

文章编号: 1001-0920(2005)11-1307-04

## 基于支持向量机的软测量方法研究

常玉清<sup>1</sup>, 邹伟<sup>2</sup>, 王福利<sup>1</sup>, 毛志忠<sup>1</sup>

(1. 东北大学 a 教育部暨辽宁省流程工业综合自动化重点实验室, b 信息科学与工程学院, 沈阳 110004; 2 陕西宝鸡钛业股份有限公司, 陕西 宝鸡 721014)

**摘要:** 针对所有样本点均出现在最小二乘支持向量机模型中的缺陷, 提出一种改进的最小二乘支持向量机回归方法。根据最小二乘支持向量机模型学习误差的大小, 去除原变量空间中大部分误差较小的样本点, 从而获得回归模型的“稀疏”特性, 大大简化了模型复杂程度。同时, 将此方法应用于生物发酵过程, 建立青霉素发酵过程中产物浓度的软测量模型, 实现青霉素浓度的在线预估。实验结果表明, 该方法为生物发酵过程中难于在线测量质量参数的实时监测提供了一个有效的手段。

**关键词:** 软测量; 最小二乘支持向量机; 生物发酵; 青霉素浓度

**中图分类号:** TP274 **文献标识码:** A

## Research on Soft Sensing Method Based on Support Vector Machines

CHANG Yu-qing<sup>1</sup>, ZOU Wei<sup>2</sup>, WANG Fu-li<sup>1</sup>, MAO Zhi-zhong<sup>1</sup>

(1a Key Laboratory of Process Industry Automation of Ministry of Education, 1b School of Information Science and Engineering, Northeastern University, Shenyang 110004, China; 2 Shanxi Baoji Titanium Stock Limited Corporation, Baoji 721014, China Correspondent: CHANG Yu-qing, E-mail: cyq730522@sina.com)

**Abstract:** An improved least square-support vector machine (LS-SVM) regression method is proposed to overcome the drawback that all the original learning samples are in the LS-SVM model. By the learning errors of the LS-SVM model, most sample points of small errors are deleted from the original sample space, and thus the sparseness of the LS-SVM is obtained. Based on the proposed LS-SVM method, a soft sensor model is built to estimate the product concentration of the penicillin fermentation process. The experiment shows that the proposed LS-SVM is of the sparseness characteristic, and a novel procedure is provided for the realtime monitoring of quality variables, which are hard to measure on-line in fermentation processes.

**Key words:** Soft sensing; Least square-support vector machine; Fermentation process; Penicillin concentration

### 1 引言

支持向量机方法是近几年发展起来的机器学习新方法, 它较好地解决了小样本、非线性、高维数、局部极小等实际问题, 具有很强的泛化能力, 成为继神经网络研究之后新的研究热点<sup>[1-3]</sup>。最小二乘支持向量机 (LS-SVM) 是标准支持向量机的一种扩展, 其求解方法简单, 在软测量技术领域中得到了一定的应用<sup>[4-6]</sup>。但是, LS-SVM 和标准支持向量机方法

相比, 失去了“稀疏”特性。在这种情况下, 每一个样本数据都出现在回归模型中, 使模型结构非常庞大, 且在线估计时计算量大大增加。

综上, 本文提出一种改进的 LS-SVM 方法。根据 LS-SVM 模型误差的大小, 去除原变量空间中大部分学习误差较小的样本点, 使得 LS-SVM 方法也体现出一定的“稀疏”性。同时, 提出一种基于该改进 LS-SVM 模型的软测量方法, 并将该方法应用于生

收稿日期: 2004-12-14; 修回日期: 2005-01-26

基金项目: 国家自然科学基金项目 (60374003); 国家 973 子课题项目 (2002CB 312200)。

作者简介: 常玉清 (1973—), 女, 沈阳人, 副教授, 博士, 从事软测量技术的研究; 王福利 (1957—), 男, 辽宁辽阳人, 教授, 博士生导师, 从事控制系统故障诊断与容错、系统优化等研究。

物发酵过程, 实现青霉素发酵过程产物浓度的在线估计.

### 2 改进的LS-SVM 回归方法

#### 2.1 LS-SVM 回归方法

给定观测数据  $(x_i, y_i) \in R^m \times R, i = 1, 2, \dots, N$ . 首先通过非线性变换  $z = \mathcal{Q}(x_i) \in R^l$ , 将原  $m$  维输入空间映射到  $l(l \gg m)$  维特征空间. 在这个高维特征空间中采用线性函数  $f(x) = w\mathcal{Q}(x) + b$  来对其进行拟和, 并允许出现拟和误差

LS-SVM 回归算法的优化目标为

$$\begin{aligned} \min R(w) &= \frac{1}{2} \|w\|^2 + \frac{1}{2} c \sum_{i=1}^N \xi_i^2, \\ \text{s.t. } y_i &= w\mathcal{Q}(x_i) + b + \xi_i, \end{aligned} \quad (1)$$

式中:  $R(w)$  为结构风险,  $c$  为经验风险调整因子,  $\xi_i$  为容许误差. 统计学习理论指出,  $w^2$  控制着模型的推广能力<sup>[3]</sup>.

利用 Lagrange 法求解式(1)的优化问题, 定义 Lagrange 函数

$$\begin{aligned} L(w, \xi, b, \alpha) &= \frac{1}{2} \|w\|^2 + \frac{1}{2} c \sum_{i=1}^N (\xi_i)^2 - \sum_{i=1}^N \alpha_i (w\mathcal{Q}(x_i) + b + \xi_i - y_i). \end{aligned} \quad (2)$$

式中  $\alpha$  为 Lagrange 乘子. 根据优化条件

$$\frac{\partial L}{\partial w} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \xi_i} = 0, \frac{\partial L}{\partial \alpha_i} = 0,$$

有

$$w = \sum_{i=1}^N \alpha_i \mathcal{Q}(x_i), \quad \alpha_i = 0, \alpha_i = c \xi_i \quad (3)$$

$$w\mathcal{Q}(x_i) + b + \xi_i - y_i = 0 \quad (4)$$

由式(3)和(4)有

$$\begin{cases} \sum_{j=1}^N \alpha_j \mathcal{Q}(x_j) \mathcal{Q}(x_i) + b + \frac{\alpha_i}{c} - y_i = 0, \\ \alpha_i = 0 \end{cases} \quad (5)$$

定义核函数  $K(x_i, x_j) = \mathcal{Q}(x_i) \mathcal{Q}(x_j)$  为高维特征空间的内积运算,  $K(x_i, x_j)$  是满足 Mercer 条件的对称函数. 典型的核函数有多项式核、径向基核及网络核.

综上, 式(5)所示的优化问题转化为求解线性方程

$$\begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & K(x_1, x_1) + 1/c & \dots & K(x_1, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & K(x_N, x_1) & \dots & K(x_N, x_N) + 1/c \end{bmatrix} \times$$

$$\begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix}.$$

最后得到最小二乘支持向量机回归模型

$$f(x) = b + \sum_{i=1}^N \alpha_i K(x, x_i), \quad x \in R^m. \quad (7)$$

式中:  $N$  为输入样本的个数,  $x_i$  为第  $i$  个输入样本,  $x$  为某一输入变量,  $\alpha_i = 0$  所对应的输入样本  $x_i$  称为支持向量,  $\alpha_i$  称为支持向量系数.

本文采用径向基形式的核函数, 即

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right), \quad \sigma > 0 \quad (8)$$

式中  $\sigma$  为径向基宽度.

#### 2.2 改进的LS-SVM 回归方法

在标准支持向量机回归方法中, 只有小部分系数  $\alpha_i = 0$ , 也就是说只有一小部分输入样本出现在支持向量机中<sup>[1-3]</sup>, 这就是标准支持向量机的“稀疏”特性. 然而在LS-SVM中, 由于  $\alpha_i = c \xi_i$  的限制, 通常使得极少的系数  $\alpha_i$  为零, 这意味着几乎每一个输入样本都出现在回归模型中, 使得LS-SVM方法失去了“稀疏”特性, 模型结构变得非常庞大, 在线估计时计算量也大大增加. 综上, 本文提出一种改进的LS-SVM方法, 使其具有“稀疏”性, 减少回归模型中样本点的数量, 从而减少在线估计时预测模型的计算量.

标准支持向量机回归方法中, 支持向量是在函数关系变化比较剧烈的样本点位置上, 且这样的点在所有学习样本中只占很小一部分. 由于最小二乘支持向量机回归模型包含了所有的样本点, 函数关系变化剧烈样本点的学习误差通常会较大. 本文利用这个特性, 将学习误差较小的样本点从原样本空间中去除. 为了保证模型的泛化性能, 再利用剩余的样本重新建立LS-SVM模型. 重复上述过程直至LS-SVM模型在保证一定泛化性能的同时获得比较理想的“稀疏”特性为止.

改进的LS-SVM方法描述如下:

Step 1: 设  $N$  为原样本点数量.

Step 2: 根据给定的  $N$  个样本数据, 按 2.1 节中所述算法, 整定出合适的参数  $\sigma$  和  $c$ , 获得如式(7)所示的LS-SVM模型. 第  $i$  个样本点的学习误差表示为  $e_i$ , 有

$$e_i = |y_i - f(x_i)|, \quad i = 1, 2, \dots, N. \quad (9)$$

式中:  $y_i$  为第  $i$  个样本点输出变量的实测值,  $f(x_i)$  为对应的LS-SVM模型预测值.

Step 3: 去除一小部分误差较小的样本点(假设

为  $M$  个).

Step 4: 利用剩余的  $N - M$  个样本, 重新建立 LS-SVM 回归模型, 置  $N = N - M$ .

Step 5: 返回 Step 2, 直至在保证性能指标的同时获得较好的“稀疏”特性为止(本文采用的性能指标为泛化误差).

上述算法是一个迭代过程, 在每一步迭代过程中都需要重新求解一个如式(6)所示的线性方程组. 然而, 每经过一步迭代, 样本总数会减少, 所以求解的线性方程组维数也会逐渐降低.

### 3 生物发酵软测量建模的提出

生物发酵过程中, 一些重要质量参数如菌体浓度、产物浓度及基质浓度等很难在线测量, 软测量技术是解决此问题的有效途径<sup>[7-9]</sup>. 本文将利用改进 LS-SVM 方法建立青霉素发酵过程产物浓度在线预估的软测量模型.

通过对实验室青霉素发酵过程的分析, 选取容氧值 DO、发酵液 pH 值、发酵罐温度  $T$ 、发酵罐压力  $P$ 、尾气中  $\text{CO}_2$  含量  $\eta_{\text{CO}_2}$ 、尾气中  $\text{O}_2$  含量  $\eta_{\text{O}_2}$  的当前值及历史时刻值作为辅助变量, 建立它们与产物(青霉素)浓度  $c_p$  之间的软测量模型. 软测量模型可描述为

$$\begin{aligned} \hat{c}_p(k) = & f(\text{DO}(k), \text{pH}(k), T(k), P(k), \eta_{\text{CO}_2}(k), \\ & \eta_{\text{O}_2}(k), \text{DO}(k-1), \text{pH}(k-1), T(k-1), \dots \\ & P(k-1), \eta_{\text{CO}_2}(k-1), \eta_{\text{O}_2}(k-1), \dots, \\ & \text{DO}(k-d), \text{pH}(k-d), T(k-d), \\ & P(k-d), \eta_{\text{CO}_2}(k-d), \eta_{\text{O}_2}(k-d)). \end{aligned} \quad (10)$$

其中:  $\hat{c}_p(k)$  为当前时刻青霉素浓度的估计值, 即软测量模型的输出;  $\text{DO}(k-d)$  为当前时刻前  $d$  个时刻的容氧值, 其他辅助变量相同.

本文利用改进的 LS-SVM 方法来拟和式(10)所示的非线性关系, 并对其进行仿真验证.

### 4 仿真结果

利用生物发酵实验室青霉素发酵过程, 先以 1 h 为采样周期, 采集的 450 个样本数据作为学习样本, 分别利用 LS-SVM 和改进的 LS-SVM 建立青霉素浓度软测量模型 ( $d = 3$ ); 再以 6.3 h 为采样周期, 采集的 70 个样本数据作为测试样本, 对模型的“稀疏”特性及推广性能进行测试.

首先, 利用所有样本点建立 LS-SVM 软测量模型  $M_1$ , 图 1 为 450 个样本点所对应的 LS-SVM 模型的学习误差. 然后, 将学习误差  $e < 0.001$  的 158 个样本点从原样本空间中去除, 再利用剩余的 292 个样本点重新建立 LS-SVM 软测量模型  $M_2$ . 以此类

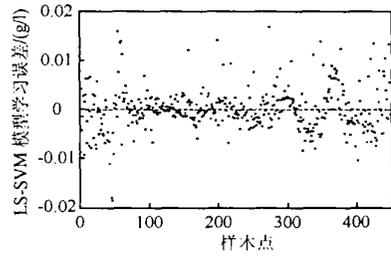


图 1 样本点所对应的 LS-SVM 模型学习误差

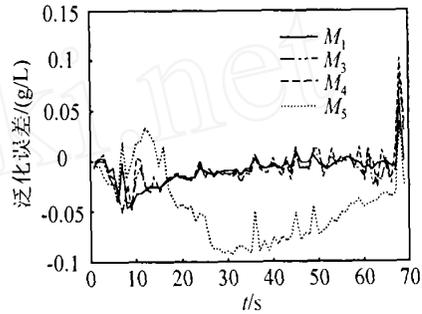


图 2 去除不同样本点后模型泛化误差比较

表 1 改进 LS-SVM 模型中样本点数及泛化误差

模型	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$
模型中样本数	450	292	159	101	72
泛化误差 $\bar{e}_g$	0.0113	0.0119	0.0135	0.0149	0.0203

推, 经 4 步迭代后 LS-SVM 软测量模型分别为  $M_2$ ,  $M_3$ ,  $M_4$  和  $M_5$ , 模型中的样本点数及平均泛化误差如表 1 所示, 模型  $M_1$ ,  $M_3$ ,  $M_4$  和  $M_5$  的泛化误差对比见图 2. 平均泛化误差为

$$\bar{e}_g = \frac{1}{N_l} \sum_{i=1}^{N_l} |e_i| \quad (11)$$

式中:  $N_l$  为测试样本点总数,  $e_i$  为测试样本点所对应的泛化误差, 可按式(9)计算.

从表 1 和图 2 可以看出, 第 5 步迭代后, 模型泛化误差明显增大. 因此, 利用本文提出的改进 LS-SVM 算法得到的软测量模型为第 4 步迭代后的 LS-SVM 模型.

从仿真结果可以看出, 本文提出的方法在保证模型泛化误差的同时, 去除了原样本空间中的大部分样本点, 这意味着改进后的 LS-SVM 模型中支持向量个数大大减少, 从而使 LS-SVM 模型具有了“稀疏”特性, 进而简化了 LS-SVM 回归模型的结构, 提高了模型在线预估时的计算速度.

图 3 为采用本文提出的改进 LS-SVM 方法, 经 4 步迭代后, 去除了原样本空间中 349 个样本点, 所得的青霉素发酵过程产物浓度软测量模型的估计结果. 可见, 本文方法在简化模型结构、提高模型在线计算速度的同时, 可较好地实现青霉素发酵过程产

物浓度的在线估计. 为生物发酵过程中难于在线测量质量参数的实时监测提供了有效的手段

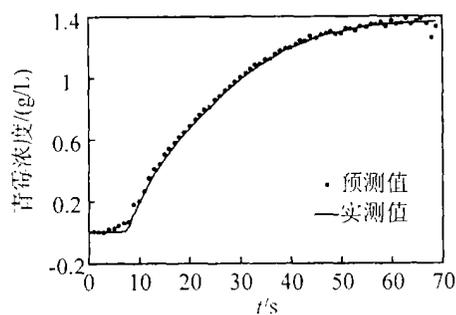


图3 青霉素浓度软测量模型泛化结果

## 5 结论

1) 针对LS-SVM 方法不具有“稀疏”特性的缺陷, 提出了改进的LS-SVM 回归方法, 使得回归模型结构得到简化; 在线使用时, 计算量大大减少, 估计速度得到了提高

2) 应用改进的LS-SVM 方法, 建立了青霉素发酵过程产物(青霉素)浓度的软测量模型. 仿真结果表明, 该软测量方法可以满足青霉素发酵过程产物浓度的辨识要求. 虽然该方法只在该过程中进行了实验, 但也同样适用于其他的生化过程

## 参考文献(References)

- [1] Cortes C, Vapnik V. Support-vector Networks [J]. *Machine Learning*, 1995, 20(1): 273-297.
- [2] Vapnik V N. *The Nature of Statistical Learning Theory* [M]. 1st ed. New York: Springer-Verlag, 1995.
- [3] 张学工. 关于统计学习理论与支持向量机[J]. *自动化学报*, 2000, 26(1): 32-42.

(Zhang X G. Statistical Learning Theory and Support Vector Machine [J]. *J of Automation*, 2000, 26(1): 32-42.)

- [4] Yan W W, Shao H H, Wang X F. Soft Sensing Modeling Based on Support Vector Machine and Bayesian Model Selection [J]. *Computers and Chemical Engineering*, 2004, 28(10): 1489-1498.
- [5] 阎威武, 朱宏栋, 邵惠鹤. 基于最小二乘支持向量机的软测量建模[J]. *系统仿真学报*, 2003, 15(10): 1494-1496.
- (Yan W W, Zhu H D, Shao H H. Soft Sensor Modeling Based on Support Vector Machines [J]. *J of System Simulation*, 2003, 15(10): 1494-1496.)
- [6] Suykens J A K. Nonlinear Modeling and Support Vector Machine [A]. *In Proc of the IEEE Instrumentation and Measurement Technology Conf* [C]. Budapest: Hungary, 2001: 287-294.
- [7] Mejdell T, Skogestad S. Output Estimation Using Multiple Secondary Measurements: High-purity Distillation [J]. *Process Systems Engineering*, 1993, 9(10): 1641-1653.
- [8] Yang S H, Wang X Z, Mcgreavy C, et al. Soft Sensor Based Predictive Control of Industrial Fluid Catalytic Cracking Processes [J]. *Institution of Chemical Engineers Trans IChemE*, 1998, 76(5): 499-508.
- [9] 常玉清, 王小刚, 王福利. PCA-RBFN 方法在精馏塔精苯干点估计中的应用[J]. *东北大学学报*, 2004, 25(2): 103-105.
- (Chang Y Q, Wang X G, Wang F L. PCA-DRBFN Model in Application to Estimating Dry Point of Pure Benzene in Rectifying Tower [J]. *J of Northeastern University*, 2004, 25(2): 103-105.)

(上接第1306页)

- [3] 聂宏, 赵军. 一类不确定切换组合系统的分散H<sub>∞</sub>鲁棒镇定[J]. *自动化学报*, 2004, 30(4): 635-640.
- (Nie H, Zhao J. Decentralized H<sub>∞</sub> Robust Stabilization for a Class of Uncertain Switched Composite Systems [J]. *Acta Automatica Sinica*, 2004, 30(4): 635-640.)
- [4] 聂宏, 王明顺, 赵军. 线性不确定时滞系统混杂反馈H<sub>∞</sub>鲁棒镇定[J]. *控制与决策*, 2004, 19(6): 642-646.

- (Nie H, Wang M S, Zhao J. Hybrid Feedback H<sub>∞</sub> Robust Stabilization for a Class of Uncertain Linear Time-delay Systems [J]. *Control and Decision*, 2004, 19(6): 642-646.)
- [5] Johansson M, Rantzer A. Computation of Piecewise Quadratic Lyapunov Functions [J]. *IEEE Trans on Automatic Control*, 1998, 43(4): 555-559.