

基于改进 SMO 算法的热工参数灰色软测量建模

乔 弘, 张全壮, 吴 蓉

(广东省电力设计研究院, 广东 广州 510663)

摘要:介绍了适宜支持向量机处理大规模数据回归问题的序列最小优化 (SMO) 学习算法, 针对 SVR 进行二次规划处理大规模数据时计算复杂度高和学习机参数选择方法复杂的问题, 从算法结构和参数选择两个方面对 SMO 算法进行了改进, 使运算速度和建模效率得到了进一步提高。结合灰色理论进行辅助变量选取, 并应用改进的 SMO 算法建立了火电厂烟气含氧量软仪表, 通过电厂的实测历史数据仿真表明, 改进的算法较传统的 SMO 算法在计算速度和性能上有较大提高, 建立的软仪表模型具有更高的精度, 能满足应用要求。

关键词:序列最小优化(SMO); 灰色关联分析; 氧量; 软测量

中图分类号: TP13 文献标识码: A 文章编号: 1003-7241(2010)10-0004-04

Thermal Parameters Grey Soft-Sensing Modeling Based on the Improved Algorithm of SMO

QIAO Hong, ZHANG Quan-zhuang, WU Rong

(Guangdong Electric Power Design Institute, Guangzhou 510663 China)

Abstract: In order to solve the problem that normal SVM algorithm can not deal with large scale data, the improved SVR algorithm of sequential minimal optimization (SMO) is introduced in this paper. The improved algorithm makes improvements in two aspects of structure and parametric selection to increase operational speed. It uses grey theory to select the auxiliary variables and build a model of soft instrument for the flue gas oxygen content in power plant. The simulation with historical data measured by plant shows that compared with the normal SMO algorithm, the operating speed and the soft instrument precise of the improved SMO algorithm is better.

Key words: Sequential Minimal Optimization (SMO); grey relational analysis; uolume of oxygen; soft-sensing

1 引言

对烟气含氧量进行及时、准确测量是保证锅炉燃烧效率的重要前提^[1]。目前火电厂广泛采用的基于热磁式或氧化锆传感器的氧量分析仪精度不高、投资大、使用寿命短, 而且测量滞后较大, 不利于燃烧过程的在线监视和实时提供闭环控制所需的反馈信号, 从而直接影响燃烧效率。采用软测量技术是解决以上问题的有效途径之一。

近年来, 基于支持向量回归(SVR)的算法已经广泛应用于工业对象的软测量建模, 并取得了良好的试验和测量效果^[2]。Platt^[3]提出的序列最小优化(SMO)算法, 其

主要思想是把一个大的优化问题分解成一系列只含两个变量的优化问题。起初该算法主要用于分类问题, 后来 Smola 和 Sch ö lkopf^[3]提出了一种训练回归 SVM 的 SMO 算法, 是 Platt 算法的类比扩展, 为支持向量机处理大规模数据回归问题提供了可能。SMO 算法已成为目前较为有效的 SVR 训练算法之一, 不仅在整体速度上对比其它算法有明显优势, 而且它没有矩阵操作不会占用大量内存, 算法容易实现, 将其应用于软测量建模具有很大优势。

本文在 SMO 算法的基础上对其进行改进, 并结合灰色理论应用于火电厂锅炉尾部烟气氧量的软测量建模, 仿真结果表明, 改进的算法较传统算法具有

收稿日期: 2010-06-30

更快的收敛速度,所建立的软测量模型能够满足应用精度要求。

2 SVR 和参数自适应选取

针对非线性回归问题, SVM 通过定义适当的核函数进行非线性变换,将输入空间变换到一个高维空间,然后在这个新空间应用线性回归寻找支持向量,即距离最优分类面最近,并且根据平行于最优分类面的超平面上的训练样本构造最优分类面,增加了对测试样本的泛化能力,其泛化能力远好于神经网络模型和模糊模型。

支持向量回归(SVR)的基本思想是通过核函数(非线性映射 Φ)将输入向量 x 映射到高维特征 Hilbert 空间,并在这个空间进行线性回归^[4]。

应用具有固定核函数形式(如 RBF 核)的 ϵ -SVR 进行非线性系统建模时,需要确定学习参数 C , ϵ 以及核函数参数,它们直接影响 SVR 建模的精度和复杂程度,而传统基于经验的“试凑法”不仅对建模人员的经验要求较高,而且使参数选择成为一个黑箱过程。Cherkassky 和 Ma^[5]提出了一种基于训练数据的参数选择方法, C 和 ϵ 的值均基于训练数据得到,而不采用抽样选取的办法,如 1-3 式所示。

$$C = \max(|\bar{y} + 3\delta_y|, |\bar{y} - 3\delta_y|) \quad (1)$$

$$\epsilon = 3\delta \times [\ln n \times n^{-1}]^{-1/2} \quad (2)$$

$$\delta = \sqrt{\frac{n^{1/5}k}{n^{1/5}k-1} \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

Where $2 \leq k \leq 6$

式 1 中 \bar{y} 和 δ_y 分别是训练数据中 y 值得均值和标准差,式 2 采用中心极限定理求解 ϵ 。 δ 是输入噪声的标准差,而 n 是训练采样的个数。 δ 的值事先未知,可以运用如 3 式的基于 k -邻域思想方法进行估计,对训练数据进行线性回归来估算该噪声方差。式中 k 是表征偏差低/高的方差估计值, \hat{y}_i 是 y 的估计值。

3 SMO 算法与改进

SMO 算法能有效解决 SVM 的二次规划(QP)问题。其按照 Osuna 的理论,在保证收敛的情况下,将 SVR 的 QP 问题分解为一系列最小规模的 QP 子问题来解决。SMO 算法主要包括两个部分^[6],一是采用解析的方法对成对的拉格朗日乘子进行优化,二是启发式(Heuristics)

的选择需要优化的拉格朗日乘子。在对 SMO 算法的原理和实现步骤进行研究的基础上可以进行改进,通过优化步长和步长阈值可以使运算速度和性能得到进一步提高。具体改进措施如下:

1. 内层循环中最大优化步长改进为 $|(E_1 - E_2)/\eta|$, 避免了因使用 $|E_1 - E_2|$ 作为粗略步长估计带来的问题(例如 $\eta = 0$, 最小优化问题的二次型只是半正定的)。

2. 在内层循环中引入衡量调整步长的标准,即阈值。由于在内层循环中, Lagrange 乘子的调整小到一定程度时,就不必再调整,因此引入阈值 ξ 来限定调整范围,如果调整量小于该阈值就停止调整。这样既避免了因选阈值不当而引起的重要调整被忽略,又避免了一些微不足道的调整上浪费计算时间。

3. 由于支持向量回归大多用于非线性回归的情况,当处理大规模数据时,内层循环中不断进行核函数计算,极大的影响了计算速度。因此,在算法的计算机平台实现时,可以将核函数定义为全局变量,一次赋值避免反复计算,节约大量计算时间。但这只是实现手段上内存和运算速度两种性能的取舍。

4 烟气含氧量灰色软测量建模

4.1 基于灰色关联分析的辅助变量选取

对于某软测量对象,一般来说,根据初步理论分析其原始辅助变量数目和类型会很多,而且其间存在相互耦合的程度较大,为了提高模型的性能和精度,有必要对输入变量进行适当的降维处理。灰色关联分析是灰色理论的基本内容,其基本思想是根据曲线间相似程度来判断因素间的关联程度,它对样本量的大小没有过高的要求,不需要典型的分布规律,而且其定量分析的结果与定性分析的结果一般能够吻合^[7]。因此,可以将机理分析与灰色关联分析结合进行软测量模型辅助变量选取,从而提高模型精度和性能。

根据机理分析烟气含氧量主要受煤质变化、锅炉炉膛漏风、未完全燃烧等因素的影响,可以初步选择主蒸汽流量、主汽压力、主汽温度、给水流量、燃料量、燃料风开度、低位发热量、给粉机转速、一次风量、二次风量、引风机电流、送风机电流、机组负荷、炉膛温度等工艺参数作为相关参数集,然后利用灰色关联分析方法依次分析各变量与烟气含氧量潜在相关性的强弱,以关联性强弱为依据筛选出最终的辅助变量。如表 1 所

示,采用邓氏关联度对辅助变量进行关联分析和筛选,指定相关性阈值为0.76,则主蒸汽流量、主汽压力、主汽温度、给水流量、给粉机转速、一次风量、送风机电流、燃料量等8个工艺参数与氧量的灰关联度满足要求,本文选取这些参数作为辅助变量。

表1 相关工艺参数与氧量邓氏灰色关联分析表

工艺参数	关联度	工艺参数	关联度
主蒸汽流量	0.81	一次风量	0.76
主汽压力	0.76	二次风量	0.74
主汽温度	0.77	引风机电流	0.74
给水流量	0.76	送风机电流	0.77
燃料风开度	0.69	燃料量	0.76
低位发热量	0.69	机组负荷	0.70
给粉机转速	0.78	炉膛温度	0.75

4.2 软测量建模与仿真研究

针对实际电厂运行对锅炉尾部烟气含氧量在线检测的需要,考虑到目前尚无成熟的技术或方法应用于实际,本文对SMO算法进行改进的基础上,结合上述的灰色一致关联理论,进行了基于改进SMO算法的氧量灰色软测量建模研究。模型采用如图1所示结构。

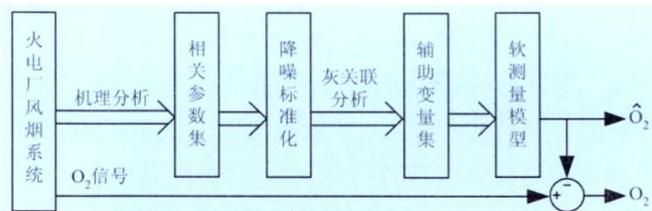


图1 O₂灰色软测量结构框图

在建模中,以4.1节所选择的辅助变量为输入,氧量作为模型输出。仿真数据为天津盘山电厂600MW火电机组正常稳定连续运行24小时的400组历史数据。为了得到更好的建模效果,在建模前对仿真数据进行了曲线拟合滤波降噪处理,并采用极差标准化对数据进行标准化处理,标准化后的数值在[0,1]范围内。

取200组数据进行模型训练,应用SVR参数自适应方法计算出模型参数为: $C=6.3244$, $\epsilon=0.094$,并用剩余200组数据进行模型校验,最终确定RBF核函数参数 $\sigma=0.4$ 时,模型检验均方差达到最优。最终模型的检验均方差 $E=0.018$, $E<0.03$ 能够满足实际应用需要,软测量模型仿真检验结果如图2所示。

前馈神经网络能够以任意精度逼近非线性函数且具有很好的泛化能力,因此该学习算法被广泛应用到工

业过程的建模和控制中,并取得了较好的效果^[6]。BP神经网络是前馈神经网络中应用最为广泛的网络模型,这里应用BP神经网络建立灰色氧量软测量模型,同基于改进SMO算法的软测量检验曲线进行对比。仿真结果如图3所示,该模型的检验均方差 $E=0.033$ 。

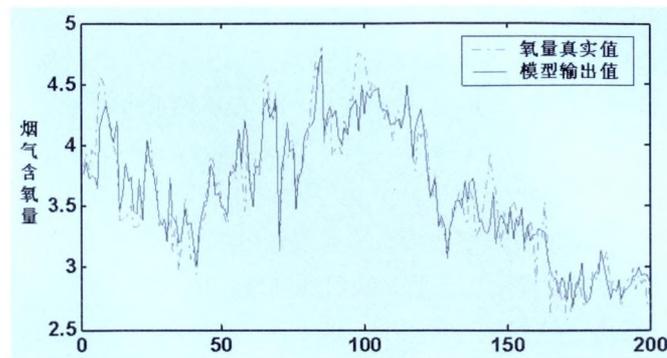


图2 基于改进SMO算法的软测量模型检验曲线

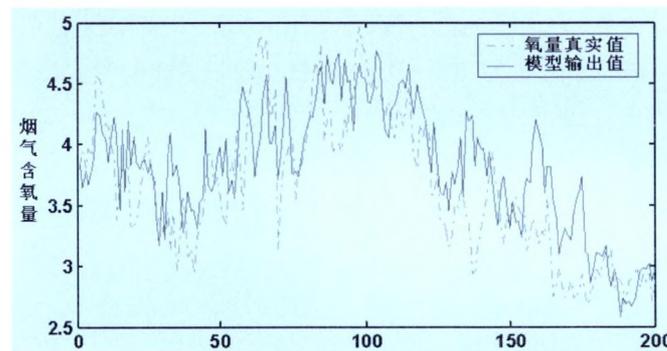


图3 基于BP算法的氧量软测量模型检验曲线

为了验证基于改进SMO算法的灰色软测量模型与普通SMO算法建模精度和收敛速度上的差别,做了如表2所示的试验组(SMO建模选如表1所示的14个工艺参数作为辅助变量)。结果表明,基于改进SMO算法的灰色模型在精度上高于基于SMO算法的普通软测量模型,在收敛时间上改进SMO算法也有比较明显的优势。

表2 模型效果对比试验表

训练 样本数	测试 样本数	改进SMO		SMO	
		检验均方差	收敛时间/s	检验均方差	收敛时间/s
100	300	0.035	14.22	0.037	15.02
150	250	0.026	43.85	0.029	46.33
200	200	0.018	65.53	0.019	76.64

5 结束语

(下转第18页)

静态指标和动态指标,已经很好的满足了设计的要求。上升时间小于5s,超调量小于8%,约为6.67。具体值可由程序计算出。

6 结束语

从该设计我们可以看到,对于一般的控制系统来说,应用PID控制是比较有效的,而且基本不用分析被控对象的机理,只根据 K_p , K_i 和 K_d 的参数特性以及MATLAB绘制的阶跃响应曲线进行设计即可。在MATLAB环境下,我们可以根据仿真曲线来选择PID参数。根据系统的性能指标和一些基本的整定参数的经验,选择不同的PID参数进行仿真,最终确定满意的参数。这样做一方面比较直观,另一方面计算量也比较小,并且便于调整。

(上接第6页)

本文将改进后的SMO算法结合灰色理论应用于大型火电厂烟气含氧量的软测量建模:

1. 基于数据的SVR学习参数获取方法,能有效降低对建模人员经验要求,减少数据噪声对模型精度影响,提高建模效率和模型精度。

2. 对SMO算法内循环迭代优化步长的优化以及阈值评定规则的引入,可以有效提高算法性能和效率。

3. 应用灰色关联分析进行辅助变量选取,降低了模型输入的维数及输入间的耦合程度,提高了软测量模型的精度,具有一定的应用潜力。

改进后的SMO算法结合灰色理论建立的大型火电厂氧量软测量模型能够满足应用的精度要求,为火电厂氧量参数的软测量提供了可能。

参考文献:

[1] 赵征,曾德良,田亮等.基于数据融合的氧量软测量研究[J].中国电机工程学报,2005,25(7):7-12.

[2] 常玉清,邹伟,王福利等.基于支持向量机的软测量方法研究[J].控制与决策,2005,20(11):1307-1310.

(上接第10页)

荷预测过程的优化研究[J].制冷学报,2002,(2):35-38.

[3] 杨行峻,郑君里.人工神经网络[M].北京:高等教育出版社,1992.

[4] 骆维军,李吉生,赵庆珠.蓄冰系统优化控制算法[J].暖通空调,2002,32(2):82-84.

参考文献:

[1] 苏金明,阮沈勇编著.MATLAB6.1使用指南[M].北京:电子工业出版社,2002,1.

[2] 赵文峰等编著.MATLAB控制系统设计与仿真[M].西安:西安电子科技大学出版社,2002,3.

[3] 尹泽明,丁春利等编著.精通MATLAB6[M].北京:清华大学出版社,2002,6.

[4] 梅晓榕主编.自动控制原理[M].北京:科学出版社,2002,9.

作者简介:刘美丽(1979-),女,硕士,助教,研究领域:控制理论与控制工程,计算机。

[3] PLATT.J.Fast training of support vector machines using sequential minimal optimization[M].In B.Schölkopf, C.Burges,and A.Smola,editors,Advances in Kernel Methods-Support vector Learning,MIT Press,1998.

[4] STEVE G.Support vector Machines Classification and Regression[R].ISIS Technical Report.Image Speech & Intelligent Systems Group,University of Southampton.May 10,1998.

[5] V.CHERKASSKE & Y.MA,' Practical selection of SVM parameters and noise estimation for SVM regression' [J].Neural Networks,2004,17(1):113-126.

[6] 翟永杰,杨金芳,徐大平等.应用序列最小优化算法的火电厂协调系统的预测[J].动力工程,2005,25(6):849-854.

[7] 黄福臣,杨永国.基于灰色关联分析的动力煤质量评价[J].能源技术,2007,28(1):11-13,16.

作者简介:乔弘(1981-),男,博士,研究方向:系统建模与智能优化控制。

[5] 袁东立,张钦,朱娜.某冰蓄冷空调系统优化设计探讨[J].暖通空调,2007,(5):24-27.

作者简介:代如静(1984-),女,硕士研究生,研究方向:智能化自动化系统与装置。