

文章编号: 1001-0920(2004)04-0383-05

基于 LS-SVM 的非线性预测控制技术

王宇红^{1,2}, 黄德先¹, 高东杰², 金以慧¹

(1 清华大学 自动化系, 北京 100084; 2 中国科学院 自动化研究所, 北京 100080)

摘 要: 探讨了利用最小二乘支持向量机(LS-SVM)进行非线性系统辨识的方法, LS-SVM 用等式约束代替传统支持向量机中不等式约束, 求解过程从解 QP 问题变成解一组等式方程. 将得到的 LS-SVM 模型应用到非线性预测控制, 提出了基于 LS-SVM 模型的非线性预测控制算法. 通过 CSTR 过程仿真表明, 最小二乘支持向量机学习速度快, 在小样本情况下具有良好的非线性建模和泛化能力. 基于 LS-SVM 的预测控制算法具有很好的控制性能.

关键词: 最小二乘支持向量机; 非线性建模; 预测控制; 非线性控制

中图分类号: TP273

文献标识码: A

Nonlinear predictive control based on LS-SVM

WANG Yu-hong^{1,2}, HUANG De-xian¹, GAO Dong-jie², JIN Yi-hui¹

(1 Department of Automation, Tsinghua University, Beijing 100084, China; 2 Institute of Automation, Chinese Academy of Science, Beijing 100080, China Correspondent: WANG Yu-hong, E-mail: yuhongwang@mail.tsinghua.edu.cn)

Abstract: An identification method of nonlinear systems using least squares support vector machine (LS-SVM) is proposed. The constraints of inequalities in the classical SVM approach are replaced by equality-type constraints in LS-SVM. The LS-SVM solution follows directly from solving a set of linear equations instead of quadratic programming. A kind of nonlinear predictive control scheme based on the LS-SVM model is presented. Simulation results for a CSTR process show that LS-SVM can be trained fastly. The LS-SVM has good ability of modeling nonlinear process and good generalization under small data set available. The nonlinear predictive control strategy based on LS-SVM model shows satisfactory performance.

Key words: LS-SVM; nonlinear modeling; predictive control; nonlinear control

1 引 言

预测控制技术是当今工业界应用最为成功的先进控制技术之一. 目前常用的是线性模型预测控制技术, 但对具有高度非线性过程, 用线性模型预测控制技术难以获得好的控制性能. 由于生产规模的扩大和复杂程度的提高, 使得建立系统的机理模型变得越来越困难. 这导致了通过非线性微分方程模型将线性模型预测控制技术扩展到非线性过程的方法^[1~3]受到极大的限制.

对于难以进行机理建模的复杂过程, 利用实际可测量的输入输出数据来辨识模型成为必然选择. 神经网络被证明是建立非线性过程模型的有效手段. 它通过一个连接模型将输入空间映射到输出空间, 这个模型能以任意精度逼近非线性函数. 因此, 神经网络在很大程度上解决了控制模型难以建立的问题, 基于神经网络的非线性预测控制技术也得到了相应的发展. 但利用神经网络建模还存在许多待解决的问题, 特别是神经网络的训练问题和

收稿日期: 2003-04-11; 修回日期: 2003-06-12

基金项目: 国家 863 计划项目(AA 413130).

作者简介: 王宇红(1970—), 男, 河北新乐人, 博士, 从事混杂系统控制、预测控制等研究; 黄德先(1958—), 男, 河南潢川人, 教授, 博士, 从事系统辨识、过程控制等研究.

局部最小值问题以及神经网络拓扑结构的确定

近年来成为研究热点的统计学习理论(SLT)是一种专门研究小样本情况下机器学习规律的理论^[4,5].统计学习理论为解决有限样本学习问题提供了一个统一的框架,在这一理论框架基础上发展了一种新的通用学习方法——支持向量机(SVM).支持向量机根据有限的样本信息,在模型的复杂性和学习能力之间寻求最佳折衷,以期获得最好的泛化能力.与传统的神经网络相比,支持向量机算法可转化为一个二次型寻优问题,从理论上说,得到的将是全局最优点,解决了在神经网络方法中无法避免的局部极值问题;支持向量机拓扑结构由支持向量决定,避免了传统神经网络拓扑结构需要经验试凑的方法.支持向量机也能以任意的精度逼近任意函数.

与经典支持向量机相比,最小二乘支持向量机(LS-SVM)^[6]用等式约束代替不等式约束,求解过程变成了解一组等式方程,避免了求解耗时的QP问题,求解速度相对加快.相对于常用的 ϵ -不敏感损失函数^[5],LS-SVM不再需要指定逼近精度 ϵ .

非线性预测控制技术得以实现的最重要的因素是实用的非线性预测模型.本文利用LS-SVM建立非线性系统的预测模型,并提出了基于LS-SVM的非线性预测控制算法.通过一个CSTR仿真例子验证了该非线性预测控制方法的有效性.

2 支持向量机原理

2.1 Vapnik 支持向量机建模原理

根据传统统计学,在非线性系统的黑箱建模过程中,一般是给定训练样本集 $\{x_j, y_j\}, j = 1, \dots, l$,寻找非线性关系 f 以使得下面经验风险最小:

$$R_{\text{emp}}(w) = \frac{1}{l} \sum_{j=1}^l L(y_j, f(x_j; w)), \quad (1)$$

其中: x 与 y 为对应的输入输出向量, $\{f(x, w)\}$ 是预测函数集, w 为广义参数, L 为损失函数.

统计学习理论指出^[5],根据有限样本进行非线性估计的真实风险用经验风险代替是不合理的,还应考虑到泛化风险,这样在小样本条件下估计才能取得较好的效果.这种思想称为结构风险最小化(SRM).

支持向量机是SRM准则的一种具体实现.在经典支持向量机学习中,取损失函数为^[5]

$$\begin{cases} |y - f(x; w)|_{\epsilon} = \\ 0, & |y - f(x; w)| < \epsilon \\ |y - f(x; w)| - \epsilon, & \text{otherwise} \end{cases} \quad (2)$$

其中 $\epsilon \geq 0$,表示逼近精度.该函数称为 ϵ -不敏感函

数,取非线性估计函数为

$$f(x) = \langle w, \Phi(x) \rangle + b = \sum_{j=1}^l w_j \phi_j(x) + b, \quad (3)$$

其中: ϕ 为非线性函数, w 为权值.

为了使上式的估计真实风险最小,利用SRM准则构造下面的最小化目标函数:

$$\begin{aligned} \min & \frac{1}{2} \|w\|^2 + C \sum_{j=1}^l (\xi_j + \xi_j^*), \\ \text{s.t.} & y_i - f(x_i) \leq \epsilon + \xi_i, \quad j = 1, \dots, l, \\ & -y_i + f(x_i) \leq \epsilon + \xi_i^*, \quad j = 1, \dots, l, \\ & \xi_j \geq 0, \xi_j^* \geq 0, \quad j = 1, \dots, l \end{aligned} \quad (4)$$

其中: C 是容量因子, ξ 和 ξ^* 为松弛变量.

优化问题(4)可以转化为其对偶问题求解,即

$$\begin{aligned} \max & -\epsilon \sum_{j=1}^l (\alpha_j + \alpha_j^*) + \sum_{j=1}^l y_j (\alpha_j - \alpha_j^*) - \\ & \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle \Phi(x_i), \Phi(x_j) \rangle, \\ \text{s.t.} & \sum_{j=1}^l (\alpha_j - \alpha_j^*) = 0, \\ & \alpha_j, \alpha_j^* \in [0, C], \quad j = 1, \dots, l \end{aligned} \quad (5)$$

α_j 为Lagrange乘子.

由式(5)计算出Lagrange乘子 α_j ,并得到权值 w 后,相应的非线性估计关系为

$$\begin{aligned} f(x) &= \langle w, \Phi(x) \rangle + b = \\ & \sum_{j=1}^l (\alpha_j - \alpha_j^*) \langle \Phi(x_j), \Phi(x) \rangle + b \end{aligned} \quad (6)$$

为了避免在高维空间计算内积 $\langle w, \Phi(x) \rangle$ 而引起维数灾难问题,利用再生核的概念将高维空间内积的计算转为在输入空间的计算.

定理1^[4] 如果 K 为正定对称函数,且满足Mercer条件,则特征空间内积与输入空间有相同的核,即

$$K(x, y) = K(x)K(y). \quad (7)$$

因此非线性估计关系可以改写为

$$\begin{aligned} f(x) &= \sum_{j=1}^l (\alpha_j - \alpha_j^*) K(x_j, x) + b, \\ b &= \sum_{s,r=1}^l (\alpha_r - \alpha_r^*) [K(x_r, x_j) + K(x_s, x_j)], \end{aligned} \quad (8)$$

x_r 和 x_s 为两个任意支持向量.

2.2 最小二乘支持向量机(LS-SVM)建模原理

与Vapnik支持向量机不同,利用SRM准则构造下面的最小化目标函数^[6]:

$$\min_{w, b, e} J = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^l e_k^2, \\ \text{s.t. } y_k = w^T \varphi(x_k) + b + e_k, k = 1, \dots, l \quad (9)$$

定义Lagrange 函数为

$$L = J - \sum_{k=1}^l a_k [w^T \varphi(x_k) + b + e_k - y_k], \quad (10)$$

其中 a_k 为Lagrange 乘子. 根据KKT 条件:

$$\begin{aligned} \partial L / \partial w &= 0 \rightarrow w = \sum_{k=1}^l a_k \varphi(x_k), \\ \partial L / \partial b &= 0 \rightarrow \sum_{k=1}^l a_k = 0, \\ \partial L / \partial e_k &= 0 \rightarrow a_k = -e_k, \\ \partial L / \partial a_k &= 0 \rightarrow w^T \varphi(x_k) + b + e_k - y_k = 0, \end{aligned} \quad (11)$$

对于 $k = 1, \dots, l$, 消去 w 和 e , 得到如下方程:

$$\begin{bmatrix} 0 & \rho_1^T \\ \rho_1 & ZZ^T + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix}, \quad (12)$$

其中

$$\rho_1 = [1 \dots 1]^T, Z = \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_l)^T \end{bmatrix}, \quad (13)$$

$$Y = [y_1 \dots y_l]^T, a = [a_1 \dots a_l]^T.$$

用最小二乘法求出 a 和 b , 由此得到预测输出

$$y(x) = \sum_{k=1}^l a_k \varphi(x)^T \varphi(x_k) + b \quad (14)$$

利用核函数的方法, 令

$$K(x, x_k) = \varphi(x)^T \varphi(x_k), \quad (15)$$

则预测输出为

$$y(x) = \sum_{k=1}^l a_k K(x, x_k) + b \quad (16)$$

3 聚合过程的LS-SVM 模型

本文用支持向量机来建立一个聚合过程^[3]的

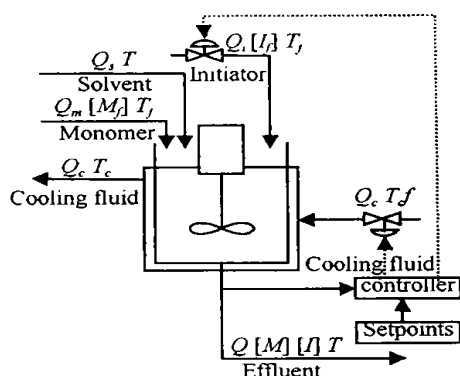


图1 聚合流程图

模型, 该聚合过程如图1所示. 控制的目标是通过调整反应中止剂进料流量 Q_i 来控制聚合产物的分子量NAMW. 该聚合过程模型如下:

$$\begin{aligned} \frac{dC_m}{dt} &= -(k_p + k_{f_m})C_m P_0 + \frac{F(C_{m_{in}} - C_m)}{V}, \\ \frac{dC_I}{dt} &= -k_I C_I + \frac{Q_i C_{I_{in}} - F C_I}{V}, \\ \frac{dD_0}{dt} &= (0.5k_{T_c} + k_{T_d})P_0^2 + k_{f_m}C_m P_0 - \frac{F D_0}{V}, \\ \frac{dD_1}{dt} &= M_m(k_p + k_{f_m})C_m P_0 - \frac{F D_1}{V}, \\ u &= Q_i, y = D_1/D_0 \end{aligned} \quad (17)$$

其中

$$P_0 = \left[\frac{2f k_I C_I}{k_{T_d} + k_{T_c}} \right]^{0.5}.$$

稳态操作条件为

$$\begin{aligned} x_1 &= C_m = 5.506774 \text{ kmol/m}^3, \\ x_2 &= C_I = 0.132906 \text{ kmol/m}^3, \\ x_3 &= D_0 = 0.0019752 \text{ kmol/m}^3, \\ x_4 &= D_1 = 49.38182 \text{ kmol/m}^3, \\ u &= Q_i = 0.016783 \text{ m}^3/\text{h}, \\ y &= D_1/D_0 = 25.0005 \text{ kg/kmol} \end{aligned}$$

利用四阶Runge-Kutta 法则对模型进行仿真, 产生长度为200步的两批数据, 第一批数据没有噪声, 第二批数据加入噪信比为24%的正态白噪声, 采样时间为0.03 h. 为了体现支持向量机的小样本学习能力, 每批数据仅用前100步作为训练集, 后100步作为测试集. 用RMSE来评价模型的估计效果, 即

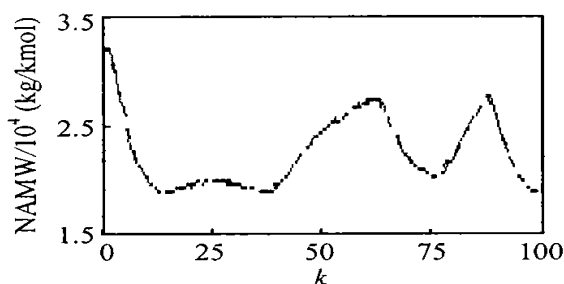
$$\text{RMSE} = \sqrt{\frac{1}{l} \sum_{j=1}^l (y_j^{\text{predicted}} - y_j^{\text{actual}})^2}. \quad (18)$$

支持向量机建模时取RBF核函数 $K(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$.

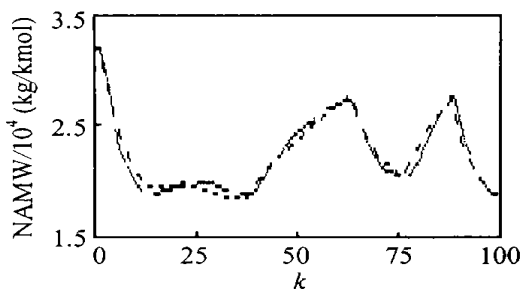
考虑到预测控制需要利用模型在线进行长期预测, 同时考察该模型的长期预测能力(即给定初态), 考察支持向量机模型的100步预测能力.

3.1 无噪声情况下辨识结果

利用第一批数据前100步训练网络, 采用Cross-validation 方法在无噪声时取容量控制因子 $C = 80$, 核函数参数 $\sigma = 140$. 利用第一批测试集对辨识得到的模型进行测试, 模型一步预测结果和长期预测结果如图2所示. 图中点划线为模型预测值, 短划线为输出测试值.



(a) 一步预测结果(KM S = 251)

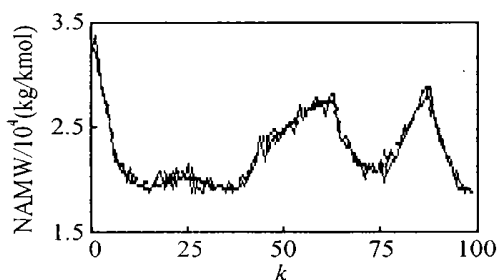


(b) 长期预测结果(RM S = 610)

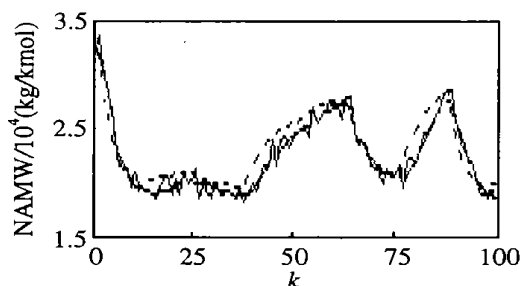
图2 无噪声情况下的预测结果

3.2 有噪声情况下辨识结果

利用第二批数据前100步训练网络,采用Cross-Validation方法在有噪声时取容量控制因子 $C = 88$,核函数参数 $\sigma = 180$ 。利用第二批测试集对辨识得到的模型进行测试,模型一步预测结果和长期预测结果如图3所示。图中点划线为模型预测值,短划线为不含噪声的输出测试值,实线为含有噪声的输出测试值。



(a) 一步预测结果(RM S = 659)



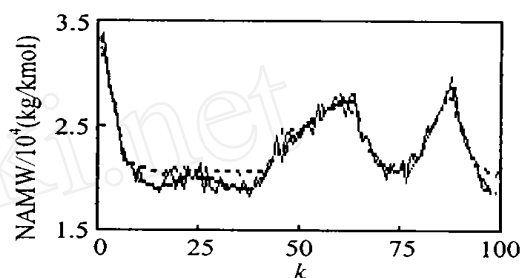
(b) 长期预测结果(RM S = 1 267)

图3 有噪声情况下的预测结果

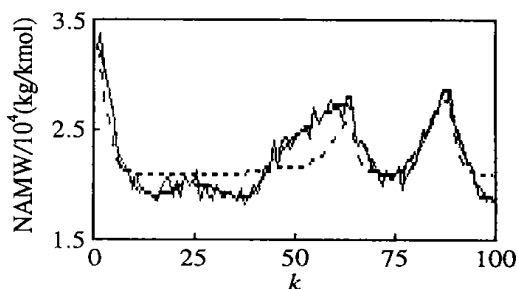
3.3 BP网络辨识结果

本文同时考察了采用BP网络对聚合过程进行建模的辨识结果。所采用的BP网络为具有调整因子的BP网络,引入调整因子是为了提高模型的泛化能力。与LS-SVM相比,两者在训练指标上类似,即在训练指标中都含有对权值的惩罚项,以提高泛化能力。

利用第二批数据前100步训练具有调整因子的BP网络,通过第二批测试集进行测试,模型一步预测结果和长期预测结果如图4所示。图中点划线为模型预测值,短划线为不含噪声的输出测试值,实线为含有噪声的输出测试值。



(a) 一步预测结果(RM S = 932)



(b) 长期预测结果(RM S = 1 765)

图4 有噪声时BP网络的预测结果

3.4 辨识结果讨论

通过仿真结果可以看出,在没有噪声时,LS-SVM在小样本的情况下仍能取得很好的辨识效果。即使存在严重噪声时,支持向量机的辨识精度也是令人满意的。

通过图3和图4的比较,LS-SVM在小样本情况下的泛化能力要比BP网络好,LS-SVM在测试集上的RM S远小于BP网络在测试集上的RM S,尤其是在长期预测效果上,LS-SVM的建模能力比BP网络有了很大的改进。

LS-SVM长期预测的RM S要比一步预测的RM S大,但结果仍令人满意。考虑到预测控制对模型失配具有鲁棒性,所以该结果能够满足预测控制的精度要求。这在下面的预测控制仿真中得到了验证。

证

支持向量机建立在小样本统计理论上, 有严密的理论体系, 因此在非线性建模中有着广泛的应用前景。应用LS-SVM 进行非线性建模, 能够提高模型的泛化能力、建模精度和训练速度。

4 基于支持向量机的预测控制

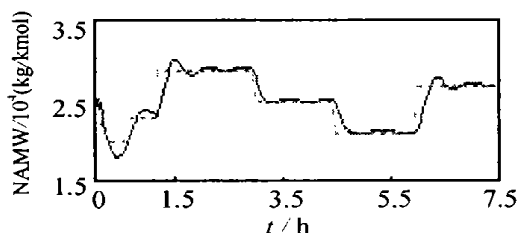
预测控制技术由于具有实时预测、实时优化及实时反馈校正的特点而获得广泛的应用。本文采用支持向量机建立过程模型并进行预测控制, 同时利用预测误差对模型进行校正以提高控制器的鲁棒性。

取优化性能指标为二次性能加权指标, 预测控制器是通过在 k 时刻对下面的性能指标进行优化以获得控制作用。

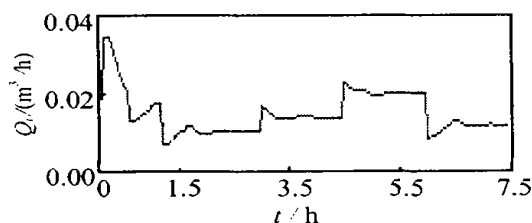
$$\begin{aligned} \min_{\Delta u(k), \dots, \Delta u(k+M-1)} & \sum_{j=1}^M \|\Delta u(k+j-1|k)\|_R^2 + \\ & \sum_{i=1}^P \|y^s(k+i) - \hat{y}(k+i|k) - \\ & y(k) + \hat{y}(k|k-P-1)\|_Q^2; \\ \text{s.t. } & u_{\min} \leq u \leq u_{\max}, \\ & \Delta u_{\min} \leq \Delta u \leq \Delta u_{\max}. \end{aligned} \quad (19)$$

其中: P 为预测时域, M 为控制时域, Q 和 R 为加权矩阵, $\Delta u(k), \dots, \Delta u(k+M-1)$ 为优化变量。在 k 时刻对该性能指标进行优化以获得控制作用序列 $\Delta u(k), \dots, \Delta u(k+M-1)$, 取 $\Delta u(k)$ 作用于系统。在 k 时刻以后重复上述过程进行滚动优化。

将基于LS-SVM 模型的预测控制算法应用到



(a) 分子量NAMW 的控制效果



(b) 反应中止剂进料流量 Q_i 的控制效果

图5 基于LS-SVM 的聚合过程的预测控制效果

一个时变的非线性聚合生产过程, 该过程如图1所示。通过将反应原料(单体)、溶剂和反应中止剂加入反应器中进行搅拌来获得反应产品混合物(主要为聚酯)。这里考虑单入单出情况, 通过反应中止剂流量 Q_i 来调节产物的分子量NAMW。

取预测时域 $P = 12$, 控制时域 $M = 1$, 加权矩阵 $R = 0.4, Q = 1$ 。控制效果如图5所示。系统跟踪给定值的曲线平稳, 表明基于LS-SVM 的非线性预测控制算法是有效的, 并具有良好的控制性能。

5 结 论

本文探讨了利用LS-SVM 进行非线性系统辨识的方法。仿真结果表明, 用LS-SVM 建立非线性系统模型是可行的。支持向量机作为统计学习理论的一种实现, 具有严密的理论体系。其拓扑结构可由支持向量确定, 在小样本情况下能以较高的精度建立非线性系统的模型。LS-SVM 用等式约束代替经典支持向量机中的不等式约束, 求解过程变成解一组等式方程, 避免了耗时的QP 求解问题, 求解速度相对加快。由此获得的非线性模型具有很好的泛化能力, 能满足动态控制的要求。

进一步, 本文将LS-SVM 模型应用到非线性预测控制, 提出了基于LS-SVM 模型的非线性预测控制算法。通过对一聚合过程的仿真实验, 验证了其可行性和有效性, 并具有良好的控制性能。基于LS-SVM 的预测控制为实现非线性系统控制的通用性提供了一种新的途径。

参考文献(References):

- [1] Bequette B W. Nonlinear control of chemical process: A review [J]. *Ind Eng Chem Res*, 1991, 30(4): 1391-1413.
- [2] Evelio H, Yaman A. Control of nonlinear systems using polynomial ARMA models [J]. *AIChE Journal*, 1993, 39(3): 446-460.
- [3] Maner R B, Doyle F J. Polymerization reactor control using autoregressive-plus volterra-based MPC [J]. *AIChE Journal*, 1997, 43(7): 1763-1784.
- [4] Cortes C, Vapnik V. Support vector machine [J]. *Machine Learning*, 1995, 20: 273-297.
- [5] Vladimir N Vapnik. 统计学习理论的本质[M]. 张学工译. 北京: 清华大学出版社, 2000.
- [6] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers [J]. *Neural Processing Letters*, 1999, 9(3): 293-300.