文章编号: 1000−8152(2008)02−0193−06

基于支持向量机的非线性离散动力学系统控制

刘 丁, 刘 涵

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

摘要: 提出了一种用支持向量机辨识系统状态空间模型的非线性离散动力学系统控制新方法. 在本方法中, 采用最小二乘支持向量机在每一个工作点辨识非线性系统的局部最优线性化模型. 针对该模型, 采用常规的线性控制方法在每个工作点设计局部线性控制器, 并在整个控制任务的每个工作点重复此设计过程.用该方法对两个典型的非线性离散系统采用极点配置技术进行了仿真验证, 结果显示系统对参考输入具有满意的跟踪性能, 证明该方法是有效和可行的.

关键词: 支持向量机; 最小二乘支持向量机; 非线性离散系统; 线性控制技术

中图分类号: TP181 **文献标识码**: A

Nonlinear discrete dynamic system based on support vector machines

LIU Ding, LIU Han

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an Shaanxi 710048, China)

Abstract: A new approach to control a nonlinear discrete dynamic system based on support vector machines (SVM) is proposed in this paper, which depends on the identification of a state space model by SVM. Firstly, a local optimal linearization model is identified at every operating point by least squares support vector machines (LS-SVM), which belongs to the least squares version of SVM. For a linearization model, any linear controller design technique can be applied to design local linear controller at the operating point, and design procedure is repeated at every operating point in the control task. The proposed approach is applied to two typical examples. Pole placement technique is chosen as the linear design technique. Finally, simulation results show that the system has satisfactory tracking performance with reference input because of the desirable ability of SVM.

Key words: support vector machines; least squares support vector machines; nonlinear discrete system; linear control technique

1 引言(Introduction)

非线性系统控制的难点之一是非线性系统的多样性和复杂性使其缺乏通用性强的模型, 而神经网络具有逼近非线性函数的能力. 20世纪80年代末、90年代初, 有关神经网络对于任意非线性函数逼近能力的研究取得了重要进展[1], 神经网络应用于非线性系统的控制也成为当今自动控制领域的研究热点之一[2,3]. 实际中的非线性系统一般都具有数学模型难以建立、扰动有界但未知、只有控制对象的输入/输出数据、采用常规的控制理论一般难以奏效等特点, 基于数据的神经网络方法则为非线性系统的控制问题提供了解决途径. 但是在神经网络控制中, 神经网络的逼近误差对神经网络自适应控制的性能影响较大, 这与具体的神经网络结构有关、与神经网络的参数选取有关[4] .

统计学习理论(statistical learning theory, SLT)是由Vapnik建立的一种专门研究小样本下机器学习规律的理论[5,6], 支持向量机(support vector machines, SVM)是在这一理论基础上发展起来的一种新的分类和回归工具. 支持向量机通过结构风险最小化原理来提高泛化能力, 将求解最优分类面问题转化为求解凸二次规划问题, 较好的解决了小样本、非线性、高维数、局部极小点等实际问题, 克服了神经网络结构和参数难以选择、易陷入局部极值、过拟合等缺陷, 已经被成功应用于模式识别和系统控制领域[7~9]. 本文提出了一种基于支持向量机的非线性离散系统控制新方法. 用支持向量机回归的方法辨识非线性离散系统的状态空间模型, 在每一个工作点, 对支持向量机回归模型进行线性化, 依此设计线性控制器并控制系统到下一个工作点, 在新的工作

收稿日期: 2006−11−16; 收修改稿日期: 2007−04−09.
基金项目: 国家自然科学基金资助项目(60675048); 高等学校博士学科点专项科研基金资助项目(20040700010).

点上重复上述的设计过程, 最终完成对系统的有效控制. 这种方法不需要非线性系统的精确模型, 只需要系统的输入/输出数据; 基于支持向量机回归的线性化方法计算量小, 每一步只需更新拉格朗日乘子; 同时在每个工作点, 传统的状态空间线性控制技术都可以用来设计控制器. 最后对两个非线性离散动力学系统的仿真研究验证了该方法的有效性.

2 非线性系统的局部线性化状态空间模型(Local linearization model of nonlinear system)

考虑如下非线性离散动力学系统:

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{W}(\boldsymbol{x}(k), \boldsymbol{u}(k)), \\ \boldsymbol{y}(k+1) = \boldsymbol{V}(\boldsymbol{x}(k), \boldsymbol{u}(k)). \end{cases} \tag{1}$$

其中: $\boldsymbol{W}: \mathbb{R}^{n+m} \to \mathbb{R}^n$和$\boldsymbol{V}: \mathbb{R}^{n+m} \to \mathbb{R}^p$为非线性函数, $\boldsymbol{x} \in \mathbb{R}^n$,$\boldsymbol{u} \in \mathbb{R}^m$和$\boldsymbol{y} \in \mathbb{R}^p$分别为系统的状态、控制量和输出. 在任一工作点式(1)可以写成等价的线性模型的形式:

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{A}_k\boldsymbol{x}(k) + \boldsymbol{B}_k\boldsymbol{u}(k), \\ \boldsymbol{y}(k+1) = \boldsymbol{C}_k\boldsymbol{x}(k) + \boldsymbol{D}_k\boldsymbol{u}(k). \end{cases} \tag{2}$$

其中$\boldsymbol{A}_k$,$\boldsymbol{B}_k$,$\boldsymbol{C}_k$和$\boldsymbol{D}_k$分别为适当维数的矩阵, 由式(1)和(2), 则在工作点$(\boldsymbol{x}_k, \boldsymbol{u}_k)$处满足

$$\begin{cases} \boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k, \\ \boldsymbol{V}(\boldsymbol{x}_k, \boldsymbol{u}_k) = \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{D}_k\boldsymbol{u}_k. \end{cases} \tag{3}$$

其中$(\boldsymbol{x}_k, \boldsymbol{u}_k) = (\boldsymbol{x}(k), \boldsymbol{u}(k))$. 那么在工作点$(\boldsymbol{x}_k, \boldsymbol{u}_k)$的邻域, 有下面的等式成立:

$$\boldsymbol{W}(\boldsymbol{x}, \boldsymbol{u}) \approx \boldsymbol{A}_k\boldsymbol{x}(k) + \boldsymbol{B}_k\boldsymbol{u}(k), \tag{4}$$

$$\boldsymbol{V}(\boldsymbol{x}, \boldsymbol{u}) \approx \boldsymbol{C}_k\boldsymbol{x}(k) + \boldsymbol{D}_k\boldsymbol{u}(k). \tag{5}$$

为了获得非线性系统(1)在工作点处的线性化模型, 必须估计系数矩阵$\boldsymbol{A}_k$、$\boldsymbol{B}_k$、$\boldsymbol{C}_k$和$\boldsymbol{D}_k$的值. 函数$\boldsymbol{W}(\boldsymbol{x}, \boldsymbol{u})$在工作点$(\boldsymbol{x}_k, \boldsymbol{u}_k)$附近可以用泰勒级数近似展开为

$$\boldsymbol{W}(\boldsymbol{x}, \boldsymbol{u}) \approx \boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \nabla_{\boldsymbol{x}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k)[\boldsymbol{x} - \boldsymbol{x}_k] + \nabla_{\boldsymbol{u}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k)[\boldsymbol{u} - \boldsymbol{u}_k]. \tag{6}$$

其中$\nabla_{\boldsymbol{x}_k}\boldsymbol{W}$和$\nabla_{\boldsymbol{u}_k}\boldsymbol{W}$分别为$\boldsymbol{W}(\boldsymbol{x}, \boldsymbol{u})$对向量$\boldsymbol{x}_k$和$\boldsymbol{u}_k$的梯度:

$$\begin{cases} \nabla_{\boldsymbol{x}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) = \dfrac{\partial \boldsymbol{W}(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}_k}, \\ \nabla_{\boldsymbol{u}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) = \dfrac{\partial \boldsymbol{W}(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{u}}|_{\boldsymbol{u}=\boldsymbol{u}_k}. \end{cases} \tag{7}$$

由式(4)和(6)可得

$$\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \nabla_{\boldsymbol{x}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k)[\boldsymbol{x} - \boldsymbol{x}_k] + \nabla_{\boldsymbol{u}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k)[\boldsymbol{u} - \boldsymbol{u}_k] \approx \boldsymbol{A}_k\boldsymbol{x} + \boldsymbol{B}_k\boldsymbol{u}. \tag{8}$$

将式(3)代入式(8)

$$\nabla_{\boldsymbol{x}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k)[\boldsymbol{x} - \boldsymbol{x}_k] + \nabla_{\boldsymbol{u}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k)[\boldsymbol{u} - \boldsymbol{u}_k] \approx \boldsymbol{A}_k[\boldsymbol{x} - \boldsymbol{x}_k] + \boldsymbol{B}_k[\boldsymbol{u} - \boldsymbol{u}_k]. \tag{9}$$

为了使式(9)成立, 则必须满足

$$\nabla_{\boldsymbol{x}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) \approx \boldsymbol{A}_k. \tag{10}$$

$$\nabla_{\boldsymbol{u}_k}\boldsymbol{W}(\boldsymbol{x}_k, \boldsymbol{u}_k) \approx \boldsymbol{B}_k. \tag{11}$$

定义$\boldsymbol{w}_j$为$\boldsymbol{W}$的第j行, $\boldsymbol{a}_{kj}$和$\boldsymbol{b}_{kj}$分别为系数矩阵$\boldsymbol{A}_k$和$\boldsymbol{B}_k$的第j行, $1 \leqslant j \leqslant n$, 则式(10)(11)可分别写为

$$\nabla_{\boldsymbol{x}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) \approx \boldsymbol{a}_k, \tag{12}$$

$$\nabla_{\boldsymbol{u}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) \approx \boldsymbol{b}_k. \tag{13}$$

构造下面的目标函数:

$$J = \frac{1}{2} \| \nabla_{\boldsymbol{x}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \boldsymbol{a}_k \|^2 + \frac{1}{2} \| \nabla_{\boldsymbol{u}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \boldsymbol{b}_k \|^2 . \tag{14}$$

当在等式(3)约束下

$$\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) = \boldsymbol{a}_{kj}^{\mathrm{T}}\boldsymbol{x}_k + \boldsymbol{b}_{kj}^{\mathrm{T}}\boldsymbol{u}_k. \tag{15}$$

最小化目标函数J即可获得系统的最优的局部线性化模型. 根据拉格朗日乘子法, 重新构造目标函数

$$J_a = \frac{1}{2} \| \nabla_{\boldsymbol{x}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \boldsymbol{a}_k \|^2 + \frac{1}{2} \| \nabla_{\boldsymbol{u}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \boldsymbol{b}_k \|^2 + \lambda[\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \boldsymbol{a}_{kj}^{\mathrm{T}}\boldsymbol{x}_k - \boldsymbol{b}_{kj}^{\mathrm{T}}\boldsymbol{u}_k]. \tag{16}$$

其中λ为拉格朗日乘子, 分别求取J_a对λ,$\boldsymbol{a}_{kj}^{\mathrm{T}}$和$\boldsymbol{b}_{kj}^{\mathrm{T}}$的梯度, 并使其等于0:

$$\nabla_{\boldsymbol{a}_{kj}^{\mathrm{T}}} J_a = \boldsymbol{a}_{kj}^{\mathrm{T}} - \nabla_{\boldsymbol{x}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \lambda\boldsymbol{x}_k, \tag{17}$$

$$\nabla_{\boldsymbol{b}_{kj}^{\mathrm{T}}} J_a = \boldsymbol{b}_{kj}^{\mathrm{T}} - \nabla_{\boldsymbol{u}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \lambda\boldsymbol{u}_k, \tag{18}$$

$$\nabla_{\lambda} J_a = \boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \boldsymbol{a}_{kj}^{\mathrm{T}}\boldsymbol{x}_k - \boldsymbol{b}_{kj}^{\mathrm{T}}\boldsymbol{u}_k. \tag{19}$$

从式(17)(18)解出$\boldsymbol{a}_{kj}^{\mathrm{T}}$和$\boldsymbol{b}_{kj}^{\mathrm{T}}$并代入式(19), 可解得

$$\lambda = \frac{\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k) - \nabla_{\boldsymbol{x}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k)\boldsymbol{x}_k - \nabla_{\boldsymbol{u}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k)\boldsymbol{u}_k}{\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{x}_k + \boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{u}_k}. \tag{20}$$

将式(20)代入式(18)和(19), 可计算出局部线性化模型的系数矩阵$\boldsymbol{A}_k$,$\boldsymbol{B}_k$的行:

$$\boldsymbol{a}_{kj}^{\mathrm{T}} = \lambda\boldsymbol{x}_k + \nabla_{\boldsymbol{x}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k), \tag{21}$$

$$\boldsymbol{b}_{kj}^{\mathrm{T}} = \lambda\boldsymbol{u}_k + \nabla_{\boldsymbol{u}_k}\boldsymbol{w}_j(\boldsymbol{x}_k, \boldsymbol{u}_k). \tag{22}$$

同理, 式(1)输出方程中的系数矩阵$\boldsymbol{C}_k$和$\boldsymbol{D}_k$的行$\boldsymbol{c}_{kj}^{\mathrm{T}}$和$\boldsymbol{d}_{kj}^{\mathrm{T}}$也可以用上述同样的方法求得

$$\boldsymbol{c}_{kj}^{\mathrm{T}} =$$

$$\frac{\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k)-\nabla_{\boldsymbol{x}_k}\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k)\boldsymbol{x}_k-\nabla_{\boldsymbol{u}_k}\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k)\boldsymbol{u}_k}{\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{x}_k+\boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{u}_k}\boldsymbol{x}_k+ \nabla_{\boldsymbol{x}_k}\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k), \tag{23}$$

$$\boldsymbol{d}_{kj}^{\mathrm{T}}= \frac{\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k)-\nabla_{\boldsymbol{x}_k}\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k)\boldsymbol{x}_k-\nabla_{\boldsymbol{u}_k}\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k)\boldsymbol{u}_k}{\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{x}_k+\boldsymbol{u}_k^{\mathrm{T}}\boldsymbol{u}_k}\boldsymbol{u}_k+ \nabla_{\boldsymbol{u}_k}\boldsymbol{v}_j(\boldsymbol{x}_k,\boldsymbol{u}_k). \tag{24}$$

当式(1)输出方程仅为$\boldsymbol{x}$的函数时, 即

$$\boldsymbol{y}(k)=\boldsymbol{V}(\boldsymbol{x}(k)), \tag{25}$$

则其近似的线性化的模型为

$$\boldsymbol{y}(k)=\boldsymbol{C}_k\boldsymbol{x}(k). \tag{26}$$

这时, 式(23)则改为如下的形式:

$$\boldsymbol{c}_{kj}^{\mathrm{T}}=\nabla_{\boldsymbol{x}_k}\boldsymbol{v}_j(\boldsymbol{x}_k)+\frac{\boldsymbol{v}_j(\boldsymbol{x}_k)-\nabla_{\boldsymbol{x}_k}\boldsymbol{v}_j(\boldsymbol{x}_k)\boldsymbol{x}_k}{\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{x}_k}. \tag{27}$$

通过计算式(21)~(24), 就可以获得非线性离散系统(1)在最小建模误差(14)意义下的局部线性化模型(2), 对模型(2)设计常规的线性控制器, 就可以完成对非线性离散系统(1)的控制.

3 基于支持向量回归的非线性系统建模(Nonlinear system modeling based on SVM)

3.1 支持向量机回归的原理(The principle of support vector machines)

支持向量机回归的基本思想为: 选择一个非线性变换$\Phi(\cdot)$把n 维输入, 1维输出样本向量

$$\begin{aligned}&(\boldsymbol{r}_1,o_1),(\boldsymbol{r}_2,o_2),\cdots,(\boldsymbol{r}_l,o_l)\\&\boldsymbol{r}_i\in\mathbb{R}^n,\ o_i\in\mathbb{R},\ i=1,\cdots,l.\end{aligned} \tag{28}$$

从原空间映射到高维特征空间$\mathbb{F}$, 在此空间构造最优线性回归函数

$$o=f(\boldsymbol{r})=\boldsymbol{\omega}^{\mathrm{T}}\cdot\Phi(\boldsymbol{r})+b. \tag{29}$$

同时, 利用结构风险最小化原则引入了间隔的概念, 并利用原空间的核函数取代高维特征空间的点积运算, 避免了复杂计算. 标准支持向量机和最小二乘支持向量机[10](least square support vector machines, LS-SVM)在利用结构风险原则时, 在优化目标中选择了不同的损失函数, 它们分别为松弛因子ξ_i 和误差ξ_i的二范数.

对于标准的支持向量机, 优化问题为

$$\min_{\boldsymbol{\omega},b,\xi}\ \frac12\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}+\gamma\sum_{i=1}^{l}\xi_i. \tag{30}$$

约束条件为

$$\begin{cases}o_i-\boldsymbol{\omega}^{\mathrm{T}}\Phi(\boldsymbol{r}_i)-b\leqslant\varepsilon+\xi_i,\\ \boldsymbol{\omega}^{\mathrm{T}}\Phi(\boldsymbol{r}_i)+b-o_i\leqslant\varepsilon+\xi_i^*,\\ \xi_i,\xi_i^*\geqslant0,\ i=1,\cdots,l.\end{cases} \tag{31}$$

常数$\gamma>0$, 它控制对超出误差ε的样本的惩罚的程度, ε为Vapnik-ε不敏感代价函数定义的误差[5].

对于最小二乘支持向量机(LS-SVM), 优化问题变为

$$\min_{\boldsymbol{\omega},b,\xi}\frac12\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}+\gamma\frac12\sum_{i=1}^{l}\xi_i^2. \tag{32}$$

约束条件变为等式约束:

$$o_i=\boldsymbol{\omega}^{\mathrm{T}}\Phi(\boldsymbol{r}_i)+b+\xi_i,\ i=1,\cdots,l. \tag{33}$$

LS-SVM定义了与标准支持向量机不同的代价函数, 并将其不等式约束改为等式约束. 为求解式(32)的优化问题, 引入拉格朗日函数:

$$\begin{aligned}L=&\frac12\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}+\gamma\frac12\sum_{i=1}^{l}\xi_i^2-\\&\sum_{i=1}^{l}\alpha_i[\boldsymbol{\omega}^{\mathrm{T}}\Phi(\boldsymbol{r}_i)+b-o_i+\xi_i].\end{aligned} \tag{34}$$

其中α_i为拉格朗日乘子. 根据KKT条件[5], 得到如下等式和约束条件:

$$\begin{cases}\boldsymbol{\omega}=\sum\limits_{i=1}^{l}\alpha_i\Phi(\boldsymbol{r}_i),\ \sum\limits_{i=1}^{l}\alpha_i=0,\\ \alpha_i=\gamma\cdot\xi_i,\ \boldsymbol{\omega}^{\mathrm{T}}\Phi(\boldsymbol{r}_i)+b-o_i+\xi_i=0.\end{cases} \tag{35}$$

对于$i=1,\cdots,l$, 上式消去$\boldsymbol{\omega}$和ξ得到式(36)的线性方程

$$\begin{bmatrix}0 & \boldsymbol{O}^{\mathrm{T}}\\ \boldsymbol{O} & \boldsymbol{ZZ}^{\mathrm{T}}+\gamma^{-1}I\end{bmatrix}\begin{bmatrix}b\\ \alpha\end{bmatrix}=\begin{bmatrix}0\\ \vec{1}\end{bmatrix}. \tag{36}$$

其中:

$$\begin{cases}\boldsymbol{Z}^{\mathrm{T}}=[o_1\Phi(\boldsymbol{r}_1),\cdots,o_l\Phi(\boldsymbol{r}_l)],\ \boldsymbol{O}=[o_1,\cdots,o_l]^{\mathrm{T}},\\ \vec{1}=[1,\cdots,1]^{\mathrm{T}},\ \alpha^{\mathrm{T}}=[\alpha_1,\cdots,\alpha_l].\end{cases} \tag{37}$$

由式(36), 根据Mercer条件[5], 可以令$\boldsymbol{Q}=\boldsymbol{ZZ}^{\mathrm{T}}$, 则

$$\begin{cases}\boldsymbol{Q}_{kh}=o_ko_h\Phi(\boldsymbol{r}_k)^{\mathrm{T}}\Phi(\boldsymbol{r}_h)=o_ko_h\cdot K(\boldsymbol{r}_k,\boldsymbol{r}_h),\\ k,h=1,\cdots,l.\end{cases} \tag{38}$$

其中$K(\boldsymbol{r}_k,\boldsymbol{r}_h)=\Phi(\boldsymbol{r}_k)^{\mathrm{T}}\Phi(\boldsymbol{r}_h)$定义为核函数, 则式(36)修改为

$$\begin{bmatrix}0 & o_1 & \cdots & o_l\\ o_1 & o_1o_1K(\boldsymbol{r}_1,\boldsymbol{r}_1)+\dfrac{2}{\gamma} & \cdots & o_1o_lK(\boldsymbol{r}_1,\boldsymbol{r}_l)\\ \vdots & \vdots & \ddots & \vdots\\ o_l & o_lo_1K(\boldsymbol{r}_l,\boldsymbol{r}_1) & \cdots & o_lo_lK(\boldsymbol{r}_l,\boldsymbol{r}_l)+\dfrac{2}{\gamma}\end{bmatrix}\times \begin{bmatrix}b\\ \alpha_1\\ \vdots\\ \alpha_l\end{bmatrix}=\begin{bmatrix}0\\1\\ \vdots\\1\end{bmatrix}. \tag{39}$$

因此支持向量的优化问题转化为解上述线性方程的问题. 该线性方程可以用最小二乘的方法求解, 并且要比SVM中求解二次规划快的多, 且所需计算资源少. 应用LS-SVM对非线性函数回归的结果为

$$o=f(\boldsymbol{r})=\sum_{i=1}^{l}\alpha_iK(\boldsymbol{r},\boldsymbol{r}_i)+b,\ i=1,\cdots,l.\tag{40}$$

3.2 基于LS–SVM的局部线性化模型(Local linearization model based on LS–SVM)

重写式(1)的非线性离散系统的模型:

$$\boldsymbol{x}(k+1)=\boldsymbol{W}(\boldsymbol{x}(k),\boldsymbol{u}(k)).\tag{41}$$

$$\boldsymbol{y}(k+1)=\boldsymbol{V}(\boldsymbol{x}(k),\boldsymbol{u}(k)).\tag{42}$$

由上面的讨论可知, 如果获得局部线性化模型的系数矩阵$\boldsymbol{A}_k$,$\boldsymbol{B}_k$,$\boldsymbol{C}_k$和$\boldsymbol{D}_k$, 即求解式(21)~(24), 就可以通过设计适当的线性控制器对(1)施加有效的控制. 而式(21)~(24)的求解需要获得$\boldsymbol{W}(\boldsymbol{x},\boldsymbol{u})$, $\boldsymbol{V}(\boldsymbol{x},\boldsymbol{u})$和它们的对$\boldsymbol{x}$,$\boldsymbol{u}$的梯度在工作点处的值. 本文中, 非线性函数关系$\boldsymbol{W}$和$\boldsymbol{V}$分别用最小二乘支持向量机(LS–SVM)来辨识. 假定系统阶数n已知, 系统状态都可以获得, 则基于支持向量回归的系统模型为

$$\boldsymbol{x}(k+1)=\widehat{\boldsymbol{W}}(\boldsymbol{x}(k),\boldsymbol{u}(k)),\tag{43}$$

$$\boldsymbol{y}(k+1)=\widehat{\boldsymbol{V}}(\boldsymbol{x}(k),\boldsymbol{u}(k)).\tag{44}$$

其中: $\boldsymbol{x}\in\mathbb{R}^n$, $\boldsymbol{u}\in\mathbb{R}^m$和$\boldsymbol{y}\in\mathbb{R}^p$分别为系统的状态、控制量和输出, $\widehat{\boldsymbol{W}}$和$\widehat{\boldsymbol{V}}$分别代表近似$\boldsymbol{W}$和$\boldsymbol{V}$的支持向量机回归模型; $\widehat{\boldsymbol{W}}$有$n+m$个输入, 对应于k时刻的n个状态和m个输入, n个输出对应$k+1$时刻系统的状态; $\widehat{\boldsymbol{V}}$有$n+m$个输入(当输出方程为式(25)时, 为n个输入), p个输出对应k时刻系统的输出.

在用LS–SVM建立非线性模型时, 选取RBF函数为核函数

$$K(\boldsymbol{r},\boldsymbol{r}_i)=\exp(-\frac{\|\boldsymbol{r}-\boldsymbol{r}_i\|_2^2}{2\sigma^2}).\tag{45}$$

其中: σ为RBF核的宽度, 需事先确定; $\boldsymbol{r}_i$为获得的支持向量:

$$\boldsymbol{r}(k)=[x_1(k),\cdots,x_n(k),u_1(k),\cdots,u_m(k)],\tag{46}$$

$$\boldsymbol{r}_i=[x_{1\alpha},\cdots,x_{n\alpha},u_{1\alpha},\cdots,u_{m\alpha}].\tag{47}$$

这时核函数(45)可写为

$$K(\boldsymbol{r},\boldsymbol{r}_i)=\exp(-\frac{\|\boldsymbol{r}-\boldsymbol{r}_i\|_2^2}{2\sigma^2})=\exp(-\frac{\sum\limits_{q=1}^{n}(x_q-x_{q\alpha})^2+\sum\limits_{s=1}^{m}(u_s-u_{s\alpha})^2}{2\sigma^2}).\tag{48}$$

因此, 回归函数o对输入$x_i(k)$和$u_i(k)$的偏导数分别为

$$\frac{\partial o}{\partial x_i}=-\frac{1}{\sigma^2}\cdot\sum_{q=1}^{n}(x_q(k)-x_{q\alpha})\cdot\sum_{i=1}^{m}\alpha_i\exp(-\tfrac{H}{2\sigma^2}),\tag{49}$$

$$\frac{\partial o}{\partial u_i}=-\frac{1}{\sigma^2}\cdot\sum_{s=1}^{m}(u_q(k)-u_{q\alpha})\cdot\sum_{i=1}^{m}\alpha_i\exp(-\tfrac{H}{2\sigma^2}).\tag{50}$$

其中

$$H=\sum_{q=1}^{n}(x_q(k)-x_{q\alpha})^2+\sum_{s=1}^{m}(u_s(k)-u_{s\alpha})^2.$$

上式确定了LS–SVM模型输出对输入的微分关系, 因此可以用来计算式(21)~(24)中对$\boldsymbol{x}$和$\boldsymbol{u}$的梯度, 从而可以确定$\boldsymbol{a}_j$,$\boldsymbol{b}_j$和$\boldsymbol{A}$, $\boldsymbol{B}$. 同理, 式(2)中的系数矩阵$\boldsymbol{C}$和$\boldsymbol{D}$也可以用上述同样的方法确定.

4 仿真实验(Simulation experiments)

利用LS–SVM良好的非线性函数逼近能力, 建立起非线性离散系统在工作点处的最优局部线性化状态空间模型, 并在该工作点设计线性状态空间控制器, 可以完成对非线性离散系统的有效控制, 控制方法的框图如图1所示. 本文通过对两个典型非线性离散系统的仿真, 验证了该方法的有效性.

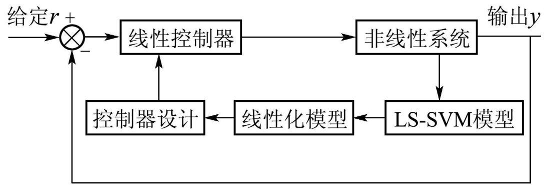

图 1 基于LS–SVM的非线性离散系统控制框图

Fig. 1 Diagram of nonlinear system control based on LS-SVM

算例 1 考虑如下的2阶非线性单输入单输出系统:

$$\begin{cases}x_1(k+1)=\dfrac{x_2^3(k)+x_2(k)}{1+2x_2^2(k)},\\ x_2(k+1)=\dfrac{x_1(k)+x_2(k)}{1+x_1^2(k)}+u(k),\\ y(k)=x_1(k).\end{cases}\tag{51}$$

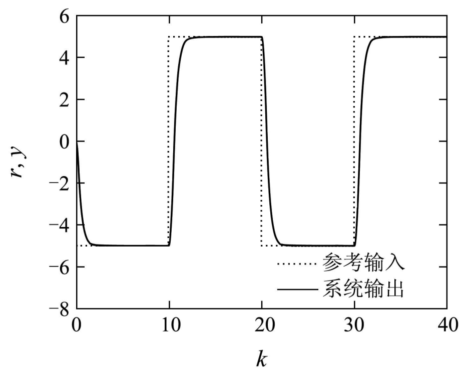

图 2(a) 参考信号为方波时系统的输出

Fig. 2(a) System output when reference input is a square wave

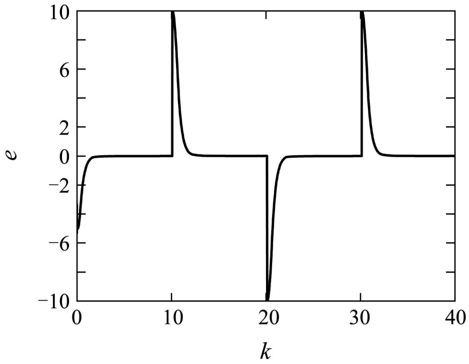

图 2(b)　参考信号为方波时的误差曲线
Fig. 2(b)　Error when reference input is a square wave

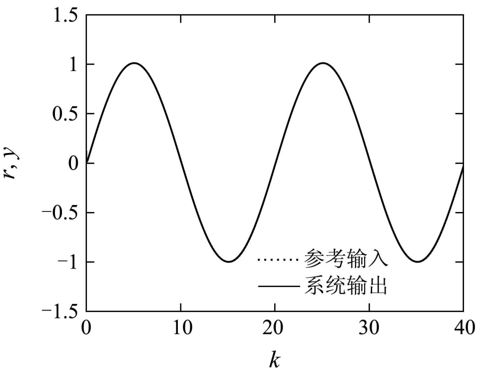

图 2(c)　参考信号为正弦波时系统的输出
Fig. 2(c)　System output when reference input is a sinusoid

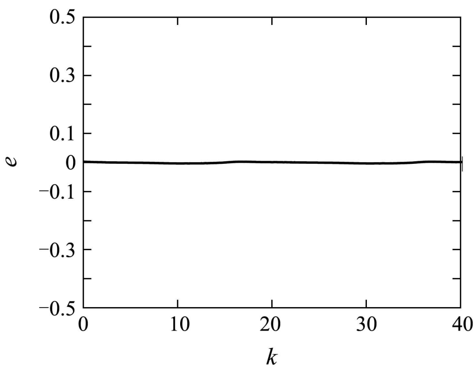

图 2(d)　参考信号为正弦波时的误差曲线
Fig. 2(d)　Error when reference input is a sinusoid

首先确定系统的2个状态和1个输入的取值范围, 对由状态和输入构成的三维空间离散化并随机采样, 每一个采样点构成训练向量的输入, 其中前两个分量为状态值, 最后的分量为系统的控制输入; 训练向量的输出由系统的下一个状态构成或系统的输出构成, 如式(43)(44)所示. 在本例中共获取120组输入/输出向量, 其中100组用于训练, 20组用于测试. 控制目标是控制系统(51)跟踪参考输入为一个幅值为±5的方波及幅值为1 的正弦波的变化. 用LS–SVM完成对非线性系统的局部线性化, 核函数选取RBF核函数, 核参数通过交叉检验的方法得到, $\gamma = 17$, $\sigma^2 = 0.2$; 采用极点配置的方法完成对线性化以后系统的控制, 设定的极点被配置在0.2和0.4. 图2(a)为参考输入为方波时系统的闭环输出信号, 图2(b)为相应的误差信号, 图2(c)为参考输入为正弦波时系统的闭环输出信号, 图2(d)为相应的误差信号. 由图可知, 系统对于两种参考输入信号具有很好的跟踪效果, 并且由于支持向量机好的逼近能力, 跟踪误差很小, 如图2(b)、图2(d)所示, 系统进入稳态后最小跟踪误差分别为0.00627, 0.00817.

算例 2　考虑如下的3阶非线性单输入单输出系统[11]:

$$\begin{cases} x_1(k+1) = 0.5x_3(k), \\ x_2(k+1) = x_1(k) + [1 - 0.4x_2(k)]u(k), \\ x_3(k+1) = 0.3x_1(k)x_3(k) - x_2(k), \\ y(k) = x_1(k). \end{cases} \tag{52}$$

样本的获取方法同上, 在本例中共获取270组输入/输出向量, 其中220组用于训练, 50组用于测试.

本例的控制目标是控制系统(52)跟踪参考输入为一个幅值为±1的方波及幅值为1 的正弦波的变化. 用LS–SVM完成对非线性系统的局部线性化, 核函数选取RBF核函数, 核参数通过交叉检验的方法得到, $\gamma = 40$, $\sigma^2 = 0.2$; 采用极点配置的方法来完成对线性化以后系统的控制, 设定的极点被配置在0.4,0.5和0.5. 图3(a)为参考输入为方波时系统的闭环输出信号, 图3(b)为相应的误差信号, 图3(c)为参考输入为正弦波时系统的闭环输出信号, 图3(d)为相应的误差信号. 由图可知, 系统对于两种参考输入信号的跟踪都获得了满意的效果, 并且由于支持向量机良好的逼近能力, 跟踪误差很小, 如图3(b)、图3(d)所示, 系统进入稳态后最小跟踪误差为0.00671, 0.01017.

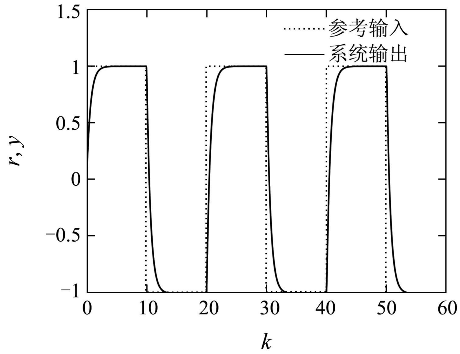

图 3(a)　参考信号为方波时系统的输出
Fig. 3(a)　System output when reference input is a square wave

文献[12]采用前馈神经网络辨识非线性对象的模型, 并在工作点处将其线性化, 然后采用极点配置技术、最优控制技术对非线性系统进行控制, 取得了很好的结果. 本文也采用前馈神经网络辨识的方法对上述两个例子分别进行了仿真, 参考输入信号分别选取相同的方波输入信号. 虽然可以获得对

参考输入信号很好的跟踪, 但是跟踪误差总有一个较小的偏置量, 算例1和算例2的最小跟踪误差分别为0.10325和0.14731. 可看出, 由于神经网络本身泛化能力的限制, 导致建模误差较大, 使得系统跟踪参考输入信号总有一定误差, 要消除这样的误差, 在控制器的设计上必须考虑在前向通道增加积分器.

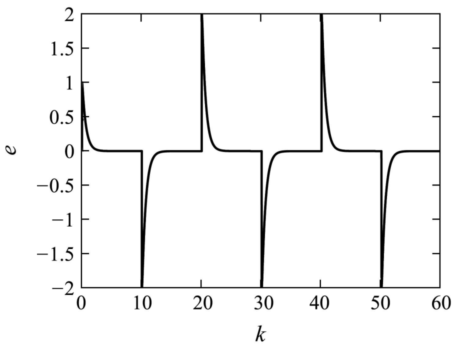

图 3(b) 参考信号为方波时的误差曲线

Fig. 3(b) Error when reference input is a square wave

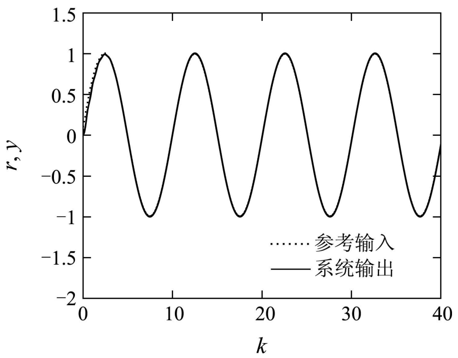

图 3(c) 参考信号为正弦波时系统的输出

Fig. 3(c) System output when reference input is a sinusoid

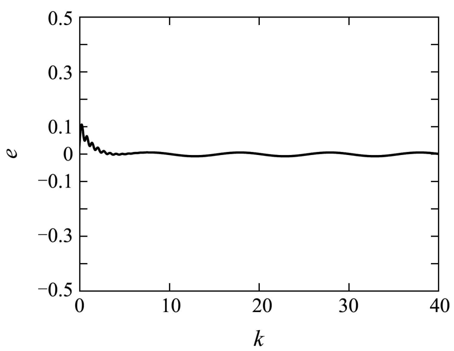

图 3(d) 参考信号为正弦波时的误差曲线

Fig. 3(d) Error when reference input is a sinusoid

5 结论(Conclusions)

本文提出了一种基于支持向量机回归的非线性离散系统控制新方法. 非线性系统的状态空间模型在工作点处用LS-SVM的方法进行线性化, 并对系统设计合适的线性控制器施加控制, 对两个典型非线性系统的仿真实验验证了该方法的有效性. 该方法不依赖于非线性系统的物理模型, 所获得的系统局部线性化模型结构固定, 可以使用任何传统的线性控制方法来控制, 计算复杂度小.

当然, 目前非线性离散动力学系统的控制和连续系统控制相比还有较多的困难, 比如: 为保证系统的稳定, 离散系统如何激励的问题; 当系统的相对阶大于1时, 当前输入将依赖于系统的未来状态, 即非因果的问题; 对多输入/多输出系统的解耦问题等等. 采用支持向量机研究非线性离散动力学系统的控制, 同样要面对上述的问题并且有新的表现形式, 这也是进一步需要研究和解决的.

参考文献(References):

[1] HORNIK K, STINCHCOMBE M, WHITE H. Multilayer feedforward networks are universal approximators[J]. *Neural Networks*, 1989, 2(5): 359 – 366.

[2] HUNT K J, SBARBARO D, ZBIKOWSKI R, et al. Neural networks for control systems-survey[J]. *Automatica*, 1992, 28(6): 1083 – 1112.

[3] AGARWAL M. A systematic classification of neural-network based control[J]. *IEEE Control Systems Magazine*, 1997, 17(2): 75 – 93.

[4] 孙富春, 李莉, 孙增圻. 非线性系统神经网络自适应控制的发展现状及展望[J]. 控制理论与应用, 2005, 22(2): 254 – 260.
(SUN Fuchun, LI Li, SUN Zengqi. Survey on adaptive control of nonlinear s ystems using neural networks[J]. *Control Theory & Applications*, 2005, 22(2): 254 – 260.)

[5] VAPNIK V. *The Nature of Statistical Learning Theory*[M]. New York: Springer-Verlag, 1999.

[6] 张学工. 关于统计学习理论和支持向量机[J]. 自动化学报, 2000, 26(1): 32 – 42.
(ZHANG Xuegong. Introduction to statistical learning theory and support vector machines[J]. *Acta Automatica Sinica*, 2000, 26(1): 32 – 42.)

[7] LIN C T, YEH C M, LIANG S F, CHUNG J F and KUMAR N. Support-vector-based fuzzy neural network for pattern classification[J]. *IEEE Transactions on Fuzzy Systems*, 2006, 14(1): 31 – 41.

[8] 刘涵, 刘丁. 基于模糊sigmoid核的支持向量回归建模[J]. 控制理论与应用, 2006, 23(2): 204 – 208.
(LIU Han, LIU Ding. Support vector regression based on fuzzy sigmoid kernel[J]. *Control Theory & Applications*, 2006, 23(2): 204 – 208.)

[9] 刘涵, 刘丁, 任海鹏. 基于最小二乘支持向量机的混沌控制[J]. 物理学报, 2005, 54(9): 4019 – 4025.
(LIU Han, LIU Ding, REN Haipeng. Chaos control based on least square support vector machines[J]. *Acta Physica Sinica*, 2005, 54(9): 4019 – 4025.)

[10] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifiers[J]. *Neural Processing Letter*, 1999, 9(3): 293 – 300.

[11] LEVIN A U, NARENDRA K S. Control of dynamical systems using neural networks-Part Ⅱ: Observability, identification and control[J]. *IEEE Transactions on Neural Networks*, 1996, 7(1): 30 – 42.

[12] AHAMED M S, TASADDUQ I A. Neural-net controller for nonlinear plants: design approach through linearization[J]. *IEE Proceedings: Control Theory and Applications*, 1994, 141(5): 315 – 322.

作者简介:

刘 丁 (1957—), 男, 工学博士, 西安理工大学教授, 博士生导师, 研究方向为智能控制、复杂系统建模等, E-mail: liud@xaut.edu.cn;

刘 涵 (1972—), 男, 工学博士, 西安理工大学副教授, 研究方向为复杂系统建模与控制、机器学习、智能信息处理等, E-mail: liuhan@xaut.edu.cn.