < 0$ 时,

$$u(k-m+r) = u(0), r = 0, 1, 2, \dots, m,$$

$$y(k-m+r) = y(0), r = 0, 1, 2, \dots, n,$$

则相应的输出为 $y(k+1)$.

设

$$U(i) = y(i), y(i-1), \dots, y(i-n), u(i), u(i-1), \dots, u(i-m)),$$

$$i = 1, 2, \dots, n,$$

则

$$y(k+1) = f(U(i)). \quad (10)$$

构造学习样本集 $(U(i), y(i+1))$, 采用支持向量机回归的方法就可以将非线性的样本数据映射为高维空间的线性输出, 即

$$\hat{y}(k+1) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(U(i), U(k)) + b, \quad (11)$$

$U(i)$ 为支持向量, 对应于 Vapnik 的 ϵ 不敏感损失函数的 $\beta = \alpha_i - \alpha_i^*$ 不为 0 的向量. $U(k)$ 为第 k 个支持向量的输入. β 为最小化式 (2) 的结果, b 由式 (8) 计算得到.

令 $|y(k+1) - \hat{y}(k+1)| = V$, 其中 V 为模型预测误差, 支持向量机回归模型的损失函数可以采用 ϵ 不敏感损失函数作为损失函数, 这是因为 ϵ 不敏感损失函数具有较好的稀疏性^[9-10]; 也可以采用最小二乘损失函数, 其稀疏性的解采用文献[11]介绍的方法来获得. 对于非线性模型的核函数一般采用 RBF 核函数

$$K(U(i), U(k)) = \exp(-|U(k) - U(i)|^2 / 2\sigma^2),$$

其中, 中心为支持向量 $U(i)$, 宽度 σ 是需要预先指定的一个常量.

模型可以通过离线学习也可以在线学习. 已经训练好的支持向量机回归模型可以进行在线模型修正. 由于在 ϵ 不敏感区域内没有支持向量^[11], 因此可以通过比较系统的实际输出 $y(k+1)$ 和估计的输出 $\hat{y}(k+1)$, 若 $|y(k+1) - \hat{y}(k+1)| < \epsilon$, 仍然采用原来的模型; 否则, 将该时刻的输入输出数据与原来的支持向量的点一起重新进行优化, 建立新的过程模型. 如此反复, 进行在线的模型修正.

4.2 预测控制滚动优化控制

支持向量机滚动优化控制规律采用非线性最小二乘优化控制方法, 其目标函数采用 N 步优化控制规律, 可以表示为如下形式:

$$\min J = \sum_{i=1}^p \alpha_i [y_r(k+i) - y_p(k+i)]^2 + \beta_j \sum_{j=1}^c [u(k+j-1) - u(k+j-2)]^2, \quad (12)$$

优化方法采用 Levenberg - Marquardt 优化方法,这是由于该方法在大量的非线性优化问题中得到较成功的应用,并被证明比 Gauss - Newton 法具有更好的鲁棒性,比无限定条件方法具有更好的迭代效率.

4.3 反馈校正

由于实际存在模型失配、环境干扰等未知因素,由式(11)给出的预测值有可能偏离实际值,如果不及时利用实时信息进行反馈校正,进一步的优化就会建立在虚假的基础上,为此必须引入反馈校正.本文采用的校正方法为

$$y_p(k+1) = \hat{y}(k+1) + he(k), \quad (13)$$

其中, h 为补偿系数,根据实际应用的效果进行调整; $e(k)$ 为系统在 k 时刻系统实际输出与模型预测之间的误差

$$e(k) = y(k) - \hat{y}(k). \quad (14)$$

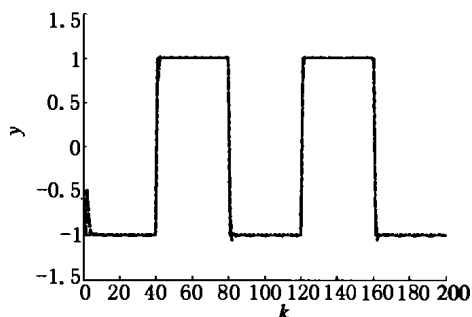
5 SVMR 预测控制的实验与分析

5.1 没有噪声的非线性系统的控制

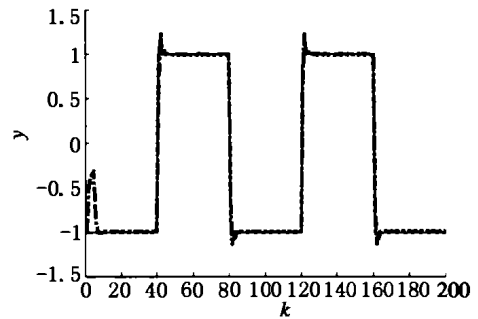
设有非线性系统

$$y(k+1) = \frac{y(k)}{1+y^2(k)} + u^3(k) + v, \quad (15)$$

采用 SVMR 预测控制和神经网络预测控制来跟踪方波信号.取 SVM 回归的损失函数为 ϵ 不敏感损失函数,核函数为 RBF 核函数,核函数的宽度 $\sigma = 10$,误差 $\epsilon = 0.01$,样本数据为 $u_k \in (-2,2)$ 的 100 个随机样本数据,仿真实验结果如图 3(a) 所示.神经网络采用 2-20-1 的 RBF 神经网络,样本数据为 $u_k \in (-2,2)$ 的 6 000 个随机样本,仿真结果如图 3(b) 所示.图中实线为被跟踪的方波信号,方点线为内部模型输出信号,短划线为系统的输出信号,三个曲线基本重合.从图中可以看出在没有噪声干扰的情况下,SVMR 预测控制的控制效果优于神经网络预测控制的效果.

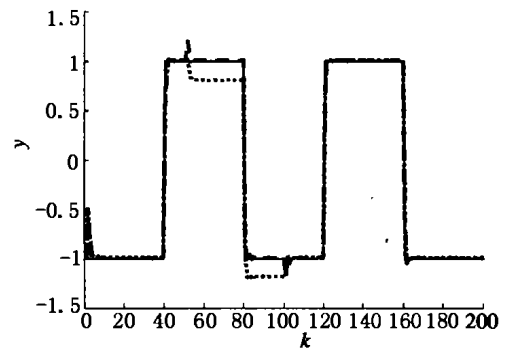


(a)SVMR 预测控制的方波参考输入响应
(a)SVMR - LM - MPC's response of square wave

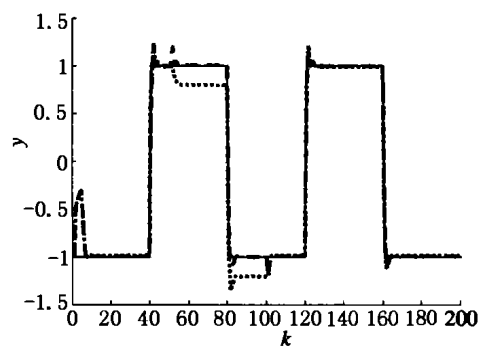


(b)神经网络预测控制的方波参考输入响应
(b)NN - LM - MPC's response of square wave
图 3 无干扰的方波参考输入响应

Fig. 3 SVMR - LM - MPC's response of square wave without noise



(a)SVMR 预测控制的方波参考输入响应
(a)VMR - LM - MPC's response of square wave



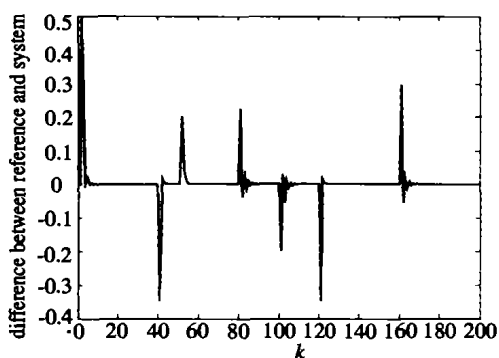
(b)神经网络预测控制的方波参考输入响应
(b)NN - LM - MPC's response of square wave
图 4 有干扰的预测控制的方波参考输入响应

Fig. 4 SVMR - LM - MPC's response of square wave with noise

5.2 有噪声的非线性系统的控制

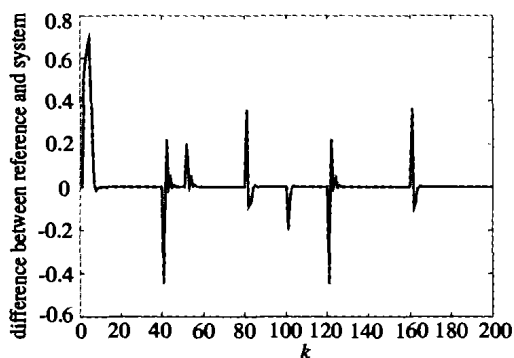
系统模型采用式(15)所示的模型,系统的干扰为随机阶跃干扰.仿真实验结果如图 4 所示.在图 4 中实线为被响应的方波参考信号,短划线为系统的输出信号,方点线为内部模型的输出.仿真中取系统的初值与跟踪的初值相同,出现了系统输出 $y(2)$ 和

模型输出 $\hat{y}(2)$ 偏离跟踪参考值,但随着控制系统反馈和滚动优化的输出,系统输出 $y(3)$ 和模型输出 $\hat{y}(3)$ 就较好地跟踪参考值.在 $k = 50$ 至 $k = 100$ 出现一个随机阶跃干扰,在 $k = 50$ 处出现了偏离跟踪参考值,但在下一个采样周期,抑制了干扰的影响,表现出了较好的鲁棒性.神经网络预测控制虽然也能够较好地抑制干扰,但从控制系统的效果来看,与系统实际输出的误差较大,而且动态性能也没有 SVMR 预测控制的效果好.



(a)SVMR 预测控制的误差

(a)SVMR-LM-MPC's difference between the reference and the system



(b)神经网络预测控制的误差

(b)NN-LM-MPC's difference between the reference and the system

图 5 有干扰的预测控制参考轨迹与系统输出误差

Fig. 5 SVMR-LM-MPC's difference between the reference and the system output

6 结束语

支持向量机回归将低维非线性的输入映射到高维线性的输出,模型简单,有完备的理论支持,具有较好的非线性建模特性,而预测控制具有较好的鲁棒性和稳定性.用 SVMR 进行预测控制的研究将二者较好的结合起来,是一种新的尝试.这种控制方式

不需要了解对象太多的先验知识,只要研究对象的输入输出数据就可以对非线性系统进行控制.仿真实验说明这种方法效果较好,是预测控制中的一种重要的新方法.

参 考 文 献

- [1] Xu L. H. Predictive Control Current State and Problem. Control Theory and Applications, 1994, 11(1): 121-125 (in Chinese) (徐立鸿. 预测控制的研究现状及问题. 控制理论及应用, 1994, 11(1): 121-125)
- [2] Xi Y. G. Predictive Control. Beijing, China: National Defence Industry Press, 1993 (in Chinese) (席裕庚. 预测控制. 北京: 国防工业出版社, 1993)
- [3] Draeger A, Engell S, Ranke H. Model Predictive Control Using Neural Networks. IEEE Control Systems Magazine, 1995, 15(15): 61-66
- [4] Wang D C, Fang T J, Tang Y, et al. Review of Support Vector Machines Regression Theory and Control. Pattern Recognition and Artificial Intelligence, 2003, 16(2): 192-196 (in Chinese) (王定成, 方廷健, 唐毅, 等. 支持向量机回归与控制的综述. 模式识别与人工智能, 2003, 16(2): 192-196)
- [5] Müller K-R, Smola A J, Rätsch G, et al. Predicting Time Series with Support Vector Machines. In: Gerstner W, Germond A, Hasler M. et al. eds. Proc of the International Conference on Artificial Neural Networks. New York, USA: Springer, 1997, 999-1004
- [6] Drucker H, Burges C J C, Kaufman L, et al. Support Vector Regression Machines. In: Mozer M, Jordan M, Petsche T, eds. Advances in Neural Information Processing Systems 9. Cambridge, USA: Massachusetts Institute of Technology Press, 1997, 155-161
- [7] Vapnik V, Golowich S, Smola A. Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing. In: Mozer M, Jordan M, Petsche T, eds. Advances in Neural Information Processing Systems 9. Cambridge, USA: Massachusetts Institute of Technology Press, 1997, 281-287
- [8] Boser B E, Guyon I M, Vapnik V. A Training Algorithm for Optimal Margin Classifiers. In: Haussler D, ed. Proc of the 5th ACM Workshop on Computational Learning Theory. Pittsburgh, USA, 1992, 144-152
- [9] Vapnik V. The Nature of Statistical Learning Theory. New York, USA: Springer, 1999
- [10] Gunn S. Support Vector Machines Classification and Regression. Technical Report, Image Speech & Intelligent Systems Group, University of Southampton, Southampton, UK, 1998
- [11] Suykens J A K, Vandewalle J. Recurrent Least Squares Support Vector Machines. IEEE Trans on Circuits and Systems, 2000, 47(7): 1109-1114