

基于支持向量机的非线性系统预测控制

张日东^{1,2} 王树青¹ 李平³

摘要 针对离散非线性系统, 提出一种可用于非线性过程的支持向量机预测控制方法, 并给出了控制律的收敛性分析. 该方法将复杂的非线性预测方程转化成直观而有效的线性形式, 同时利用线性预测控制方法求得解析的控制律, 避免了复杂的非线性优化求解, 对非线性工业焦化装置温度控制的仿真结果表明了算法的有效性.

关键词 支持向量机, 预测控制, 非线性过程, 工业焦化装置
中图分类号 TP273

Support Vector Machine Based Predictive Control for Nonlinear Systems

ZHANG Ri-Dong^{1,2} WANG Shu-Qing¹ LI Ping³

Abstract A support vector machine based predictive control method and its convergence analysis for nonlinear systems is presented. The method gives a direct and effective multi-step predicting method and uses linear methods to get the control law which avoids the complicated nonlinear optimization. Simulation results of temperature control of industrial coking equipment are presented in the paper showing the efficiency of this method.

Key words Support vector machine, predictive control, nonlinear process, industrial coking equipment

1 引言

模型预测控制 (Model predictive control, MPC)^[1] 具有预测模型、滚动优化和反馈校正三个基本特征和优良的控制性能及鲁棒性, 已成功应用于工业过程中. 对于非线性系统而言, 由于难以得到精确的模型, 因此对于研究非线性模型预测控制是不利的. 近几年来, 用支持向量机来描述非线性系统并对其控制已成为国际上研究的热点之一, 而把支持向量机应用于预测控制的研究近年来也逐渐增多, 已经出现了一些基于支持向量机的非线性预测控制方法^[2~8]. 这些方法大体上是利用支持向量机对非线性系统建模, 然后利用非线性搜索或线性化等方法求取控制律. 这些方法为支持向量机非线性预测控制的分析与解决问题提供了一个基本思路, 使得预测控制器的设计有了基本的依据.

然而, 支持向量机非线性预测控制方法或多或少还存在以下问题: 1) 虽然支持向量机比较圆满解决了非线性系统建模问题, 但由于它通过内积函数

定义的非线性变换将输入变换到一个高维空间, 因此得到的是高维复杂的过程模型, 在进行预测控制时导致系统多步预测方程非常复杂, 由于这个原因, 目前的一些支持向量机预测控制只局限在一步预测控制算法. 2) 由于支持向量机预测控制算法的优化目标是高维非线性的, 因此在进行多步预测控制时, 目前的优化方法通常是一些数值解法, 如梯度下降法等, 这些方法有时不能保证优化结果是全局最优解. 3) 支持向量机核函数所特有的点积结构使得基于支持向量机的非线性预测控制很难得到解析的控制律.

本文为解决以上问题提出了一种支持向量机非线性预测控制算法. 其主要思路是利用支持向量机非线性核函数的线性化表示, 将复杂的非线性多步预测方程转化为一系列简单直观的线性多步预测形式, 并利用线性预测控制^[9] 方法求取控制律, 从而不需要复杂的非线性优化搜索等求解方法而获取解析的预测控制律.

2 非线性系统模型及其表示

2.1 过程模型

采用最小二乘支持向量机^[10] 对非线性过程进行建模, 该过程用支持向量机模型描述为

$$y_m(k) = y_L(k) + y_{NL}(k) \quad (1)$$

$$y_L(k) = \bar{\Phi}^T \mathbf{X}, \bar{\Phi} = [a_1, a_2, \dots, a_n, b_0, b_1, \dots, b_{m-1}]^T$$

收稿日期 2006-5-23 收修改稿日期 2006-8-22
Received May 23, 2006; in revised form August 22, 2006
国家自然科学基金 (60421002) 资助
Supported by National Natural Science Foundation of China (60421002)
1. 浙江大学先进控制技术研究所工业控制技术国家重点实验室 杭州 310027 2. 杭州电子科技大学信息与控制研究所 杭州 310018 3. 辽宁石油化工大学信息与控制工程学院 抚顺 113001
1. National Key Laboratory of Industrial Control Technology, Institute of Advanced Process Control, Zhejiang University, Hangzhou 310027 2. Institute of Information and Control, Hangzhou Dianzi University, Hangzhou 310018 3. College of Information and Control, Liaoning University of Petroleum and Chemical Technology, Fushun 113001
DOI: 10.1360/aas-007-1066

$$\begin{aligned} \mathbf{X} &= [y(k-1), \dots, y(k-n), \\ &u(t-d-1), \dots, u(t-d-m)]^T \\ y_{NL}(k) &= \sum_{i=1}^N w_i \exp\{-\|\mathbf{X} - \mathbf{X}_i\|_2^2 / \sigma^2\} + \beta \end{aligned} \quad (2)$$

其中, 给定过程输入输出数据样本集为 $\{\mathbf{X}_i, y(i)\}_{i=1}^N$, $y_L(k)$ 为利用过程输入输出数据样本集建立的局部线性模型, $y_{NL}(k)$ 为利用过程输入输出数据样本集与基于线性模型的输出 $y_L(k)$ 之差建立的非线性部分模型, \mathbf{X} 为输入数据向量, \mathbf{X}_i 为不同采样时刻的输入数据向量, $y(i)$ 为与 \mathbf{X}_i 对应的过程输出. w_i, σ^2, β 为最小二乘支持向量机的参数.

$n, m, d + 1$ 分别为过程的输出、输入阶次、时滞, 也可以是依经验确定的三个模型参数, 由于它们的取值以及系统非线性引起的模型误差将通过支持向量机进行进一步辨识.

2.2 辨识算法

1) 采用带遗忘因子的最小二乘法, 用过程的输入输出数据对线性部分的参数 $\bar{\Phi}$ 辨识.

$$\begin{aligned} \bar{\Phi}_k &= \bar{\Phi}_{k-1} + \bar{K}(k)[y(k) - \bar{\Phi}_k^T \mathbf{X}_k] \\ \bar{K}(k) &= P(k-1)\mathbf{X}_k[\mathbf{X}_k^T P(k-1)\mathbf{X}_k + \gamma]^{-1} \\ P(k) &= [\bar{I} - \bar{K}(k)\mathbf{X}_k^T]P(k-1)/\gamma \end{aligned} \quad (3)$$

其中, k 代表第 k 次参数的值, $P(0) = \delta \hat{I} (\delta > 10^6)$, γ 为遗忘因子, \hat{I} 为适当维数的单位阵.

2) 定义 $\bar{e}(k) = y(k) - y_L(k)$, 用最小二乘支持向量机对非线性部分辨识, 得到的参数为

$$\begin{bmatrix} \beta \\ \mathbf{W} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \Omega + \gamma^{-1}I \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{E} \end{bmatrix} \quad (4)$$

其中, $\mathbf{W} = [w_1, w_2, \dots, w_N]^T$, $\mathbf{1} = [1, 1, \dots, 1]^T$, $\mathbf{E} = [\bar{e}(1), \bar{e}(2), \dots, \bar{e}(N)]^T$, Ω 为一个方阵, 其第 i 行 j 列的元素为 $\Omega_{ij} = \varphi(\mathbf{X}_i)^T \varphi(\mathbf{X}_j) = K(\mathbf{X}_i, \mathbf{X}_j)$, $K(\cdot, \cdot)$ 为核函数, 本文采用 RBF 核函数: $K(\mathbf{X}, \mathbf{X}_i) = \exp\{-\|\mathbf{X} - \mathbf{X}_i\|_2^2 / \sigma^2\}$, I 为适当维数的单位阵.

2.3 非线性模型的处理

由式 (1) 和式 (2) 得到非线性过程的模型为

$$\begin{aligned} y_m(k) &= y_L(k) + y_{NL}(k) \\ &= \bar{\Phi}^T \mathbf{X} + \sum_{i=1}^N w_i \exp\{-\|\mathbf{X} - \mathbf{X}_i\|_2^2 / \sigma^2\} + \beta \\ \bar{\Phi} &= [a_1, a_2, \dots, a_n, b_0, b_2, \dots, b_{m-1}]^T \\ \mathbf{X} &= [y(k-1), \dots, y(k-n), u(t-d-1), \dots, \\ &u(t-d-m)]^T \end{aligned} \quad (5)$$

下面对式 (5) 中的非线性部分进行处理.

令 $F_i(\mathbf{X}) = \exp\{-\|\mathbf{X} - \mathbf{X}_i\|_2^2 / \sigma^2\}$ ($i = 1, 2, \dots, N$), 将 $F_i(\mathbf{X})$ 利用泰勒展开, 其中心点设为 \mathbf{X}_0 :

$$\begin{aligned} F_i(\mathbf{X}) &= F_i(\mathbf{X}_0) + \frac{\partial F_i}{\partial \mathbf{X}(1)} \Big|_{\mathbf{x}=\mathbf{x}_0} [\mathbf{X}(1) - \mathbf{X}_0(1)] + \\ &\dots + \frac{\partial F_i}{\partial \mathbf{X}(n+m)} \Big|_{\mathbf{x}=\mathbf{x}_0} [\mathbf{X}(n+m) - \mathbf{X}_0(n+m)] + \\ \varepsilon_i(\mathbf{X}) &= \frac{\partial F_i}{\partial \mathbf{X}(1)} \Big|_{\mathbf{x}=\mathbf{x}_0} \mathbf{X}(1) + \dots + \frac{\partial F_i}{\partial \mathbf{X}(n+m)} \Big|_{\mathbf{x}=\mathbf{x}_0} \\ &\mathbf{X}(n+m) + F_i(\mathbf{X}_0) - \frac{\partial F_i}{\partial \mathbf{X}(1)} \Big|_{\mathbf{x}=\mathbf{x}_0} \mathbf{X}_0(1) - \dots - \\ &\frac{\partial F_i}{\partial \mathbf{X}(n+m)} \Big|_{\mathbf{x}=\mathbf{x}_0} \mathbf{X}_0(n+m) + \varepsilon_i(\mathbf{X}) = \frac{\partial F_i}{\partial \mathbf{X}(1)} \Big|_{\mathbf{x}=\mathbf{x}_0} \\ &\mathbf{X}(1) + \dots + \frac{\partial F_i}{\partial \mathbf{X}(n+m)} \Big|_{\mathbf{x}=\mathbf{x}_0} \mathbf{X}(n+m) + C_i + \\ &\varepsilon_i(\mathbf{X}) \\ C_i &= F_i(\mathbf{X}_0) - \frac{\partial F_i}{\partial \mathbf{X}(1)} \Big|_{\mathbf{x}=\mathbf{x}_0} \mathbf{X}_0(1) - \dots - \\ &\frac{\partial F_i}{\partial \mathbf{X}(n+m)} \Big|_{\mathbf{x}=\mathbf{x}_0} \mathbf{X}_0(n+m) \end{aligned} \quad (6)$$

其中, ε_i 是非线性函数符号, C_i 为某一常数, 其值与中心点 \mathbf{X}_0 有关, 通常取 $\mathbf{X}_0 = \mathbf{0}$, 这时 $C_i = F_i(\mathbf{X}_0)$.

将式 (6) 代入式 (5) 得到

$$\begin{aligned} y_m(k) &= \bar{\Phi}^T \mathbf{X} + \sum_{i=1}^N w_i F_i(\mathbf{X}) + \beta = \\ &\bar{\Phi}^T \mathbf{X} + \sum_{i=1}^N w_i [\frac{\partial F_i}{\partial \mathbf{X}} \mathbf{X} + C_i + \varepsilon_i(\mathbf{X})] + \beta = \\ &(\bar{\Phi}^T + \sum_{i=1}^N w_i \frac{\partial F_i}{\partial \mathbf{X}}) \mathbf{X} + \sum_{i=1}^N w_i C_i + \beta + \\ &\sum_{i=1}^N w_i \varepsilon_i(\mathbf{X}) \end{aligned} \quad (7)$$

其中, $\frac{\partial F_i}{\partial \mathbf{X}} = [\frac{\partial F_i}{\partial \mathbf{X}(1)} \Big|_{\mathbf{x}=\mathbf{x}_0}, \dots, \frac{\partial F_i}{\partial \mathbf{X}(n+m)} \Big|_{\mathbf{x}=\mathbf{x}_0}]^T$.

将 $y_m(k)$ 写成离散差分方程形式

$$y_m(k) = \bar{a}_1 y(k-1) + \cdots + \bar{a}_n y(k-n) + \bar{b}_0 u(k-d-1) + \cdots + \bar{b}_{m-1} u(k-d-m) + \sum_{i=1}^N w_i \varepsilon_i(\mathbf{X}) + \sum_{i=1}^N w_i C_i + \beta \quad (8)$$

式 (7) 与式 (8) 对比系数得

$$\begin{aligned} \bar{a}_1 &= a_1 + \sum_{i=1}^N w_i \partial \mathbf{F}_i(1) \\ \bar{a}_2 &= a_2 + \sum_{i=1}^N w_i \partial \mathbf{F}_i(2) \\ &\vdots \\ \bar{a}_n &= a_n + \sum_{i=1}^N w_i \partial \mathbf{F}_i(n) \\ \bar{b}_0 &= b_0 + \sum_{i=1}^N w_i \partial \mathbf{F}_i(n+1) \\ \bar{b}_1 &= b_1 + \sum_{i=1}^N w_i \partial \mathbf{F}_i(n+2) \\ &\vdots \\ \bar{b}_{m-1} &= b_{m-1} + \sum_{i=1}^N w_i \partial \mathbf{F}_i(n+m) \end{aligned} \quad (9)$$

经过处理后的模型等价表达式 (8) 包含三个部分: 第一部分为线性部分, 其系数由式 (9) 求取; 第二部分为非线性部分, 记为 $NL(\mathbf{X}) = \sum_{i=1}^N w_i \varepsilon_i(\mathbf{X})$; 第三部分为常数项, 记为 $C = \sum_{i=1}^N w_i C_i + \beta$. 从而式 (8) 表示为

$$y_m(k) = \bar{a}_1 y(k-1) + \cdots + \bar{a}_n y(k-n) + \bar{b}_0 u(k-d-1) + \cdots + \bar{b}_{m-1} u(k-d-m) + NL(\mathbf{X}) + C \quad (10)$$

3 预测控制器的设计

若采用自适应策略, 须在每次采样时刻到来时先辨识系统的模型参数 Φ , w_i , β , 然后基于这些参数组成的系统模型 (5), 利用第 2 节的方法进行处理得到式 (10), 由于辨识到的参数均为固定的数值, 因此由 C 的定义知在每个采样时刻 C 为某一常数, 但不同的采样时刻对应的 C 却是不同的数值; 若不采用自适应策略, 即认定模型 (5) 是精确的, 则参数均为固定的数值, 从而 C 为一个固定的常数, 且在每个采样时刻均不变. 因此, 无论是否采用自适应策略, 在每个采样时刻用第 2 节的方法处理模型时 C 均为某一与时间无关的常数, 考虑到这一点, 式 (10) 可

等价写为

$$y_m(k) = A_1 y(k-1) + \cdots + A_{n+1} y(k-n-1) + B_{1,0} \Delta u(k-d-1) + \cdots + B_{1,m-1} \Delta u(k-d-m) + \Delta NL(\mathbf{X}) \quad (11)$$

其中, $A_1 = 1 + \bar{a}_1$, $A_i = \bar{a}_i - \bar{a}_{i-1}$ ($i = 2, 3, \dots, n$), $A_{n+1} = -\bar{a}_n$, $B_{1,i} = \bar{b}_i$ ($i = 0, 1, \dots, m-1$).

以下将基于模型 (11) 设计非线性多步预测控制器. 将最优预测输出 $\hat{\mathbf{Y}}$ 分成三部分, 一部分取决于过去的输入和输出, 用 \mathbf{Y}_p 表示. 另一部分取决于现在和未来的输入, 表示为 \mathbf{GU} , 第三部分为预测误差, 包括非线性误差 \mathbf{E}_1 和外界干扰引起的误差 \mathbf{E}_2 . 这里 \mathbf{E}_2 的各元素相等, 均为采样时刻系统实际输出与基于模型预测输出之差. 同时为叙述清楚方便, 这里取预测时域和控制时域均为 p , 并定义

$$\begin{aligned} \mathbf{X}^i &= [y(k+i-1), \dots, y(k+i-n), \\ &u(t+i-d-1), \dots, u(t+i-d-m)]^T \\ &(i = 1, 2, \dots, p) \end{aligned}$$

因此得到

$$\hat{\mathbf{Y}} = \mathbf{Y}_p + \mathbf{GU} + \mathbf{E}_1 + \mathbf{E}_2 \quad (12)$$

其中,

$$\hat{\mathbf{Y}} = [\hat{y}(k+d+1/k), \hat{y}(k+d+2/k), \dots, \hat{y}(k+d+p/k)]^T$$

$$\mathbf{Y}_p = [y_p(k+d+1), y_p(k+d+2), \dots, y_p(k+d+p)]^T$$

$$\mathbf{U} = [\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+p-1)]^T$$

$$\mathbf{E}_1 = [\Delta NL(\mathbf{X}^1), \Delta NL(\mathbf{X}^2), \dots, \Delta NL(\mathbf{X}^p)]^T$$

$$G = \begin{bmatrix} B_{1,0} & & & \\ B_{2,0} & B_{1,0} & & 0 \\ \dots & \dots & & \\ B_{p,0} & B_{p-1,0} & \dots & B_{1,0} \end{bmatrix}$$

\mathbf{Y}_p 是系统的自由输出, 它可由其支持向量机模型计算. G 中的元素计算如下

$$\begin{aligned} B_{1,0} &= \bar{b}_0 \\ B_{i,0} &= \bar{b}_{i-1} + \sum_{j=1}^{i-1} A_j B_{i-j,0}, \quad i = 2, \dots, p \end{aligned} \quad (13)$$

取参考轨迹为

$$\begin{aligned} y_r(k+d) &= y(k) \\ y_r(k+d+i) &= \mu^i y(k) + (1-\mu^i) y_s \quad (14) \\ i &= 1, 2, \dots, p \end{aligned}$$

其中, μ 为输出柔化系数, y_s 为设定值. 另外, 分别设参考轨迹向量和目标函数为

$$\begin{aligned} \mathbf{Y}_r &= [y_r(k+d+1), y_r(k+d+2), \dots, \\ & y_r(k+d+p)]^T \\ J &= \min\{(\mathbf{Y}_r - \hat{\mathbf{Y}})^T(\mathbf{Y}_r - \hat{\mathbf{Y}}) + \lambda^2 \mathbf{U}^T \mathbf{U}\} \end{aligned} \quad (15)$$

其中, λ^2 为加权项, 由 $\partial J / \partial \mathbf{U} = 0$, 得

$$\mathbf{U} = (G^T G + \lambda^2 \tilde{I})^{-1} G^T (\mathbf{Y}_r - \mathbf{Y}_p - \mathbf{E}_1 - \mathbf{E}_2) \quad (16)$$

其中, \tilde{I} 为适当维数的单位阵, 设 \mathbf{q}^T 为 $(G^T G + \lambda^2 \tilde{I})^{-1} G^T$ 的第一行组成的行向量, 则 k 时刻的控制量为

$$u(k) = u(k-1) + \mathbf{q}^T (\mathbf{Y}_r - \mathbf{Y}_p - \mathbf{E}_1 - \mathbf{E}_2) \quad (17)$$

由于系统为非线性的, 因此 \mathbf{E}_1 是未知的, 本文采取迭代递推求取控制律, 算法为

$$\begin{aligned} \mathbf{E}_1^0 &= 0 \\ \mathbf{U}^0 &= (G^T G + \lambda^2 \tilde{I})^{-1} G^T (\mathbf{Y}_r - \mathbf{Y}_p - \mathbf{E}_2 - \mathbf{E}_1^0) \\ \mathbf{E}_1^{k+1} &= \mathbf{E}_1^k + \delta(k) (\mathbf{Y}_m^k - \hat{\mathbf{Y}}^k) \\ \mathbf{U}^{k+1} &= (G^T G + \lambda^2 \tilde{I})^{-1} G^T (\mathbf{Y}_r - \mathbf{Y}_p - \mathbf{E}_2 - \\ & \mathbf{E}_1^{k+1}) \end{aligned} \quad (18)$$

其中, \mathbf{E}_1^0 为 \mathbf{E}_1 的初始值, \mathbf{U}^0 为 \mathbf{U} 的初始值, 上标 k 代表各分量第 k 步的值. 而 \mathbf{Y}_m^k 为求出的 \mathbf{U}^k 代入支持向量机模型而求出的输出, $\delta(k)$ 为收敛因子, 它是正标量递减序列, $\delta(k) \in (0, 1)$, $\lim_{k \rightarrow \infty} \delta(k) = 0$. 适当选取 $\delta(k)$ 可以使得上述递推很快收敛, 这里取 $\delta(k) = 1/k$ 以保证控制律是收敛的.

当 $\hat{\mathbf{Y}} = \mathbf{Y}_m$ 时, 求出的 \mathbf{U}^k 即为最优控制律. 实际上若 $\|\mathbf{E}_1^{k+1} - \mathbf{E}_1^k\|$ 足够小, 则即可认为 \mathbf{U}^k 为控制律. 我们将本文提出的控制器称为 NSVMPC (Nonlinear support vector machine predictive control).

4 收敛性分析

由式 (18), 对充分大的 k 来说, $\delta(k)$ 将任意小, 为此 $\mathbf{E}_1^k, \mathbf{U}^k$ 的变化都将越来越小, 当 \mathbf{E}_1^k 收敛于真值 $\bar{\mathbf{E}}_1$ 时, $\hat{\mathbf{Y}}^k$ 将收敛于 \mathbf{Y}_m^k , 从而得到最优控制律. 因此本节从分析 \mathbf{E}_1^k 的收敛性入手, 若 \mathbf{E}_1^k 收敛于真值 $\bar{\mathbf{E}}_1$, 则由上面的分析可知, 最优控制律的求解是收敛的. 本文控制律的收敛性可以借助伴随微分方程的稳定性质来确定.

引理 1^[11]. 对于以下递推算法

$$\begin{aligned} \alpha(k) &= \alpha(k-1) + \delta(k) Q(k, \alpha(k-1), \Phi(k)) \\ \Phi(k) &= A(\alpha(k-1)) \Phi(k-1) + B(\alpha(k-1)) e(k) \end{aligned} \quad (19)$$

式中 $\alpha(k)$ 称为“估计量”, 并定义 $D_s = \{\alpha | A(\alpha) \text{ 的所有特征值都严格在单位圆内}\}$, D_R 是 D_s 的一个连通开子集, 在子集 D_R 上, 式 (19) 中的函数满足以下正则条件 C1~C5^[11].

C1: 函数 $Q(k, \alpha, \Phi)$ 在 $(\bar{\alpha}, \bar{\Phi})$ 附近关于 α 和 Φ 是 Lipschitz 连续的, 其中 $\bar{\alpha} \in D_R, \bar{\Phi}$ 是任意的; 且关于 α 和 Φ 是连续可微的.

C2: 对 $\bar{\alpha} \in D_R$, 矩阵函数 $A(\alpha), B(\alpha)$ 关于 α 是 Lipschitz 连续的.

C3: $\{e(k)\}$ 是个独立的随机向量序列, 且有 $\lim_{k \rightarrow \infty} E\{Q(k, \bar{\alpha}, \Phi(k, \bar{\alpha}))\} \rightarrow f^*(\alpha); \forall \bar{\alpha} \in D_R$.

其中 $\Phi(k, \bar{\alpha})$ 由下式定义

$$\begin{cases} \Phi(k, \bar{\alpha}) = A(\bar{\alpha}) \Phi(k-1, \bar{\alpha}) + B(\bar{\alpha}) e(k) \\ \Phi(0, \bar{\alpha}) = 0 \end{cases}$$

C4: $\sum_{k=1}^{\infty} \delta(k) = \infty, \sum_{k=1}^{\infty} [\delta(k)]^p < \infty$ ($p > 1$).

C5: $\delta(k)$ 是正标量递减序列, $\lim_{k \rightarrow \infty} \delta(k) = 0, \lim_{k \rightarrow \infty} \sup[\frac{1}{\delta(k-1)} - \frac{1}{\delta(k)}] < \infty$.

如果对估计量 $\bar{\alpha} \in D_R$ 的任何可能值能够定义平稳随机函数 $Q(k, \bar{\alpha}, \Phi(k, \bar{\alpha}))$, 则式 (19) 的伴随微分方程为

$$\frac{d}{d\tau} \alpha_D(\tau) = f^*(\alpha_D(\tau)) \quad (20)$$

其中 τ 为伪时间, $\tau = \sum_{i=1}^k \delta(k)$, 而

$$f = \lim_{k \rightarrow \infty} E\{Q(k, \bar{\delta}, \Phi(k, \bar{\alpha}))\} \quad (21)$$

若方程 (20) 存在一个正函数 $V(\alpha_D)$, 使得

$$\begin{cases} \frac{d}{d\tau} V(\alpha_D) \leq 0, & \forall \alpha_D \in D_A \\ \frac{d}{d\tau} V(\alpha_D) = 0, & \forall \alpha_D \in D_c, D_c \in D_A \end{cases} \quad (22)$$

则当 $k \rightarrow \infty$ 时, $\alpha(k)$ 以概率 1 趋于不变集 D_c , 即 $\alpha(k) \xrightarrow[k \rightarrow \infty]{w.p.1} D_c$. 同时, 如果 α^* 是微分方程 (20) 的整体渐近稳定平衡点, 则当 $k \rightarrow \infty$ 时, $\alpha(k)$ 又以概率 1 收敛于 α^* , 也就是 $\alpha(k) \xrightarrow[k \rightarrow \infty]{w.p.1} \alpha^*$.

定理 1. 对于离散非线性系统, 如果将其支持向量机模型 (1) 表示为式 (10) 的形式, 控制器设计为式 (18) 的形式, 控制器的设计参数 $\delta(k)$ 是正标量递减序列, 且 $\delta(k) \in (0, 1), \lim_{k \rightarrow \infty} \delta(k) = 0$, 则系统的预测控制律是收敛的.

证明.

1) 本文的递推算法

由式 (18) 及式 (12) 得

$$\begin{aligned} \mathbf{E}_1^{k+1} &= \mathbf{E}_1^k + \delta(k)(\mathbf{Y}_m^k - \hat{\mathbf{Y}}^k) = \\ & \mathbf{E}_1^k + \delta(k)(\mathbf{Y}_m^k - \mathbf{Y}_p - \mathbf{G}\mathbf{U}^k - \mathbf{E}_1^k - \mathbf{E}_2) = \\ & F\mathbf{E}_1^k + D \end{aligned} \quad (23)$$

其中, D, F 在每次递推时为某个常数矩阵

$$\begin{aligned} D &= \delta(k)[\mathbf{Y}_m^k - \mathbf{Y}_p - \mathbf{G}\mathbf{U}^k - \mathbf{E}_2] \\ F &= \text{diag}(1 - \delta(k), 1 - \delta(k), \dots, 1 - \delta(k)) \end{aligned}$$

2) 正则条件分析

首先定义 $D_S = \{F|F$ 的所有特征值都严格在单位圆内 $\}$. 根据 F 的特点可知

$$\lambda_k(F) = 1 - \delta(k) < 1, \quad k = 1, \dots, p \quad (24)$$

可见 D_S 为全平面, 即 $D_R = D_S = \mathbf{R}^n$. 其次, 由式 (23) 及 $\delta(k) = 1/k$, 条件 C1~C5 均可满足. 于是, 式 (23) 的伴随微分方程为

$$\begin{aligned} \frac{d\mathbf{E}_{1D}(\tau)}{d\tau} &= f^*(\mathbf{E}_{1D}(\tau)) \\ &= \lim_{k \rightarrow \infty} E\{\mathbf{Y}_m^k - \mathbf{Y}_p - \mathbf{G}\mathbf{U}^k - \mathbf{E}_{1D}(\tau) - \mathbf{E}_2\} \end{aligned} \quad (25)$$

而 $\mathbf{Y}_m^k = \mathbf{Y}_p + \mathbf{G}\mathbf{U}^k + \bar{\mathbf{E}}_1 + \mathbf{E}_2$, 进而有

$$f^*(\mathbf{E}_{1D}(\tau)) = \bar{\mathbf{E}}_1 - \mathbf{E}_{1D}(\tau) \quad (26)$$

式中, $\bar{\mathbf{E}}_1$ 为 \mathbf{E}_1 的真值.

设 \mathbf{E}_{1D}^* 是微分方程 (26) 的平衡点, 则有

$$\mathbf{E}_{1D}^* = \bar{\mathbf{E}}_1 \quad (27)$$

即, 平衡点位于 \mathbf{E}_1 的真值上.

3) 稳定性分析

取微分方程 (26) 的 Lyapunov 函数

$$V(\mathbf{E}_{1D}) = \frac{1}{2}(\bar{\mathbf{E}}_1 - \mathbf{E}_{1D})^T(\bar{\mathbf{E}}_1 - \mathbf{E}_{1D}) > 0 \quad (28)$$

则

$$\begin{aligned} \frac{d}{d\tau}V(\mathbf{E}_{1D}) &= (\bar{\mathbf{E}}_1 - \mathbf{E}_{1D})^T \frac{d}{d\tau}(\bar{\mathbf{E}}_1 - \mathbf{E}_{1D}) \\ &= -(\bar{\mathbf{E}}_1 - \mathbf{E}_{1D})^T(\bar{\mathbf{E}}_1 - \mathbf{E}_{1D}) \leq 0 \quad \forall \mathbf{E}_{1D} \end{aligned} \quad (29)$$

且 $\frac{d}{d\tau}V(\mathbf{E}_{1D}) = 0, \mathbf{E}_{1D} = \mathbf{E}_{1D}^* = \bar{\mathbf{E}}_1$ (30)

可见, 微分方程存在不变集 $D_C = \{\bar{\mathbf{E}}_1\}$, 而吸引域为全平面, 则由引理知 \mathbf{E}_1 收敛于 $\bar{\mathbf{E}}_1$. \square

5 仿真研究

我们选用镇海炼化化工公司的重油焦化装置作为研究对象, 以验证本文方法的有效性.

该装置主要任务是对原料渣油、分馏塔底循环油迅速加热, 为原油的进一步深加工提供原料. 原料渣油从南北两侧分两路送入加热炉对流室预热至 330°C 左右, 之后合并进入分馏塔底, 与焦炭塔顶来的油气接触并传热传质, 原料中轻组份蒸发, 上升至精馏段进行分离, 而原料中蜡油以上馏分与来自焦炭塔顶油气中被冷凝的重组份一起流入塔底. 约 360°C 的分馏塔底油经加热炉辐射进料泵分两路送至加热炉辐射室迅速加热至 495°C, 之后进入焦炭塔进行裂解反应. 由于整个装置是间歇操作, 因此该装置的处理过程比较复杂. 该装置的控制要求是出口温度误差控制在 $\pm 1.5^\circ\text{C}$ 之内, 以满足后续装置的运行并获得经济效益最大化.

当该装置在稳态 493°C 下运行时, 将设定值提到 495°C, 实际测取过程的加热炉出口温度数据, 得到的实际响应曲线见图 1. 从该装置的实际响应曲线可以看出, 该过程是一个具有大纯滞后、非线性、并具有反向特性的对象. 利用以上 300 组数据通过第 2 节的方法用支持向量机对该焦化装置的特性进行建模, 由于不知道该过程的确切时滞, 依经验从响应曲线中观察, 粗略设时滞 $d = 50$, 并辨识得到线性部分为

$$y_L(k) = a_1 y(k - 1) + b_1 u(k - 51)$$

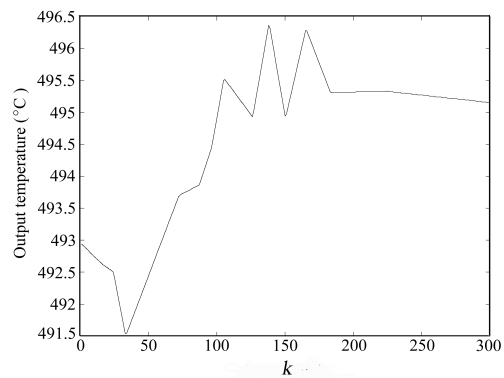


图 1 过程实际响应

Fig. 1 Real response of the system

然后利用实际输出数据与线性部分的误差进一步建模, 其中支持向量机的参数选取为: $\sigma^2 = 15$, 得到的建模结果见图 2 (实线为支持向量机的辨识效果, 虚线为过程的实际响应). 最终的模型为

$$\begin{aligned} y_m(k) &= a_1 y(k - 1) + b_1 u(k - 51) + \\ & \sum_{i=1}^N w_i \exp\{-\|\mathbf{X} - \mathbf{X}_i\|_2^2 / \sigma^2\} + \beta \end{aligned}$$

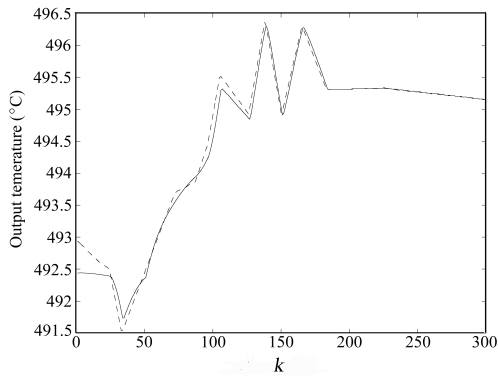


图2 支持向量机建模结果

Fig. 2 Modeling results based on SVM

以下仿真给出了本文方法 (NSVMPC) 与一般预测控制方法 (GPC) 的对比结果, 一般预测控制采用经典的最小二乘法辨识方法建立模型, 为

$$y_m(k) = a_1 y(k-1) + b_1 u(k-51)$$

5.1 模型匹配时的仿真结果

5.1.1 定值跟踪与抗干扰性能比较

仿真过程为: 在 $k=0$ 时, 设定值进行阶跃变化; 在 $k=500$ 时, 引入幅值为 -2 的负载流量干扰. 两种预测控制器的参数均为 $p=56, \lambda^2=1, \mu=0.95$, 仿真结果见图3和图4.

从仿真结果可以看出, 一般预测控制方法 (GPC) 针对强非线性、大时滞以及具有反向特性的过程难以达到理想的控制, 控制精度不是很高, 这在一些需要高精度控制的场合是不令人满意的, 而 NSVMPC 对这类工业过程达到了比较理想的控制效果.

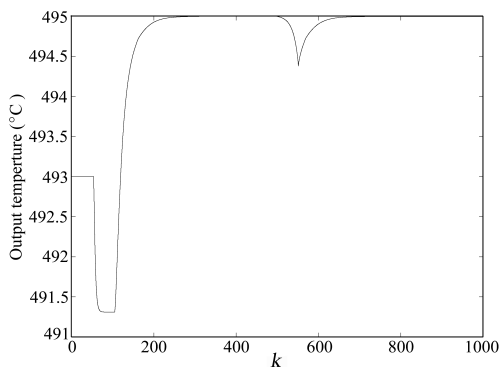


图3 NSVMPC 的仿真结果

Fig. 3 Simulation result of NSVMPC

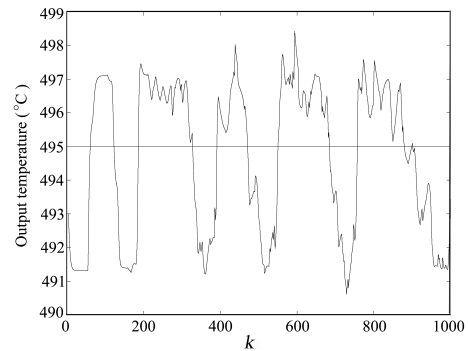


图4 GPC 的仿真结果

Fig. 4 Simulation result of GPC

5.1.2 随动跟踪与抗干扰性能比较

本节给出了两种控制器在随动跟踪矩形波性能方面的比较, 两种控制器的参数与上节相同. 仿真结果见图5和图6. 从仿真结果看, 在随动跟踪时, NSVMPC 也给出了比较理想的控制. 而一般 GPC 由于系统的非线性、大滞后以及反向特性的原因没有给出理想的控制效果.

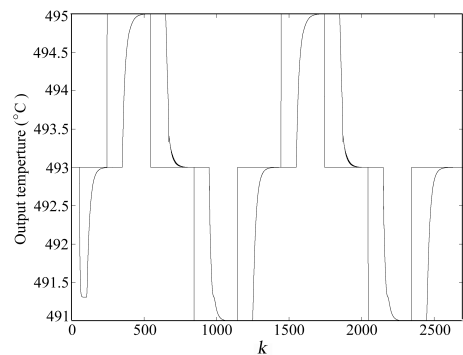


图5 NSVMPC 跟踪矩形波的仿真结果

Fig. 5 Result of rectangular signal tracking based on NSVMPC

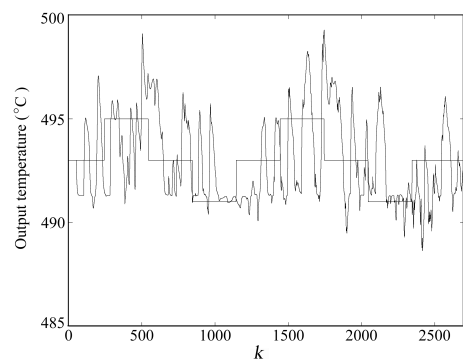


图6 GPC 跟踪矩形波的仿真结果

Fig. 6 Result of rectangular signal tracking based on GPC

5.2 模型失配时的仿真结果

5.2.1 参数 a_1, b_1 失配时的性能比较

在 $k = 0$ 时, 将模型的参数 a_1 改为 $a_1 - 0.007$, b_1 改为 $b_1 + 0.007$; 在 $k = 1200$ 时, 将参数 a_1 改为 $a_1 + 0.007$, b_1 改为 $b_1 - 0.007$. 这将使过程的时间常数和增益产生较大变化. 控制参数仍为 $p = 56, \lambda^2 = 1, \mu = 0.95$, 两种控制器的仿真结果如图 7 和图 8 所示. 从仿真结果看, NSVMPC 在参数失配情况下基本保持了比较满意的控制效果, GPC 的控制效果不是很令人满意.

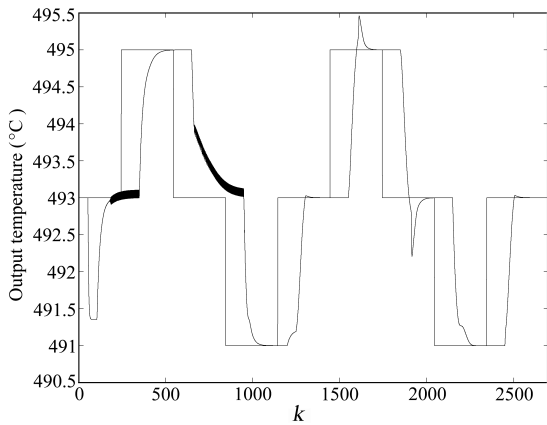


图 7 模型参数 a_1, b_1 失配时 NSVMPC 的仿真结果
Fig. 7 Result of NSVMPC with the mismatch of a_1, b_1

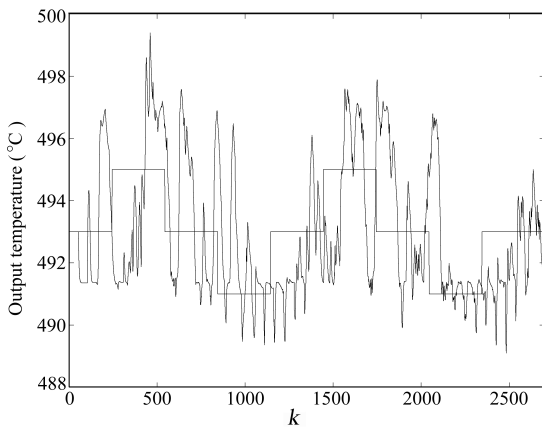


图 8 模型参数 a_1, b_1 失配时 GPC 的仿真结果
Fig. 8 Result of GPC with the mismatch of a_1, b_1

5.2.2 时滞 d 失配时的性能比较

本节分别给出了模型时滞和实际过程时滞相比估计过大和估计过小时 NSVMPC 的控制性能. 实际过程时滞为 $d = 50$. 假设模型时滞估计过大, 这时取模型的时滞为 $d_m = 54$; 当模型时滞估计过小时取模型的时滞 $d_m = 46$. NSVMPC 的控制结果如图 9 和图 10 所示, 而仿真结果显示当时滞如上述

情况变化时 GPC 给出了发散的仿真结果, 如图 11 和图 12 所示.

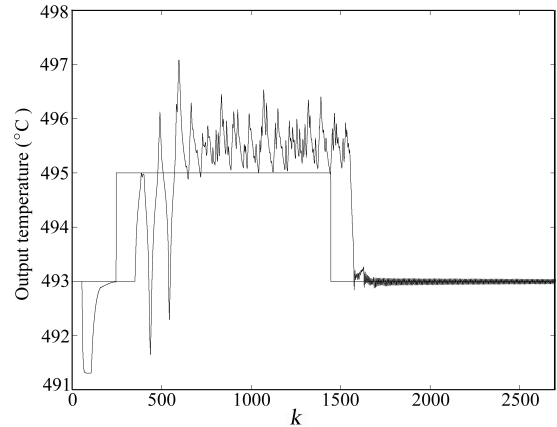


图 9 时滞失配时 ($d_m = 54$) NSVMPC 的仿真结果
Fig. 9 Result of NSVMPC with $d_m = 54$

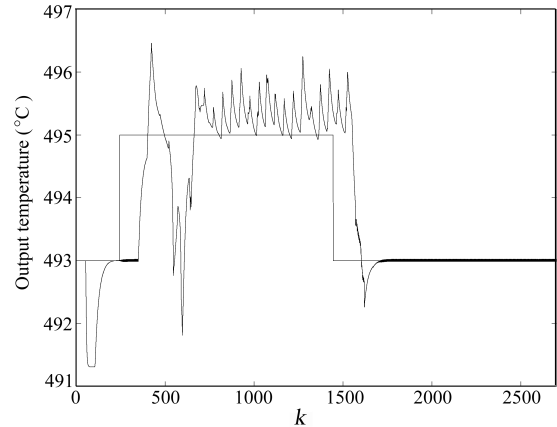


图 10 时滞失配时 ($d_m = 46$) NSVMPC 的仿真结果
Fig. 10 Result of NSVMPC with $d_m = 46$

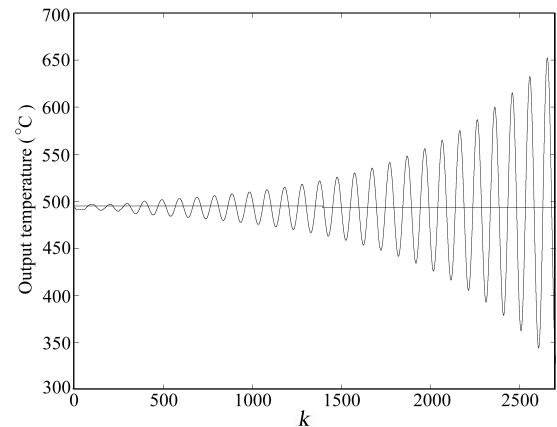


图 11 时滞失配时 ($d_m = 54$) GPC 的仿真结果
Fig. 11 Result of GPC with $d_m = 54$

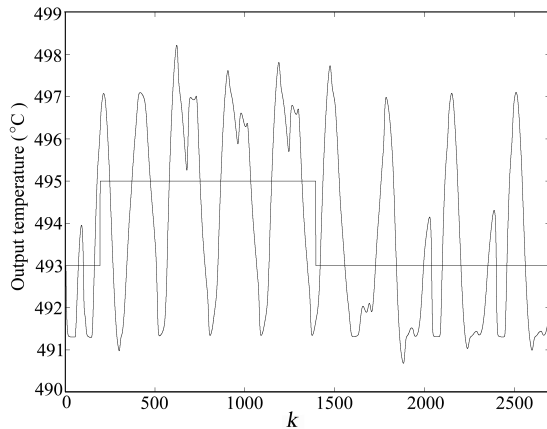


图 12 时滞失配时 ($d_m = 46$) GPC 的仿真结果

Fig. 12 Result of GPC with $d_m = 46$

6 结论

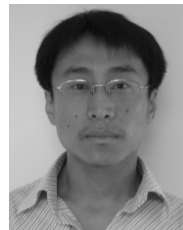
针对强非线性、大时滞以及具有反向特性的工业过程, 本文提出了一种新的支持向量机预测控制算法. 该算法将复杂的非线性多步预测转化为直观有效的线性形式, 不需要非线性求解. 同时本文还给出了它的具体实现方法以及控制律的收敛性分析. 仿真结果表明, 这种方法有良好的跟踪性能和抗干扰性能, 优于一般的预测控制方法.

References

- Clarke D W, Mohtadi C, Tuffs P S. Generalized predictive control part I: the basic algorithm. *Automatica*, 1987, **23**(2): 137~148
- Miao Q, Wang S F. Nonlinear model predictive control based on support vector regression. In: Proceedings of the First International Conference on Machine Learning and Cybernetics. Beijing, China, IEEE, 2002. 1657~1661
- Zhang Hao-Ran, Han Zheng-Zhi, Li Chang-Gang. Support vector machine based unknown systems identification and control. *Journal of Shanghai Jiao Tong University*, 2003, **37**(1): 927~930
(张浩然, 韩正之, 李昌刚. 基于支持向量机的未知非线性系统辨识与控制. 上海交通大学学报, 2003, **37**(1): 927~930)
- Zhang Hao-Ran, Han Zheng-Zhi, Li Chang-Gang. Support vector machine based nonlinear model predictive control. *System Engineering and Electronics*, 2003, **25**(3): 330~334
(张浩然, 韩正之, 李昌刚. 基于支持向量机的非线性模型预测控制. 系统工程与电子技术, 2003, **25**(3): 330~334)
- Liu Bin, Su Hong-Ye, Chu Jian. Predictive control algorithm based on least square support vector machines. *Control and Decision*, 2004, **19**(12): 1399~1402
(刘斌, 苏宏业, 褚建. 一种基于最小二乘支持向量机的预测控制算法. 控制与决策, 2004, **19**(12): 1399~1402)
- Ren Y, Cao G Y, Zhu X J. Predictive control of proton exchange membrane fuel cell (PEMFC) based on support vector regression machine. In: Proceedings of the Fourth

International Conference on Machine Learning and Cybernetics. Guangzhou, China, IEEE, 2005. 18~21

- Zhong Wei-Min, He Guo-Long, Pi Dao-Ying, Sun You-Xian. SVM with quadratic polynomial kernel function based nonlinear model one-step-ahead predictive control. *Chinese Journal of Chemical Engineering*, 2005, **13**(3): 373~379
- Li X, Cao G Y, Zhu X J. Modeling and control of PEMFC based on least squares support vector machines. *Energy Conversion and Management*, 2006, **47**(7): 1032~1050
- Jin Yuan-Yu, Gu Xing-Yuan. Modified generalized predictive control algorithm. *Information and Control*, 1990, **19**(3): 8~14
(金元郁, 顾兴源. 改进的广义预测控制算法. 信息与控制, 1990, **19**(3): 8~14)
- Suykens J A K, Vandewalle J. Least square support vector machine classifiers. *Neural Processing Letters*, 1999, **9**(3): 293~300
- Fang Chong-Zhi, Xiao De-Yun. *System Identification*. Beijing: Tsinghua University Press, 1998. 327~328
(方崇智, 萧德云. 过程辨识. 北京: 清华大学出版社, 1998. 327~328)



张日东 浙江大学先进控制技术研究所博士研究生, 杭州电子科技大学讲师. 主要研究方向为自适应控制、预测控制和工业过程优化. 本文通信作者.

E-mail: zrd-el@163.com

(ZHANG Ri-Dong Ph.D. candidate at the Institute of Advanced Process Control, Zhejiang University and

lecturer of Hangzhou Dianzi University. His research interest covers adaptive control, predictive control, and industrial process optimization. Corresponding author of this paper.)



王树青 浙江大学先进控制技术研究所教授. 研究方向为预测控制和工业过程优化. E-mail: sqwang@iipc.zju.edu.cn

(WANG Shu-Qing Professor at the Institute of Advanced Process Control, Zhejiang University. His research interest covers predictive control and industrial process optimization.)



李平 辽宁石油化工大学教授. 主要研究方向为预测控制和工业过程先进控制及优化. E-mail: liping@lnpu.edu.cn

(LI Ping Professor at Liaoning University of Petroleum and Chemical Technology. His research interest covers predictive control, advanced control and optimization of industrial processes.)