文章编号: 1002-0411(2007) 04-0451-04

基于最小二乘支持向量机的 T-S模型在线辨识

丁学明
(上海理工大学计算机与电气工程学院, 上海 200093)

摘 要: 提出一种基于时间窗最小二乘支持向量机的 T-S模型在线辨识算法, 包括结构辨识和参数辨识. 该算法以时间窗内数据的势能作为结构辨识依据, 同时采用最小二乘支持向量机辨识系统参数, 具有辨识速度快、精度高的特点. 仿真结果证明了算法的有效性.

关键词: T-S模型; 时间窗; 势能; 最小二乘支持向量机

中图分类号: TP273.4　　**文献标识码**: A

LSSVM-Based Online Identification for T-S Model

DING Xue-ming
(*College of Computer and Electrical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China*)

Abstract: An online identification method based on least square support vector machine (LSSVM) with time window for T-S model is proposed, including structure and parameter identifications. Structure identification depends on the potential of data in the time window, and parameter identification is based on LSSVM, resulting in high identification speed and precision. The simulation result illustrates the effectiveness of the proposed method.

Keywords: T-S model; time window; potential; least square support vector machine (LSSVM)

1 引言(Introduction)

实际系统通常与工作点附近的数据有较大的相关性, 而与远离工作点区域的数据的相关性并不是很大. 系统往往在工作了一段时间后, 工作域会发生迁移, 用固定的数据建立的模型随着时间和条件的变化会与实际情况有些不吻合[1]. T-S模糊模型是一个通用逼近器[2], 它把一个非线性系统当作多个线性子系统与其权重乘积之和, 易于表达复杂非线性系统的动态特性, 同时可以将线性控制理论应用到非线性系统控制中, 从而成为研究热点. T-S模型辨识包括结构辨识和参数辨识, 结构辨识用于辨识 T-S模型的前提模糊规则, 参数辨识则用于确定规则结论部分线性参数.

本文提出基于时间窗[1]的最小二乘支持向量机 T-S模型在线辨识, 结构辨识采用聚类方法, 每个类代表一条规则, 规则数等于类数量, 类中心作为该规则的隶属度中心参数, 系统辨识从当前起到过去的 l组记录数据(时间窗)中得到, 而时间窗内数据的势能[3,4]作为判断模糊聚类中心的依据.

T-S模型的参数辨识采用最小二乘支持向量机方法. 支持向量机(Support Vector Machine, SVM)是由 Vapnik等人在统计学习理论基础上建立的一种基于数据的机器学习方法[5], 它是受约束的二次规划问题, 计算复杂度大. 为减小算法复杂度, Suykens[6]提出最小二乘支持向量机(LSSVM), 它把支持向量机的学习问题转化成解线性方程组的问题, 因此运算速度较快, 在函数估计和逼近中得到了广泛应用.

2 T-S模型(T-S model)

T-S模型[7]是在 1985年提出来的, 它把一个非线性系统当作多个线性子系统与其权重乘积之和, 可表示为:

$R^{(i)}$: 如果 $\boldsymbol{x}$ 是 $\boldsymbol{F}_i$, 则

$$f(\boldsymbol{x}) = \boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x} + b_i \quad i = 1, 2, \cdots, m \tag{1}$$

其中 $\boldsymbol{x}\in \mathrm{R}^n$ 是输入变量. $\boldsymbol{w}_i^{\mathrm{T}}\in \mathrm{R}^n$ 和 $b_i\in \mathrm{R}$ 是第 i条规则的参数. 设第 i条规则模糊集 $\boldsymbol{F}_i$ 的隶属度函数

*收稿日期: 2006-10-08
基金项目: 上海高校选拔培养优秀青年教师科研专项基金资助项目(21020); 上海理工大学博士科研启动基金资助项目(X641)

为 $h_i(\boldsymbol{x})$. 采用高斯隶属度函数, 直积运算采用求积法, 则

$$h_i(\boldsymbol{x}_j)=\prod_{t=1}^{n}\exp[-\alpha(x_{jt}-v_{it})^2] \tag{2}$$

其中 $\boldsymbol{x}_j$ 表示第 j 个输入数据, x_{jt} 表示第 j 个输入数据的第 t 个分量, v_{it} 是类 i 的第 t 个分量, $\alpha=4/r^2$, r 是分布半径.

定义以下归一化处理隶属度函数

$$\mu_{ij}(\boldsymbol{x}_j)=\frac{h_i(\boldsymbol{x}_j)}{\sum_{i=1}^{m}h_i(\boldsymbol{x}_j)} \tag{3}$$

因为 $h_i(\boldsymbol{x}_j)\geqslant 0$, $\sum_{i=1}^{m}h_i(\boldsymbol{x}_j)>0$, 所以 $0\leqslant\mu_{ij}(\boldsymbol{x}_j)\leqslant 1$ 且 $\sum_{i=1}^{m}\mu_{ij}(\boldsymbol{x}_j)=1$.

模糊化采用单点模糊集合, 清晰化采用加权平均法, 则可得到全局系统状态方程

$$f(\boldsymbol{x}_j)=\sum_{i=1}^{m}\mu_{ij}(\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x}_j+b_i) \tag{4}$$

3 结构辨识 (Structure identification)

模糊辨识包括结构辨识和参数辨识, 结构辨识用于辨识 T-S 模型的前提模糊规则, 参数辨识则用于确定规则结论部分的线性参数. 当采集的数据量 $k\leqslant l$ 时, 以所有过去数据来辨识 T-S 模型, 当 $k>l$ 时, 利用从当前到过去的 l 组数据辨识模型, 建立一个随时间滚动的建模数据区间, 并保持该区间长度 l 不变. 当有一个新数据加入时, 最早的一个数据相应地从 l 区间滚动出去, 这个滚动的数据区间随时间变化, 故称这个数据区间为滚动时间窗[1].

文[3, 4]的辨识基于所有过去数据, 本文在其 SCM (subtractive clustering method) 辨识基础上, 提出了基于时间窗的 T-S 模型在线结构辨识. 改进在于:

1) 文[3, 4]中, 只有最近采集数据的势能最大时, 它才成为类中心. 而本文中只要数据在时间窗内, 其势能最大, 就成为类中心, 不管其是否为最近采集的数据. 因为有些数据虽然刚采集时势能不是最大, 但随着时间窗的滚动, 数据的势能不断变化, 一些非类中心的数据, 可能由于其势能最大而变为类中心.

2) 本文基于时间窗内的数据进行辨识, 故类中心点仅限于时间窗内的数据, 而文[3, 4]中辨识是基于所有过去的数据, 类中心点在所有过去数据中选取.

$\boldsymbol{x}_k$ 势能作为判断模糊聚类中心的依据, 其定义如下

$$P_k(\boldsymbol{x}_k)=\begin{cases}\sum_{j=1}^{k}\exp(-\alpha\|\boldsymbol{x}_k-\boldsymbol{x}_j\|^2) & k\leqslant l\\ \sum_{j=k-l+1}^{k}\exp(-\alpha\|\boldsymbol{x}_k-\boldsymbol{x}_j\|^2) & k>l\end{cases} \tag{5}$$

其中 $\boldsymbol{x}_k$ 是第 k 个采集的数据, k 是采集数据的序号, l 是时间窗宽度, 即辨识的数据数量. 此时时间窗内其它数据的势能更新为:

$$P_k(\boldsymbol{x}_j)=\begin{cases}p_{k-1}(\boldsymbol{x}_j)+\exp(-\alpha\|\boldsymbol{x}_j-\boldsymbol{x}_k\|^2) & j=1,\cdots,k-1,\ k\leqslant l\\ p_{k-1}(\boldsymbol{x}_j)+\exp(-\alpha\|\boldsymbol{x}_j-\boldsymbol{x}_k\|^2)-\exp(-\alpha\|\boldsymbol{x}_j-\boldsymbol{x}_{k-l}\|^2) & j=k-l+1,\cdots,k-1,\ k>l\end{cases} \tag{6}$$

辨识步骤:

1) 初始化

给定参数 r、α, 设第一个数据 $\boldsymbol{x}_1$ 为类中心 $\boldsymbol{v}_1$, 其势能 $P_1(\boldsymbol{x}_1)=1$, 类数量 $m=1$, 数据数量 $k=k+1$.

2) 滚动时间窗, 计算势能

采集数据 $\boldsymbol{x}_k$, 滚动时间窗, 按式(5)计算 $\boldsymbol{x}_k$ 势能, 按式(6)更新时间窗内其它数据势能.

如果 $k>l$ 且 $\boldsymbol{x}_{k-l}$ 为类中心, 则从类中心删除 $\boldsymbol{x}_{k-l}$, 即调整类序号, 假设 $\boldsymbol{x}_{k-l}$ 是类中心 $\boldsymbol{v}_i$, $\boldsymbol{v}_q=\boldsymbol{v}_{q+1}$, $q=i,\cdots,m-1$, 类数量 $m=m-1$.

3) 类中心的增加和替代

$$p_k(\boldsymbol{x}_i)=\begin{cases}\max\{p_k(\boldsymbol{x}_j),\ j=1,\cdots,k\} & k\leqslant l\\ \max\{p_k(\boldsymbol{x}_j),\ j=k-l+1,\cdots,k\} & k>l\end{cases} \tag{7}$$

如果 $\boldsymbol{x}_i$ 不是类中心, 则令 $\delta_{\min}$ 为:

$$\delta_{\min}=\min\{\exp(-\alpha\|\boldsymbol{x}_i-\boldsymbol{v}_j\|),\quad j=1,\cdots,m\} \tag{8}$$

设 $\boldsymbol{v}_j$ 是离 $\boldsymbol{x}_i$ 最近的类中心, 如果 $\frac{\delta_{\min}}{r}+\frac{p_k(\boldsymbol{v}_j)}{p_k(\boldsymbol{x}_i)}<1$, 则 $\boldsymbol{x}_i$ 替代类中心 $\boldsymbol{v}_j$, 即: $\boldsymbol{v}_j=\boldsymbol{x}_i$. 否则增加 $\boldsymbol{x}_i$ 为类中

心, 即: $m=m+1$, $\boldsymbol{v}_m=\boldsymbol{x}_i$.

4) 删除类中心

$$d_{\min}=\min\{\exp(-\alpha\|\boldsymbol{v}_i-\boldsymbol{v}_j\|),\quad i=1,\cdots,m-1,\ j=2,\cdots,m\} \tag{9}$$

$$p_{\max}=\max\{p_k(\boldsymbol{v}_q),\quad q=1,\cdots,m\} \tag{10}$$

设 $\boldsymbol{v}_i$ 和 $\boldsymbol{v}_j$ 是距离最近的两个类中心, 并设 $p_k(\boldsymbol{v}_i)<p_k(\boldsymbol{v}_j)$. 如果 $\frac{d_{\min}}{r}+\frac{p_k(\boldsymbol{v}_i)}{p_{\max}}<1$, 则删除类中心 $\boldsymbol{v}_i$, 即调整类序号 $\boldsymbol{v}_q=\boldsymbol{v}_{q+1}$, $q=i,\cdots,m-1$, 类数量 $m=m-1$.

5) $k=k+1$, 返回步骤 2), 直至辨识结束.

4 参数辨识 (Parameter identification)

需辨识的 T-S模型为:

$$f(\boldsymbol{x}_j)=\sum_{i=1}^{m}\mu_{ij}(\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x}_j+b_i) \tag{11}$$

其中: $\boldsymbol{w}_i=[w_{i1}\ \cdots\ w_{in}]^{\mathrm{T}}$, $\boldsymbol{x}_j=[x_{j1}\ \cdots\ x_{jn}]^{\mathrm{T}}$, n 为输入变量维数.

根据结构风险最小化原理, 综合考虑函数复杂度和拟合误差, 回归问题可以表示为约束优化问题:

$$\begin{aligned}&\min_{\boldsymbol{w}_i,b_i,e}\frac{1}{2}\sum_{i=1}^{m}\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{w}_i+\frac{C}{2}\sum_{j=k-l+1}^{k}e_j^2\\&\mathrm{s.t.}\ \ e_j=y_j-\sum_{i=1}^{m}\mu_{ij}(\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x}_j+b_i)\end{aligned} \tag{12}$$

为了求解上述优化问题, 把约束优化问题转化成无约束的优化问题, 构造拉格朗日方程:

$$\begin{aligned}L=&\frac{1}{2}\sum_{i=1}^{m}\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{w}_i+\frac{C}{2}\sum_{j=k-l+1}^{k}e_j^2\\&+\sum_{j=k-l+1}^{k}\alpha_j[y_j-\sum_{i=1}^{m}\mu_{ij}(\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x}_j+b_i)]\end{aligned} \tag{13}$$

根据 KKT条件有

$$\begin{aligned}&\frac{\partial L}{\partial \boldsymbol{w}_i}=0\rightarrow\ \boldsymbol{w}_i=\sum_{j=k-l+1}^{k}\alpha_j\mu_{ij}\boldsymbol{x}_j\\&\frac{\partial L}{\partial b_i}=0\rightarrow\sum_{j=k-l+1}^{k}\alpha_j\mu_{ij}=0\\&\frac{\partial L}{\partial e_j}=0\rightarrow\ \alpha_j=Ce_j\\&\frac{\partial L}{\partial \alpha_j}=0\rightarrow\sum_{i=1}^{m}\mu_{ij}(\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x}_j+b_i)+e_j-y_j=0\end{aligned} \tag{14}$$

从式 (14)方程组中消去 e_j、$\boldsymbol{w}_i$, 可得到

$$\begin{vmatrix}0&\cdots&0&\mu_{1,k-l+1}&\cdots&\mu_{1,k}\\\vdots&\cdots&\vdots&\vdots&\cdots&\vdots\\0&\cdots&0&\mu_{m,k-l+1}&\cdots&\mu_{m,k}\\\mu_{1,k-l+1}&\cdots&\mu_{m,k-l+1}&\sum_{i=1}^{m}\mu_{i,k-l+1}\mu_{i,k-l+1}Q_{k-l+1,k-l+1}+\frac{1}{C}&\cdots&\sum_{i=1}^{m}\mu_{i,k-l+1}\mu_{i,k}Q_{k,k-l+1}\\\vdots&\cdots&\vdots&\vdots&\cdots&\vdots\\\mu_{1,k}&\cdots&\mu_{m,k}&\sum_{i=1}^{m}\mu_{i,k}\mu_{i,k-l+1}Q_{k-l+1,k}&\cdots&\sum_{i=1}^{m}\mu_{i,k}\mu_{i,k}Q_{k,k}+\frac{1}{C}\end{vmatrix}\begin{vmatrix}b_1\\\vdots\\b_m\\\alpha_{k-l+1}\\\vdots\\\alpha_k\end{vmatrix}=\begin{vmatrix}0\\\vdots\\0\\y_{k-l+1}\\\vdots\\y_k\end{vmatrix} \tag{15}$$

其中: $Q_{i,j}=\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j$, $i=1,\cdots,m$, $j=k-l+1,\cdots,k$.

由式 (15)可以看出, 最小二乘支持向量机的训练问题归结为一个线性方程组的求解问题, 而不像标准支持向量机那样要求一个二次规划问题, 解线性方程组比求解二次规划要简单得多.

注: 1) 式 (12) ~ (15)是 $k>l$时的表达形式, 开始辨识时, 时间窗内的数据量 $k\leqslant l$, 此时式 (12)为 $\min_{\boldsymbol{w}_i,b_i,e}\frac{1}{2}\sum_{i=1}^{m}\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{w}_i+\frac{C}{2}\sum_{j=1}^{k}e_j^2$, 式 (13) ~ (15)作相应的变动.

2) 式 (15)中 μ下标中的 “, ”号仅仅用于分隔前后下标, 与其它没有 “, ”号 μ的前后下标含义相同.

5 仿真实例 (Simulation example)

本文辨识的对象是下列二阶系统:

$$y(k+1)=\frac{y(k)y(k-1)[y(k)+2.5]}{1+y^2(k)+y^2(k-1)}+u(k) \tag{16}$$

设输入变量为 $\boldsymbol{x}_k = [y(k)\ y(k-1)\ u(k)]^{\mathrm{T}}$, 输出变量 $y_k = y(k+1)$. 控制信号 $u(k)= 0.5\sin(\pi k/3) + 0.5\sin(\pi k/5)$.

采用第 3、4节的在线辨识算法, 令 $C = 100$、$l= 10$、$r= 2$、$\alpha= 1$, 得到下述仿真曲线, 辨识精度高、误差小, 均方差为 0.0543.

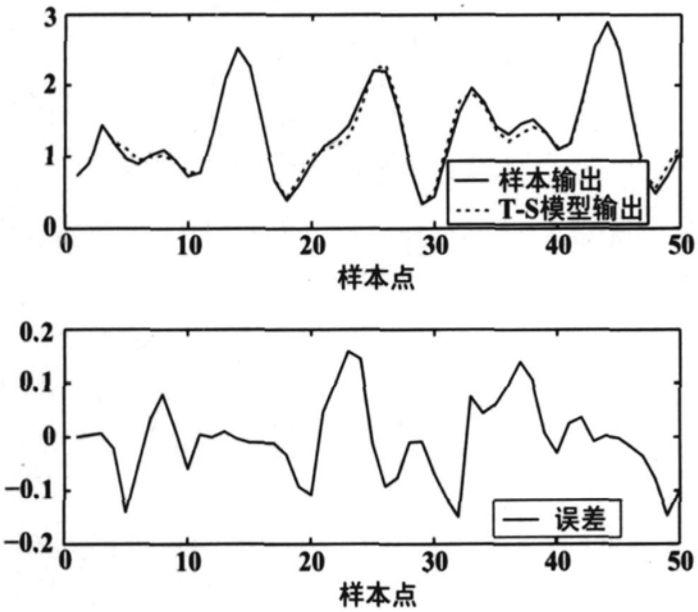

图 1 样本、T-S模型及其误差输出曲线

Fig. 1 Output curves of sample, T-S model and error

6 结论 (Conclusion)

T-S模糊模型是一个通用逼近器, 它把一个非线性系统当作多个线性子系统与其权重乘积之和. 本文提出了基于时间窗的最小二乘支持向量机T-S模型在线辨识算法, 包括结构辨识和参数辨识. 随着时间窗滚动, 不断调整数据势能, 作为结构辨识的依据, 实现系统的在线结构辨识. 同时利用最小二乘支持向量机辨识系统参数, 辨识精度高、泛化能力强. 仿真结果证明了算法的有效性.

参 考 文 献 (References)

[1] 阎威武, 常俊林, 邵惠鹤. 基于滚动时间窗的最小二乘支持向量机回归估计方法及仿真 [J]. 上海交通大学学报, 2004, 38(4): 524~526, 532.

[2] Buckley J J. Sugeno type controllers are universal controllers [J]. Fuzzy Sets and Systems, 1993, 53(3): 299.

[3] Chiu S L. Fuzzy model identification based on cluster estimation [J]. Journal of Intelligent and Fuzzy Systems, 1994, 2(3): 267~278.

[4] Qi R Y, Brdys M A. Adaptive fuzzy modelling and control for discrete-time nonlinear uncertain systems [A]. Proceedings of the 2005 American Control Conference [C]. Piscataway, NJ, USA: IEEE, 2005. 1108~1113.

[5] Vapnik V N. The Nature of Statistical Learning Theory [M]. New York, USA: Springer-Verlag, 1995.

[6] Suykens J A K, Vandewale J. Least squares support vector machine classifiers [J]. Neural Processing Letters, 1999, 9(3): 293~300.

[7] Takagi T, Sugeno M. Fuzzy identification of systems and its application to modeling and control [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1985, SMC-15(1): 116~132.

作者简介

丁学明 (1971-), 男, 博士, 讲师. 研究领域为智能控制, 运动控制, 嵌入式系统.