

文章编号: 1002-0411(2002)06-567-05

## 基于 SVM 的软测量建模

冯 瑞 张浩然 邵惠鹤

(上海交通大学自动化系 上海 200030)

**摘 要:** 支持向量机(Support Vector Machines)是一种基于统计学习理论的新型学习机,本文提出用支持向量机建立软测量模型.理论分析和仿真研究表明,该方法学习速度快、跟踪性能好、泛化能力强、对样本的依赖程度低,比基于 RBF 神经网络的软测量建模具有更好的推广能力。

**关键词:** 支持向量机(SVMs); 软测量; RBF 神经网络; 建模

中图分类号: TP274

文献标识码: B

### SOFT SENSOR MODELING BASED ON SUPPORT VECTOR MACHINE

FENG Rui ZHANG Hao-ran SHAO Hu-ke

(Institute of Automation, Shanghai Jiaotong University, Shanghai 200030)

**Abstract:** Support vector machine (SVM) is a new learning machine based on the statistical learning theory. This paper presents a soft sensor model based on the SVM. Theoretical and simulation analysis indicates that this method features high learning speed, good approximation, well generalization ability, and little dependence on the sample set. It has the better performance than the soft sensor modeling based on the RBF neural network.

**Keywords:** support vector machine (SVMs), soft sensor, RBF neural network, modeling

### 1 引言(Introduction)

软测量技术主要解决工业过程控制中过程变量的检测问题.在实际控制系统中,由于工艺和技术的限制,许多重要的过程变量难以实时在线检测.为解决由此问题带来的控制问题,以前主要存在两种方法<sup>[1]</sup>: 1) 采用间接的质量指标控制,但此方法难以保证最终质量指标的控制精度; 2) 采用在线分析仪表,并用硬件予以实现.这种方法开发费用高,硬件投资大,原理复杂,因此除了在少数特殊情况以外,并没有被工业控制界广泛接受和采纳.高级过程控制理论的出现和成熟,直接促成了软测量技术的形成和发展.

软测量技术就是建立待测变量与可测或易测的过程变量之间的非线性函数关系,通过对可测变量的变换计算,间接得到待测变量的估计值.软测量建模方法主要有机理建模方法和辨识建模方法; 辨识建模方法又分有动态模型间接建模、静态模型辨识和非线性建模.通过适当选择神经网络的隐单元数和网络层次,前馈网络能以任意精度逼近非线性函

数<sup>[3]</sup>,因此人工神经网络技术已经广泛应用到工业过程建模和控制中.目前有许多学者成功地应用神经网络进行软测量建模<sup>[4,5]</sup>,其中比较有代表性的是 RBF 神经网络<sup>[4,5,6]</sup>.

但是,神经网络依然存在一些缺陷.第一,网络结构需要事先指定或应用启发式算法在训练过程中修正,这些启发式算法难以保证网络结构的最优化,对于多层网络,这是一个很复杂的组合问题; 第二,网络权系数的调整方法存在局限性,表现在训练可能过早结束,权值衰退等等; 第三:神经网络容易陷入局部最小,有些训练算法甚至不能得到最小,这样,为了得到更好的泛化性能,只能牺牲网络的性能,而这种间接控制系统的能力,和最优相去甚远.第四:过分依赖学习数据的质量和数量,模型性能的好坏取决于模型训练过程中样本数据的数量和质量.然而,大多数情况下,数据有限,而且不一致.许多问题中的输入空间是高维的,得到的数据仅是输入空间的稀疏分布.即使有高质量的训练数据,大量数据必然会大大增加算法的训练时间.针对用于非

线性函数估计的 RBF 神经网络, 神经网络的缺陷具体表现为以下问题<sup>[6]</sup>: 1) 网络基函数中心点难以确定; 2) 输出权值计算过程中存在数值变态问题; 3) 网络的在线校正效果不明显; 4) 网络的泛化能力不强. 对于前三个问题, 文献[6]在一定程度上加以改进. 由于网络易导致过学习, 目前还没有有有效提高网络泛化能力的方法.

支持向量机最初是由 Vapnik<sup>[7]</sup>提出的一种新兴的基于统计学习理论的学习机. 相对于神经网络的启发式学习方式和实现中带有很大的经验成分相比, SVMs 具有更严格的理论和数学基础, 不存在局部最小问题, 小样本学习使它具有很强的泛化能力, 不过分依赖样本的数量和质量. 尽管它的研究和应用刚刚开始, 已经有大量的文献报道了它在各个领域的应用, 如模式识别、回归估计和时序预测等. 支持向量机基本思想是把训练数据集从输入空间非线性地映射到一个高维特征空间(Hilbert 空间), 然后在此空间中求解凸优化问题(典型二次规划问

题). SVMs 的最优求解基于结构风险最小化思想, 因此比其他非线性函数逼近方法具有更强的泛化能力.

本文应用支持向量机算法用于软测量建模, 全文组织如下: 第二部分介绍基于标准支持向量机软测量建模算法; 第三部分介绍基于改进 SVMs 算法的软测量建模; 第四部分通过理论和仿真研究分析了基于 SVMs 和 RBF 神经网络软测量模型的性能.

## 2 基于标准 SVMs 算法的软测量建模(Soft sensor modeling based on standard SVMs)

软测量技术就是先建立待测变量与其它一些可测或易测的过程变量之间的关系, 通过对可测变量的检测、变换和计算, 间接得到待测变量的估计值<sup>[2]</sup>. 软测量模型的输出可作为过程控制系统状态变量或输出变量的估计值, 送入控制装置, 参与反馈控制. 软测量系统实现的基本框架如图 1 所示.

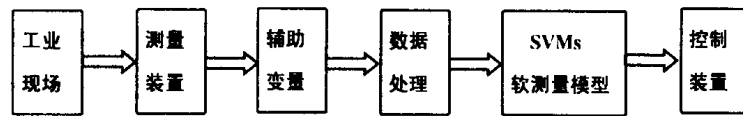


图 1 软测量系统框架

Fig. 1 Frame of the soft sensor system

测量装置从工业现场中获取用于估计主变量的辅助变量. 经过数据处理, 完成输入数据的二次变量的选择, 将存在显著误差的数据剔除, 进行随机误差的过滤, 消除冗余数据, 并实现数据的归一化处理. 通过 SVM 软测量模型对待测变量进行最优估计, 估计值作为控制系统的反馈变量, 参与对控制对象的控制.

这样, SVM 软测量模型可以看成是一个非线性函数估计器, 完成输入数据和输出数据的非线性映射, 而不必关心对象的具体结构及其机理.

给定  $l$  个样本数据  $(x_k, y_k)_{k=1}^l$ , 其中  $x_k \in R^n$  为  $n$  维软测量模型样本输入,  $y_k \in R$  为软测量模型样本输出, 则软测量建模问题就是要找出一个函数  $f$ , 使之通过样本训练后, 对于样本以外的  $x$ , 通过  $f$  找出对应的  $y$ . 利用非线性映射  $\mathcal{Q}(\cdot)$  将训练数据集非线性地映射到一个高维特征空间(Hilbert 空间), 将非线性函数估计问题转化为高维特征空间中的线性函数估计问题. 设函数具有下式的形式:

$$f(x) = \omega^T \mathcal{Q}(x_k) + b \quad \omega \in R^m, b \in R \quad (1)$$

其中非线性函数  $\mathcal{Q}(\cdot): R^n \rightarrow R^m$  将输入空间映

射到一个高维特征空间, 特征空间的维数不固定, 有可能是无穷维,  $b$  为偏置量. 求解目的是要寻找参数  $\omega^T$  和  $b$ , 使得对于样本以外的输入  $x$ , 有  $|f(x) - \omega^T \mathcal{Q}(x_k) - b| \leq \epsilon$ . 寻找式(1)的参数等价于求解下面的优化问题:

$$\text{目标函数:} \quad \min_{\omega, b} J = \frac{1}{2} \omega^T \omega \quad (2)$$

$$\text{约束条件:} \quad \begin{cases} y_i - \omega^T \mathcal{Q}(x_i) - b \leq \epsilon \\ \omega^T \mathcal{Q}(x_i) + b - y_i \leq \epsilon \end{cases} \quad i = 1, \dots, l$$

为了使上述优化问题确定有解, 引入松弛变量  $\xi, \xi^*$ . 则优化问题转化为下式问题:

目标函数:

$$\min_{\omega, b, \xi, \xi^*} J = \frac{1}{2} \omega^T \omega + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (3)$$

约束条件:

$$\begin{cases} y_i - \omega^T \mathcal{Q}(x_i) - b \leq \epsilon + \xi_i \\ \omega^T \mathcal{Q}(x_i) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad i = 1, \dots, n$$

其中常数  $C(C > 0)$  表示函数  $f$  的平滑度和允许误差大于  $\epsilon$  的数值之间的折衷.  $\epsilon$  不灵敏损失函数

的形式如下:

$$|y - f(x)|_{\epsilon} = \begin{cases} 0, & |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon, & \text{其他} \end{cases} \quad (4)$$

通过式(3)的对偶形式可以求它的最优解. 式(3)的对偶形式可以根据目标函数和约束条件建立 Lagrangian 函数:

$$L = \frac{1}{2} \omega^T \omega + C \sum_1^l (\xi_i + \xi_i^*) - \sum_1^l \alpha_i (\epsilon + \xi_i - y_i + \omega^T \mathcal{Q} x_i + b) - \sum_1^l \alpha_i^* (\epsilon + \xi_i^* + y_i - \omega^T \mathcal{Q} x_i - b) - \sum_1^l (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (5)$$

其中对偶变量  $\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0$ .

对式(5)中的原变量求偏导, 可得:

$$\partial_{\alpha_i} L = \sum_1^l (\alpha_i^* - \alpha_i) = 0 \quad (6)$$

$$\partial_{\omega} L = \omega - \sum_1^l (\alpha_i^* - \alpha_i) \mathcal{Q} x_i = 0 \quad (7)$$

$$\partial_{\xi_i^*} L = C - \alpha_i^* - \eta_i^* = 0 \quad (8)$$

将式(6)(7)(8)代入式(5)可得对偶优化问题:

目标函数:

$$\begin{cases} \max J = \max_{\alpha, \alpha^*} \min_{\omega, b, \xi, \xi^*} L = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) - \\ (\alpha_j - \alpha_j^*) \langle \mathcal{Q} x_i, \mathcal{Q} x_j \rangle - \epsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) \\ + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \end{cases} \quad (9)$$

约束条件:

$$\begin{cases} \sum_{i=1}^l (\alpha_i, \alpha_i^*) = 0 \\ \alpha_i - \alpha_i^* \in [0, C] \end{cases}$$

重写式(7), 可得  $\omega$  和待估计函数:

$$\begin{cases} \omega = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \mathcal{Q} x_i \\ f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle \mathcal{Q} x_i, \mathcal{Q} x \rangle + b \end{cases} \quad (10)$$

其中  $(\alpha_i - \alpha_i^*) \neq 0$  对应的  $x_i$  为支持向量. 变量  $\omega$  反映了函数的复杂度, 是非线性映射函数  $\mathcal{Q} \cdot$  的线性组合. 从这个意义上说, 利用支持向量进行函数估计的计算复杂度和输入空间的维数是不相关的, 依赖于支持向量的数目. 偏置量  $b$  可以通过 KTT (Karush-Kuhn-Tucker) 条件计算. 由于支持向量算

法仅依赖于输入数据的点积, 为了找到非线性映射函数  $\mathcal{Q} \cdot$ , 可以引进核函数来代替非线性映射, 即用  $k(x, x') = \langle \mathcal{Q} x, \mathcal{Q} x' \rangle$  来代替  $\mathcal{Q} \cdot$ , 则式(9)改写为:

目标函数:

$$\begin{aligned} \max_{\alpha, \alpha^*} J = & -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) - \\ & \epsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) \end{aligned} \quad (11)$$

约束条件:

$$\begin{cases} \sum_{i=1}^l (\alpha_i, \alpha_i^*) = 0 \\ \alpha_i - \alpha_i^* \in [0, C] \end{cases}$$

式(10)重写为:

$$\begin{cases} \omega = \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) \mathcal{Q} x_i \\ f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) k(x_i, x) + b \end{cases} \quad (12)$$

其中核函数  $k(x_i, x)$  是满足 Mercer 条件的任意对称函数. 常用的核函数有①线性核:  $k(x_i, x) = x_i \cdot x$ ; ②径向核:  $k(x_i, x) = \exp(-\|x - x_i\|^2 / 2\sigma^2)$ ; ③多项式核:  $k(x_i, x) = (x_i \cdot x + 1)^d$ ,  $d = 1, 2, \dots, N$ ; ④感知器核:  $k(x_i, x) = \tanh(\beta x_i \cdot x + b)$ .

最后得到的软测量模型为:

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) k(x_i, x) + b \quad (13)$$

### 3 基于最小二乘 SVMs 算法 (LS\_SVMs) 的软测量建模 (Soft sensor modeling based on the least squares SVMs)

标准 SVM 算法复杂度不依赖于输入空间的维数, 而依赖于样本数据的个数. 样本数据越大, 求解相应的二次规划问题越复杂, 计算速度越慢. 目前解决此问题的方法主要有以下几种: 1) 改进优化算法或采用其他算法得到支持向量取代二次规划求解优化问题, 如最小二乘 SVM 算法 (LS\_SVM); 2) 对待估计函数进行分段估计, 然后再组合到一起以减少算法复杂度; 3) 构建新的核函数, 使之能降低优化问题的维数. 其中比较成功的是最小二乘 SVM 算法. 它和标准 SVM 的主要区别在于损失函数项和等式约束.

最小二乘 SVM 算法的目标函数<sup>[8]</sup>:

$$\min_{\alpha, b, e} J = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^l e_i^2 \quad (14)$$

约束条件:

$$y_i = \omega^T \cdot \mathcal{Q}(x_i) + b + e_i, \quad i = 1, \dots, l$$

相应的 Lagrangian 函数为:

$$L = J - \sum_{i=1}^l \alpha_i [\omega^T \mathcal{Q}(x_i) + b + e_i - y_i] \quad i = 1, \dots, l \quad (15)$$

求解的优化问题转化为求解线性方程:

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (16)$$

其中  $y = [y_1, \dots, y_l]^T$ ,  $\vec{1} = [1, \dots, 1]^T$ ,  $\alpha = [\alpha_1, \dots, \alpha_l]^T$ ,  $\Omega_{kl} = \mathcal{Q}(x_k)^T \mathcal{Q}(x_l) = K(x_k, x_l)$

软测量模型为:

$$y(x) = \sum_{i=1}^l \alpha_i K(x, x_i) + b \quad (17)$$

从上面的推导可以看出, 等式约束在将求解优化问题转化成求解线性方程的过程中起到重要作用, 优化问题用线性方程求解显然可以大大减少算法的复杂度. 另外, 其需要预先确定的参数比 Vapnik SVMs 算法少, 如采用 RBF 核函数的 LS\_SVMs 算法的参数为  $\gamma, \delta$ , 而采用 RBF 核函数的 Vapnik SVMs 算法的参数为:  $\epsilon, C, \delta$ .

#### 4 仿真研究(Simulation and results)

本文研究的对象是某炼油厂的重油催化裂化装置中的吸收稳定系统. 重油催化裂化装置将常压渣油和掺炼榨油经催化裂化变成更多的轻质油品. 该过程生产成本低, 生产过程简单, 是当前原油深度加工经济有效的工艺方法.

吸收稳定系统是催化裂化装置中一个重要组成部分. 该系统的主要任务是将分馏系统来的低压富气(含轻汽油和液化石油气)和粗汽油(含干气和液化气)进一步分离成干气、液化气和稳定汽油三个部分. 该系统中最主要的指标是汽油收率. 本文将 LS\_SVM 的软测量建模用于汽油收率的估计. 首先

进行二次变量选取, 根据工艺分析, 对汽油收率的主要影响变量有原料油进料量, 回炼油进料量、催化反应温度、主分馏塔顶温、轻柴抽温、稳定塔底温. 确定这六个变量为软测量模型估计汽油收率的辅助变量, 作为 SVM 软测量模型的输入. 模型输出为汽油收率.

本文分别采用 LS\_SVMs 和 RBF 神经网络进行软测量建模. 其中 LS\_SVMs 采用 RBF 核函数, 核宽度  $\sigma$  取 0.60,  $\gamma$  取 10; RBF 神经网络参数  $e$  取 0.0001,  $\sigma$  取 0.60. 图 2 为支持向量机软测量模型的估计值与实际分析值的比较曲线; 表 1 为用支持向量机和 RBF 神经网络对同样的训练集和测试集的拟合结果.

从表 1 可以看出, 支持向量机的样本集均方差比 RBF 神经网络的样本均方差大, 而其测试集均方差误差要比 RBF 神经网络精度高. 当样本集数增加时, 两者的泛化误差都有所加大, 但是, RBF 神经网络的泛化误差的增加速度明显要高于支持向量机, 说明 SVMs 的泛化能力比 RBF 神经网络好. 在样本依赖程度上, 可以看出, 当样本集变化时, SVMs 泛化误差变化幅度小于 RBF 神经网络的泛化误差变化幅度, 说明 SVMs 对样本数据的依赖程度比 RBF 神经网络小.

神经网络等传统方法基于经验风险最小化原则, 应用误差反向传播等方法, 使其经验风险最小, 但是经验风险最小, 不等于实际风险最小, 这是神经网络出现过学习的原因, 即能以任意精度逼近样本集, 而泛化能力却不如人意. 从前面的推导, 不难看出, SVMs 的优化基于结构风险最小化原理, 其优化目标函数包含两项指标: 经验风险和置信空间, 两者共同决定了 SVMs 的实际风险. 因此, SVMs 的泛化能力要比神经网络等传统方法好. 另外, SVMs 小样本学习的特性也决定了它对于样本数据集依赖程度比传统方法小.

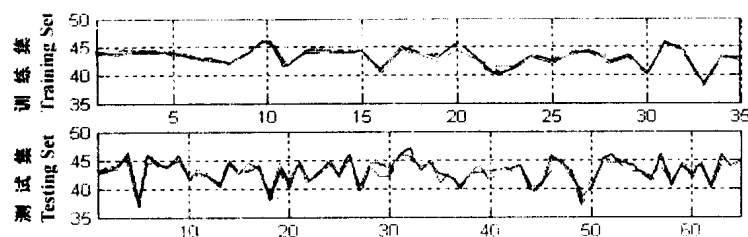


图 2 估计值与实际值分析曲线( $\sigma = 0.6, \gamma = 10$ )

Fig. 2 The analysis curves of estimation and analyzed value ( $\sigma = 0.6, \gamma = 10$ )

表 1 LS\_SVM 与 RBF 神经网络均方误差比较分析表

Tab. 1 Mean squared errors analysis of LS\_SVMs and RBFNNs

样本集数目 Training Set	测试集数目 Testing Set	LS_SVMs			RBF	
		样本集均方误差 LMSEs	测试集均方误差 GMSEs	支持向量数 SVs	样本集均方误差 LMSEs	测试集均方误差 GMSEs
20	80	0.0017	0.0281	12	0	0.0254
30	70	0.0025	0.0124	19	8.333e-004	0.0159
35	65	0.0032	0.0104	23	0.0013	0.0119
40	60	0.0031	0.0104	27	0.0014	0.0120
50	50	0.0020	0.0112	40	0.0014	0.0133
60	40	0.0020	0.0113	52	0.0016	0.0146
70	30	0.0028	0.0116	65	0.0022	0.0161

综合以上理论和仿真分析, 基于 LS\_SVMs 的软测量建模在精度、泛化能力和样本数据依赖程度上比基于 RBF 神经网络的软测量建模要好。

## 5 结论(Conclusion)

本文详细推导了用 SVMs 和改进的 Vapnik SVMs 算法(最小二乘 SVMs: LS\_SVMs) 算法进行软测量建模的实现过程, 并应用于催化裂化装置中的汽油收率的估计。理论分析和仿真结果表明, 与 RBF 神经网络建立的软测量模型相比, 改进的 Vapnik SVMs 算法(最小二乘 SVMs: LS\_SVMs) 算法具有学习能力强、泛化能力好、对样本的依赖程度低等优点, 较好地跟踪了汽油收率的变化趋势, 显示出它在工业领域巨大的应用潜力。

## 参 考 文 献 (References)

- 1 荣 冈, 金晓明, 王树青. 先进控制技术及应用. 化工自动化及仪表, 1999; 26(4): 70~ 72
- 2 徐 敏, 俞金寿. 软测量技术. 石油化工自动化, 1988, 2: 1

- 3 Funahashi K. J, On the Approximate Realization of Continuous Mapping by Neural Networks[M]. 1989, (2): 183~ 192
- 4 王旭东, 邵惠鹤. RBF 神经网络在非线性系统建模中的应用. 控制理论与应用. 1997, 14(1): 59~ 66
- 5 仲 蔚, 俞金寿. 基于 RBFN 的加氢裂化分馏塔产品质量估计 MIMO 软测量. 石油化工自动化, 1995, 5: 19
- 6 王旭东, 邵惠鹤. 神经网络建模与软测量技术. 化工自动化及仪表, 1996, 23(2): 28~ 31
- 7 V Vapnik, the Nature of Statistical Learning Theory. Springer Verlag, New York, 1995
- 8 J A K Suykens. Nonlinear Modelling and Support Vector Machine, IEEE Instrumentation and Measurement Technology Conference Budapest, Hungary, May 2001, 21~ 23

## 作者简介

冯 瑞(1971~), 男, 博士生. 研究领域为统计学习、数据挖掘、优化建模。

张浩然(1972~), 男, 博士生. 研究领域为统计学习、智能控制、非线性系统建模与控制。

邵惠鹤(1936~), 男, 教授, 博士生导师. 研究领域为工业过程控制、智能控制。