

## 在线鲁棒最小二乘支持向量机回归建模

张淑宁<sup>1</sup>, 王福利<sup>1,2</sup>, 何大阔<sup>1,2</sup>, 贾润达<sup>1</sup>

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110004;

2. 东北大学 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110004)

**摘要:** 鉴于工业过程的时变特性以及现场采集的数据通常具有非线性特性且包含离群点, 利用最小二乘支持向量机回归(least squares support vector regression, LSSVR)建模易受离群点的影响. 针对这一问题, 结合鲁棒学习算法(robust learning algorithm, RLA), 本文提出了一种在线鲁棒最小二乘支持向量机回归建模方法. 该方法首先利用LSSVR模型对过程输出进行预测, 与真实输出相比较得到预测误差; 然后利用RLA方法训练LSSVR模型的权值, 建立鲁棒LSSVR模型; 最后应用增量学习方法在线更新鲁棒LSSVR模型, 从而得到在线鲁棒LSSVR模型. 仿真验证了所提方法的有效性.

**关键词:** 鲁棒学习算法; 最小二乘支持向量机; 鲁棒性; 非线性

**中图分类号:** TP273 **文献标识码:** A

## Modeling method of online robust least-squares-support-vector regression

ZHANG Shu-ning<sup>1</sup>, WANG Fu-li<sup>1,2</sup>, HE Da-kuo<sup>1,2</sup>, JIA Run-da<sup>1</sup>

(1. School of Information Science & Engineering, Northeastern University, Shenyang Liaoning 110004, China;

2. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang Liaoning 110004, China)

**Abstract:** Industrial processes possess time-varying feature, and data from industrial field usually possess nonlinear feature and contain outliers. Modeling with least-squares-support-vector regression(LSSVR) method may suffer from these outliers. To deal with this problem, we develop an online robust LSSVR method by combining with the robust learning algorithm(RLA). The LSSVR model is used to predict process outputs, and the residuals are formed from real outputs and predicted outputs. The RLA trains the weights of LSSVR model iteratively. The trained robust LSSVR model is then updated by means of incremental updating algorithm. An online robust LSSVR model is also developed. Simulation results show the effectiveness of the proposed approach.

**Key words:** robust learning algorithm; least-squares-support-vector machine; robustness; nonlinear

### 1 引言(Introduction)

实际工业过程中, 由于工艺和技术等方面的限制, 许多重要的过程变量难以实时在线检测. 而软测量技术则是实现此类变量实时估计的一种非常有效的方法, 其核心问题是建立软测量的数学模型. 机理建模、数据建模和混合建模方法是常用的软测量建模方法. 数据建模方法能够充分利用采集到的过程数据信息, 具有较高的预测精度. 因此, 足够丰富的过程数据是数据建模方法的基础. 然而实际工业过程中采集到的数据非常有限, 并且由于过程干扰、人为因素造成数据异常, 使得许多数据建模方法的应用范围受到限制, 因此, 研究小样本建模方法具有重要意义.

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, 而基于此理论的支持向量机(support vector machines, SVM)方法<sup>[1,2]</sup>因其具有非线性拟合、结构简单、全局最优和较好泛化能力等优点, 广泛应用于许多领域. 然而, SVM每增加一个新的样本, 需进行一次二次优化求解, 运算过于复杂且不具有实时性.

为了提高求解速度, Suykens等人提出了最小二乘支持向量机回归(least-squares-support-vector regression, LSSVR)<sup>[3]</sup>. 它使用等式约束代替了不等式约束, 将二次规划问题转化为线性方程组求解, 简化计算的复杂性. 但是LSSVR算法是在假设所有训练

样本相互独立分布的情况下得到的,然而在实际工程中,由于现场干扰以及过失差错等原因,离群数据的存在是在所难免的.此时,LSSVR采用的误差平方和评价函数会导致不稳健的估计结果.

针对采样样本中存在离群点问题,文献[4~7]对离群点的影响进行了研究.结果表明离群点对模型精度有很大影响,特别是小样本建模问题,这种影响尤为突出.因此,由于离群点影响,使LSSVR模型在实际过程中的应用受到很大限制.针对此问题,文献[8]通过对样本加权来减小离群点的影响,提出了加权最小二乘支持向量机(WLSSVR).文献[9]应用模糊聚类的思想对每个样本进行加权,使模型获得鲁棒估计.二者都考虑应用对样本加权的思想减小离群点的影响来获得鲁棒性,尤其对显著离群点进行了相应处理,但是因为训练过程中模型的权重是一个常数,且易受到LSSVR参数选择的影响,所以上述两种方法对非显著离群点的处理显得无能为力.因此,文献[10]利用鲁棒学习(robust learning algorithm, RLA)提出了鲁棒RBF网络来解决这个问题,然而RBF网络的初始权值和结构较难确定.

针对上述问题,并且考虑到工业过程的时滞性以及过程动态变化特点,结合LSSVR结构容易确定的优点,本文提出了一种在线鲁棒最小二乘支持向量机回归建模方法.该方法首先在LSSVR模型基础上,通过RLA克服离群点影响来获得鲁棒估计,为了能更好克服非显著离群点影响,本文对RLA进行了改进,提出了利用离群点(outliers)检测方法的鲁棒代价函数端点确定的方法.接下来应用增量学习方法更新鲁棒LSSVR模型,达到实时跟踪过程动态特性的目的.仿真分析验证了该方法的可行性和有效性.

## 2 最小二乘支持向量机回归(Least-squares-support-vector regression)

SVM回归方法首先把原始的输入空间非线性映射到一个高维甚至无穷维的特征空间,然后在这个特征空间中寻找使得结构风险最小的线性决策函数<sup>[11]</sup>.对一组输入样本,LSSVR通过非线性变换将训练样本集映射到高维特征空间.然后在这个变换后的特征空间中应用某一线性函数 $g(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x})$ 来对函数进行估计.在LSSVR回归估计中,求解下面的优化问题:

$$\begin{aligned} \min J(\mathbf{w}, \mathbf{e}, b) &= \frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{2} \sum_{i=1}^n e_i^2, \\ \text{s.t. } y_i &= \mathbf{w}^T \varphi(\mathbf{x}_i) + b + e_i, \\ i &= 1, 2, \dots, n, \end{aligned} \quad (1)$$

其中:  $J(\mathbf{w}, \mathbf{e}, b)$  为目标函数,  $C$  为惩罚系数, 其值的大小决定了对误差的惩罚力度,  $\mathbf{e}^T = [e_1 \ e_2 \ \dots \ e_n]$  为容许误差向量.

应用Lagrange法可以将式(1)的二次规划问题化简为求解线性方程组问题,即

$$\begin{bmatrix} K + \frac{1}{C} I_n & \mathbf{1}_n \\ \mathbf{1}_n^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \quad (2)$$

其中:  $\mathbf{1}_n^T = [1 \ \dots \ 1]$ ,  $\mathbf{y}^T = [y_1 \ \dots \ y_n]$ ,  $I_n$  为  $n$  维单位阵,  $\boldsymbol{\alpha}^T = [\alpha_1 \ \dots \ \alpha_n]$  为Lagrange乘子向量,  $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j)$ . 求解出式(2)的  $\boldsymbol{\alpha}$  和  $b$  得到LSSVR模型为

$$\hat{y} = g(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}, \mathbf{x}_i) + b, \quad (3)$$

其中:  $k(\cdot, \cdot)$  为满足Mercer条件的核函数, 本文采用Gaussian核函数

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right), \quad (4)$$

其中  $\sigma$  为Gaussian核参数.

## 3 在线鲁棒LSSVR方法(Online robust LSSVR method)

为解决建模问题中离群点的影响,特别是非显著离群点对模型精度影响的问题,并且为了实时跟踪工业过程的变化,本文提出了在线鲁棒LSSVR建模方法.而为了能更好克服非显著离群点影响,对RLA方法进行了改进,提出了利用离群点(outliers)检测方法的鲁棒代价函数端点确定方法.在此基础上将改进的RLA方法引入到LSSVR模型,获得鲁棒估计.在线应用时,采用增量学习的递推策略更新鲁棒LSSVR模型,实现关键变量的在线预估.

### 3.1 改进鲁棒学习方法(Improved RLA method)

RLA的核心思想是应用鲁棒代价函数对模型权值进行更新.因此鲁棒代价函数的特征决定了RLA的有效性.鲁棒代价函数的端点确定以及权值更新策略也是RLA能较好克服离群点影响的关键因素,因此,下面从这3个方面来描述改进的RLA方法.

#### 3.1.1 鲁棒代价函数(Robust cost function)

本文RLA采用双曲正切函数为鲁棒代价函数<sup>[12]</sup>:

$$\sigma(e) = \begin{cases} \frac{1}{2} e^2, & 0 \leq |e| < a, \\ \frac{1}{2} a^2 + \frac{c_1}{c_2} \ln \left[ \frac{\cosh(c_2(b-a))}{\cosh(c_2(b-|e|))} \right], & a \leq |e| \leq b, \\ \frac{1}{2} a^2 + \frac{c_1}{c_2} \ln [\cosh(c_2(b-a))], & b < |e|, \end{cases} \quad (5)$$

其中:  $a$  和  $b$  是区间端点,  $c_1$  和  $c_2$  是常数, 本文选择  $c_1 = 1.73$ ,  $c_2 = 0.93$ .

为清晰描述RLA与最小二乘方法的区别,二者的代价函数曲线如图1所示.

从图1可以看出, 鲁棒代价函数是一个分段函数. 当  $0 \leq |e| < a$  时, 因LSSVR模型的权值已接近最优值, 所以在此范围内采用最小二乘代价函数; 当  $a \leq |e| \leq b$  时, 因误差较大, 需一定程度上限制较大误差的训练样本的影响; 当  $|e| < b$  时, 因样本误差太大而在训练时不考虑它对模型的作用. 因此, 通过上述过程把大误差点对模型的影响逐渐消除.

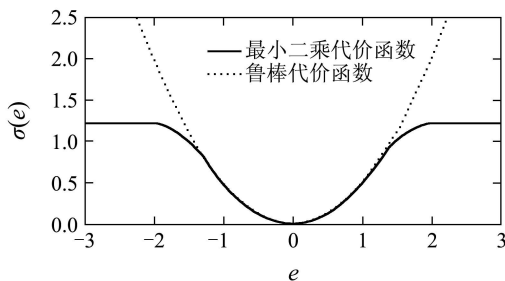


图 1 最小二乘代价函数与鲁棒代价函数

Fig. 1 Least square cost function and robust cost function

### 3.1.2 端点更新(Update of cutoff)

通过前面对鲁棒代价函数形式以及图像的分析可知, 鲁棒代价函数区间端点  $a, b$  是影响RLA效果的重要参数. 目前对于二者的确定常采用试凑法或者固定端点的方法, 然而, 这两种方法都存在一定的缺点. 为了更好更合理的确定区间端点, 本文提出了基于离群点检测的区间端点更新方法, 离群点检测的实质是鲁棒  $3\sigma$  原则<sup>[13]</sup>.

对于一组训练样本, 应用模型预测后可以得到一误差数据列为  $e_i (i = 1, \dots, n)$ , 为实现鲁棒估计, 将数据列的均值和标准差采用中位值和绝对偏差的中位值进行替换, 二者的计算公式分别为

$$\begin{cases} e_{\text{median}} = \text{median}_i \{e_i\}, \\ s_{\text{median}} = 1.4826 \text{median}_i \{|e_i - e_{\text{median}}|\}, \end{cases} \quad (6)$$

其中  $\text{median}(\cdot)$  是求中位值运算. 令第  $i$  个误差样本的偏差为  $d_i = |e_i - e_{\text{median}}|$ , 按照鲁棒  $3\sigma$  原则, 如果  $d_i \geq 3s_{\text{median}}$ , 则该样本是显著离群点. 既考虑到模型的预估精度, 又考虑到模型的在线应用, 如果  $1.5s_{\text{median}} \leq d_i < 3s_{\text{median}}$ , 则认为该样本是非显著离群点; 如果  $d_i < 1.5s_{\text{median}}$  则该样本是正常样本. 因此, 鲁棒代价函数区间端点  $a, b$  的计算公式为

$$\begin{cases} a = 1.5s_{\text{median}}, \\ b = 3s_{\text{median}}. \end{cases} \quad (7)$$

为了直观说明端点更新方法相对固定端点方法的优势, 图2绘制了区间端点由  $a, b$  变化为  $a^*, b^*$  的图像. 由图2可知, 当区间端点更新时, 鲁棒代价函数的形状发生变化. 由于本文采取了基于鲁棒  $3\sigma$  原理的端点更新方法, 所以区间端点随着模型训练误差  $e$  的分布而变化, 从而鲁棒代价函数也是随着  $e$  的分布

而变化, 因此本文方法能较好的跟踪模型误差  $e$  的变化, 减小离群点对建模过程的影响.

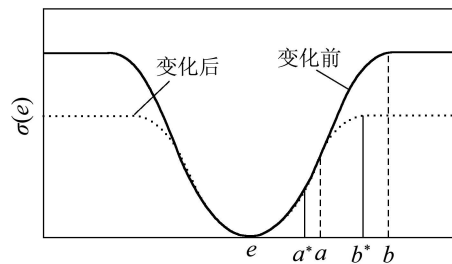


图 2 鲁棒代价函数变化

Fig. 2 Change of two robust cost function curves

本文方法相对于试凑方法的优势主要体现在: 计算简单, 避免了试凑方法重复多次才能得到较优区间端点的缺点.

### 3.1.3 模型权值更新(Update of model weight)

LSSVR模型权值  $\alpha_k$  采用梯度下降法更新, 即

$$\alpha_k(t+1) = \alpha_k(t) - \eta \nabla \alpha_k(t), \quad (8)$$

其中:  $t$  为迭代次数,  $\nabla \alpha_k(t)$  为梯度, 计算公式为

$$\nabla \alpha_k(t) = \frac{\partial R_e(t)}{\partial \alpha_k(t)} = -\frac{1}{n} \sum_{i=1}^n \psi(e_i(t)) k(\mathbf{x}_k, \mathbf{x}_i), \quad (9)$$

其中:  $\eta$  为学习速率,  $k(\mathbf{x}_k, \mathbf{x}_i)$  为高斯核函数,

$$R_e(t) = \frac{1}{n} \sum_{i=1}^n \sigma(e_i(t)),$$

$\psi(e_i(t))$  为鲁棒代价函数导数, 即

$$\psi(e_i(t)) = \frac{\partial \sigma(e_i(t))}{\partial e_i(t)} = \begin{cases} e_i(t), & 0 \leq |e_i(t)| < a(t), \\ c_1 \tanh[c_2(b(t) - |e_i(t)|)] \text{sgn}(e_i(t)), & a(t) \leq |e_i(t)| \leq b(t), \\ 0, & b(t) < |e_i(t)|. \end{cases} \quad (10)$$

其中:  $a(t), b(t)$  为第  $t$  次迭代时的区间端点,  $\text{sgn } e$  为符号函数, 当  $e < 0$  时,  $\text{sgn } e = -1$ , 当  $e > 0$  时,  $\text{sgn } e = 1$ .

可见, 权值更新与代价函数的导数  $\psi(e)$  相关. 传统最小二乘方法中, 最小二乘代价函数的导数为  $\psi(e) = e$ , 所以权值更新随着训练误差  $e$  的增大而正比例增大. 由于模型对离群点的训练误差  $e$  特别大, 从而造成了离群点所对应的权值较大, 加大了离群点对模型的影响, 导致模型精度下降. 由式(10)可以看出, 当训练误差  $e$  增大到一定程度,  $\psi(e)$  由增长较快的正比例变化变为增长相对缓慢的双曲正切曲线变化, 从而对离群点在建模过程中的影响起到限制作用, 提高模型的鲁棒性.

### 3.2 鲁棒LSSVR方法步骤(Steps of robust LSSVR method)

**Step 1** 确定LSSVR模型的核函数类型及参数, 给出鲁棒代价函数的区间端点初始值 $a(0)$ ,  $b(0)$ 及算法终止条件 $\epsilon_R$ .

**Step 2** 采集正常工况下的数据集作为建模数据, 建立LSSVR模型, 预测各输出变量 $f(\mathbf{x}_j, \boldsymbol{\alpha}(t))$ , 得到预测误差 $e_j(t) = y_j(t) - f(\mathbf{x}_j, \boldsymbol{\alpha}(t))$ , 其中:  $j = 1, \dots, N$ .

**Step 3** 由式(7)确定每次迭代的区间端点.

**Step 4** 由式(8)更新LSSVR模型的权值.

**Step 5** 如果 $R_e > \epsilon_R$ 成立, 则转Step 2进行循环, 否则终止循环学习, 得到最终的鲁棒LSSVR模型.

### 3.3 模型在线更新(Online update of model)

大多数生产过程是慢时变过程, 由于过程参数变化、建模样本不完备等原因导致鲁棒LSSVR模型老化. 虽然鲁棒LSSVR能克服训练样本中较大噪声的问题, 但它作为离线模型还是无法实时跟踪过程的变化, 准确预估过程关键变量. 因此, 本文提出了基于增量学习方法的在线鲁棒LSSVR建模方法. 其主要思想是增加过程中的实时数据对, 然后在线更新模型.

设初始建模数据集为 $Y^{(0)} = [y_1 \dots y_N]^T$ ,  $X^{(0)} = [\mathbf{x}_1 \dots \mathbf{x}_N]^T$ , 通过本文的鲁棒LSSVR模型可以得到模型参数为 $\boldsymbol{\alpha}^{(0)}$ ,  $b^{(0)}$ . 当第 $k+1$ 个新数据到来时, 模型在线训练的数据集更新为

$$Y^{(k+1)} = \begin{bmatrix} Y^{(k)} \\ y_{k+1} \end{bmatrix}, X^{(k+1)} = \begin{bmatrix} X^{(k)} \\ \mathbf{x}_{k+1}^T \end{bmatrix}, \quad (11)$$

其中:

$$Y^{(k)} = \begin{bmatrix} Y^{(0)} \\ y_1 \\ \vdots \\ y_k \end{bmatrix}, X^{(k)} = \begin{bmatrix} X^{(0)} \\ \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_k^T \end{bmatrix}.$$

由式(2)知鲁棒LSSVR实际是解线性方程组, 从而可以通过增量学习增加过程中的实时数据对, 达到跟踪过程时变特性的目的. 设第 $k$ 个采样时刻鲁棒LSSVR模型已经有 $M(M > N)$ 个样本, 当第 $k+1$ 个数据加入样本集后, 回归模型为

$$\begin{bmatrix} \Phi_M & \mathbf{p}_{M+1} \\ \mathbf{p}_{M+1}^T & c_{M+1} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha}_{M+1} \end{bmatrix} = \begin{bmatrix} 0 \\ Y^{(k+1)} \end{bmatrix}, \quad (12)$$

其中:

$$\Phi_M = \begin{bmatrix} 0 & \mathbf{1}_M^T \\ \mathbf{1}_M & K + \frac{1}{C} I_M \end{bmatrix},$$

$\boldsymbol{\alpha}_{M+1}$ 与 $b_{M+1}$ 为更新后的Lagrange乘子和模型偏置量,

$$\mathbf{p}_{M+1} = [1, k(\mathbf{x}_1 \mathbf{x}_{M+1}) \dots k(\mathbf{x}_M \mathbf{x}_{M+1})]^T, \\ c_{M+1} = \frac{1}{C} + k(\mathbf{x}_{M+1}, \mathbf{x}_{M+1}).$$

求解 $\Phi_{M+1}$ 的逆矩阵即 $H_{M+1} = \Phi_{M+1}^{-1}$ 是求解 $\boldsymbol{\alpha}_{M+1}$ 与 $b_{M+1}$ 的关键. 根据分块矩阵与矩阵和的求逆公式<sup>[14]</sup>得到 $H_{M+1}$ 的递推公式

$$H_{M+1} = \begin{bmatrix} H_M & \mathbf{0}_M \\ \mathbf{0}_M^T & 0 \end{bmatrix} + \frac{1}{\beta_{M+1}} \begin{bmatrix} \mathbf{h}_{M+1} \mathbf{h}_{M+1}^T & \mathbf{h}_{M+1} \\ \mathbf{h}_{M+1}^T & 1 \end{bmatrix}, \quad (13)$$

其中:  $H_M = \Phi_M^{-1}$ ,  $\mathbf{0}_M = [0 \dots 0]^T$ 为零向量,

$$\mathbf{h}_{M+1} = -H_M \mathbf{p}_{M+1}, \\ \beta_{M+1} = c_{M+1} - \mathbf{p}_{M+1}^T H_M \mathbf{p}_{M+1}.$$

随着样本数量的增多, 模型计算越来越复杂, 而为了简化计算, 提高模型的实时性, 应用减量学习方法剔除老样本. 为兼顾样本多样性和饱和性. 当新数据积累到一定数目时对模型参数进行减量学习.

设 $H_{M+1}$ 已知, 而 $H_M$ 未知, 当剔除第 $k+1$ 个样本时, 同理可以应用分块矩阵思想把 $H_{M+1}$ 划分为

$$H_{M+1} = \begin{bmatrix} H_{11} & \mathbf{h}_{12} \\ \mathbf{h}_{12}^T & h_{22} \end{bmatrix}, \quad (14)$$

从而得到

$$H_M = H_{11} - \frac{\mathbf{h}_{12} \mathbf{h}_{12}^T}{h_{22}}. \quad (15)$$

综上, 考虑到直接抛弃老样本会损失信息, 又要提高模型实时性, 本文在增量学习时, 选择新样本积累到一定数目(如 $N$ )时开始进行减量学习. 从而既可以提高建模速度又可以充分利用过程信息, 提高在线鲁棒LSSVR模型精度.

## 4 仿真分析(Simulation analysis)

考虑应用本文方法逼近正弦目标函数

$$y = \sin x + \varepsilon, \quad 0 \leq x \leq 2\pi. \quad (16)$$

这里 $\varepsilon$ 是服从 $N(0, 0.0016)$ 分布的噪声; 在 $[0, 2\pi]$ 之间等间隔产生31个数据用于建模, 从上述数据中随机选取3个样本, 分别为14, 20, 29(在图3中已标明), 对第14, 29个数据对的输出分别加入130%扰动, 对第20个数据对的输出加入4.5倍的扰动, 则第14, 29个数据样本可视为非显著离群点, 第20个数据样本可视为显著离群点; 另外相邻两个训练数据的中点作为测试样本.

在图3中, 将LSSVR方法、WLSSVR方法与本文方法进行比较, 可以看出, LSSVR方法所建立的模型因受到离群点的影响而偏向于离群点; 而WLSSVR

方法对于离群点则具有较好的鲁棒性, 而本文方法较WLSSVR 方法对于离群点具有更好的鲁棒性, 预测精度更高。

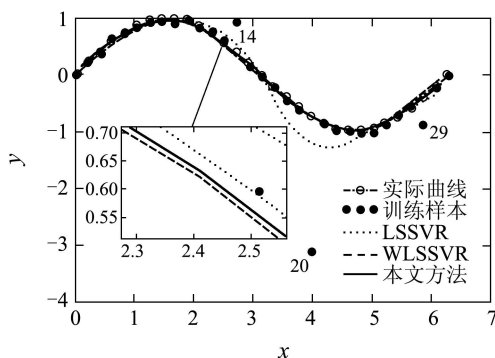


图 3 仿真结果对比

Fig. 3 Comparison of simulation results

为了更好地评价预测模型的性能, 本文分别采用最大绝对误差(MAE)和均方根误差(RMSE)对几种模型的预测准确性进行了分析, 比较结果列于表1. 所用公式分别为

$$MAE = \max |y_i - \hat{y}_i|, \quad (17)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (18)$$

其中:  $\hat{y}_i$ ,  $y_i$  分别为第  $i$  个测试样本的预测值与测量值,  $N$  为测试样本数目。

表 1 预测误差比较

Table 1 Comparison of prediction error

方法	MAE	RMSE
LSSVR方法	0.5158	0.2020
WLSSVR方法	0.0785	0.0461
本文方法	0.0633	0.0313

可以看出, WLSSVR 方法其预测精度虽然好于 LSSVR方法, 但却与本文方法存在一定差距。

为进一步说明本文方法在限制离群点对模型精度方面的影响, 图4绘制了3种方法在建模过程中, 每个训练样本所对应的权值。从图4可以看出, LSSVR方法对应每个离群点(样本14, 20, 29)的权值都比较大, 尤其是显著离群点(样本20)。从而加大了离群点在LSSVR建模过程中的影响, 导致模型过拟合, 模型精度下降。WLSSVR方法对应的显著离群点(样本20)权值虽然明显降低, 但是对非显著离群点(样本14, 29)却无能为力。而本文方法对应的两种离群点的权值都下降到了正常样本权值的范围内, 从而不仅消除了显著离群点影响, 而且对非显著离群点也起到了明显的抑制作用, 所以本文方法预测精度较高。这也证明了双曲正切估计函数能克服非显著离群点影响。

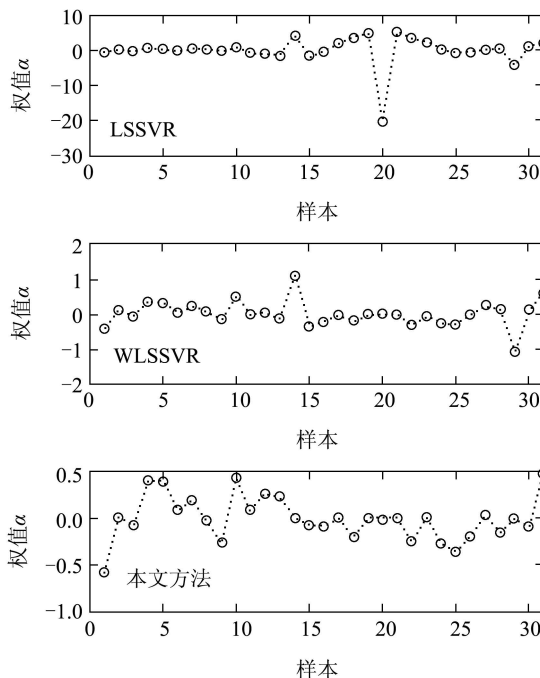


图 4 训练样本权值

Fig. 4 Weights of training samples

## 5 结论(Conclusion)

为了解决实际工业过程的在线鲁棒建模问题, 本文在最小二乘支持向量机回归建模方法的基础上, 引进鲁棒学习和增量学习方法, 提出了一种在线鲁棒LSSVR建模方法。该方法首先对RLA进行了改进, 提出了基于离群点检测的鲁棒代价函数区间端点确定方法; 接下来在LSSVR模型基础上通过改进的RLA逐步减小非显著离群点影响, 建立系统鲁棒LSSVR模型; 最后应用增量学习方法对鲁棒LSSVR模型进行更新, 建立在线鲁棒LSSVR模型, 实现对工业过程关键变量的实时鲁棒估计。本文还通过仿真实验, 验证了方法有效性。本文方法也同样适用于其他鲁棒非线性系统软测量建模问题。

## 参考文献(References):

- [1] VAPNIK V N. *The Nature of Statistical Learning Theory*[M]. New York: Springer-Verlag, 1995.
- [2] VAPNIK V N. An overview of statistical learning theory[J]. *IEEE Transactions on Neural Networks*, 1999, 10(5): 988 – 999.
- [3] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifiers[J]. *Neural Processing Letters*, 1999, 9(3): 293 – 300.
- [4] PELL R J. Multiple outlier detection for multivariate calibration using robust statistical techniques[J]. *Chemometrics and Intelligent Laboratory Systems*, 2000, 52(1): 87 – 104.
- [5] DASZYKOWSKI M, KACZMAREK K, HEYDEN Y V, et al. Robust statistics in data analysis-a review basic concepts[J]. *Chemometrics and Intelligent Laboratory Systems*, 2007, 85(1): 203 – 219.
- [6] MØLLE S F, FRESE J V, BRO R. Robust methods for multivariate data analysis[J]. *Journal of Chemometrics*, 2005, 19(10): 549 – 563.

- [7] KRUGER U, ZHOU Y, WANG X, et al. Robust partial least squares regression: part I algorithmic developments[J]. *Journal of Chemometrics*, 2008, 22(1): 1 – 13.
- [8] SUYKENS J A K, BRABANTER J D, LUKAS L, et al. Weighted least squares support vector machines: robustness and sparse approximation[J]. *Neurocomputing*, 2002, 48(1): 85 – 105.
- [9] SHIM J, HWANG C, NAU S. Robust LSSVM regression using fuzzy c-means clustering[C] // *Proceedings of the 2nd International Conference on Natural Computation*. Xi'an: Springer, 2006, 9: 157 – 166.
- [10] SÁNCHEZ A V D. Robustization of a learning method for RBF networks[J]. *Neurocomputing*, 1995, 9(1): 85 – 94.
- [11] 唐贤伦, 庄陵, 胡向东. 铁水硅含量的混沌粒子群支持向量机预报方法[J]. *控制理论与应用*, 2009, 26(8): 838 – 842.  
(TANG Xianlun, ZHUANG Ling, HU Xiangdong. The support vector regression based on the chaos particle swarm optimization algorithm for the prediction of silicon content in hot metal[J]. *Control Theory & Applications*, 2009, 26(8): 838 – 842.)
- [12] CHEN D S, JAIN R C. A robust back propagation learning algorithm for function approximation[J]. *IEEE Transactions on Neural Networks*, 1994, 5(3): 467 – 479.
- [13] HUBER P. *Robust Statistics*[M]. New York: Wiley, 1989.
- [14] TANG H S, XUE S T, CHEN R, et al. Online weighted LS-SVM for hysteretic structural system identification[J]. *Engineering Structures*, 2006, 28(12): 1728 – 1735.

#### 作者简介:

**张淑宁** (1983—), 男, 博士研究生, 主要研究方向为复杂工业过程建模与优化, E-mail: zhangshn0221@163.com;

**王福利** (1957—), 男, 教授, 博士生导师, 主要研究方向为复杂工业过程建模与优化、故障诊断等, E-mail: flwang@mail.neu.edu.cn;

**何大阔** (1975—), 男, 副教授, 主要研究方向为复杂工业过程建模、控制与优化等, E-mail: hedakuo@mail.neu.edu.cn;

**贾润达** (1981—), 男, 博士研究生, 主要研究方向为复杂工业过程建模与优化, E-mail: jiarunda@yahoo.com.cn.

#### (上接第1600页)

- [11] DIUK C, STREHL A L, LITTMAN M L. A hierarchical approach to efficient reinforcement learning in deterministic domains[C] // *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*. New York: ACM, 2006: 313 – 319.
- [12] SERI S, TADEPALLI P. Model-based hierarchical average-reward reinforcement learning[C] // *Proceedings of the 19th International Conference on Machine Learning*. San Francisco, CA: Morgan Kaufmann Publishers Inc, 2002: 562 – 569.
- [13] DAI Z H, CHEN X, Cao W H, et al. Model-based learning with Bayesian and MAXQ value function decomposition for hierarchical task[C] // *Proceedings of the 8th World Congress on Intelligent Control and Automation*. New York: IEEE, 2010: 676 – 681.
- [14] WILSON A, FERN A, RAY S, et al. Multi-task reinforcement learning: a hierarchical bayesian approach[C] // *Proceedings of the 24th International Conference on Machine Learning*. New York: ACM, 2007: 1015 – 1022.
- [15] PUTERMAN M L. Markov decision processes[M] // *Handbooks in Operations Research and Management Science*. North-Holland: Elsevier Science Publishers, 1990.
- [16] BELLMAN R E, DREYFUS S E. *Applied Dynamic Programming*[M]. Princeton, USA: Princeton University Press, 1962.

#### 作者简介:

**戴朝晖** (1969—), 男, 工程师, 在职博士研究生, 目前研究方向为多机器人系统和智能控制, E-mail: infob@mail.csu.edu.cn;

**袁姣红** (1985—), 女, 硕士研究生, 目前研究方向为多智能体强化学习, E-mail: yuanjiahong@wuhua.csu.edu.cn;

**吴敏** (1963—), 男, 博士, 博士生导师, 通讯作者, 长江学者特聘教授, 目前研究方向为智能系统、鲁棒控制和过程控制, E-mail: min@csu.edu.cn;

**陈鑫** (1977—), 男, 博士, 副教授, 目前研究方向为多机器人协作和群智能, E-mail: chenxin@mail.csu.edu.cn.