

基于最小二乘支持向量机的泵性能分析

万毅

(温州大学物理与电子信息工程学院, 温州 325035)

摘要: 泵的性能曲线是泵选型、优化调度和泵站运行的重要依据, 通常该曲线均是通过试验或是根据试验数据和性能图表上的数据进行曲线拟合而获得, 但这些方法复杂昂贵, 而且拟合精度不高。针对以上方法的缺点, 提出了一种基于交叉验证最优参数选择的最小二乘支持向量机(LSSVM)泵性能预测方法。通过最小二乘支持向量机(LSSVM)学习算法网络的设计和构建, 并应用网络搜索—交叉验证的方法对支持向量机参数进行优化选择, 模拟得到复杂和非线性很强的泵的性能曲线, 经优化模型输出值和试验值、同多项式拟合值以及径向神经网络误差的比较, 交叉验证最优参数选择的最小二乘支持向量机具有优良的非线性建模能力和泛化能力, 在有限学习样本条件下仍获得了很高的精度, 平均相对误差为0.02378%, 为泵的性能分析提供了一种简便可行的智能方法。

关键词: 泵, 支持向量机, 非线性分析, 最优化

doi: 10.3969/j.issn.1002-6819.2009.08.021

中图分类号: TH311

文献标识码: A

文章编号: 1002-6819(2009)-8-0115-04

万毅. 基于最小二乘支持向量机的泵性能分析[J]. 农业工程学报, 2009, 25(8): 115—118.

Wan Yi. Pump performance analysis based on least squares support vector machine[J]. Transactions of the CSAE, 2009, 25(8): 115—118. (in Chinese with English abstract)

0 引言

泵是农业生产中的重要机械设备。泵的性能曲线是泵选型、优化调度和泵站运行的重要依据, 设计人员以泵性能曲线作为参考, 可以合理地选取泵并可以有效地避免所选泵容量偏大, 工作点偏离高效区^[1]。但长期以来, 泵的性能曲线的获得一直是个难点, 人们只能通过大量的试验测量或以泵的试验数据或性能图上的数据进行性能曲线拟合而得到, 试验测量是由生产厂家在成批生产后, 对同一系列产品进行大量的抽样试验, 实测出样机在不同流量点下的扬程、功率、允许吸上真空度及转速等性能参数的数据, 然后进行测试数据的处理, 绘制出一组曲线, 同时用户将泵安装在实际管路系统中后, 也要对泵进行性能测定, 以获取泵实际运转的性能, 这项工作非常复杂而且费用昂贵。以泵的试验数据或性能图上的数据进行拟合而得到的特性曲线的方法虽可减少试验, 但它所得的结果只能保证已知数据附近的准确预测, 远离该数据点的预测则往往与实际差别较大, 甚至很不相同。因此, 既经济又完整可靠地获得泵的性能曲线是泵站机组控制的设计工作者长期以来追求的目标^[1-2]。论文提出了一种基于交叉验证最优参数选择的最小二乘支持向量机(least squares support vector machines, LSSVM)的泵性能预测方法。通过最小二乘支持向量机(LSSVM)学习算法网络的设计和构建, 并应用网络搜索和交叉验证的方法对支持向量机参数进行优化选择, 模拟得到复

杂和非线性很强的泵的性能曲线, 经网络输出数据和试验数据的对比, 交叉验证最优参数选择的最小二乘支持向量机算法具有很强的非线性映射和函数逼近能力, 在小样本条件下, 有效地避免了局部极小和过学习的现象, 获得了很好的效果, 为泵的研究、设计、制造和使用提供很便利的工具。

1 最优参数选择的 SVM 算法

1.1 改进的支持向量机算法 LSSVM

改进的支持向量机—最小二乘支持向量机(LSSVM), 它与标准支持向量机的主要区别在于采用不同的优化目标函数, 并且用等式约束代替不等式约束^[3-7]。

LSSVM算法的目标优化函数为

$$\min_{w,b,e} J(w,e) = \frac{1}{2} W^T W + \frac{1}{2} \gamma \sum_{k=1}^n e_k^2 \quad (1)$$

$$s.t. \quad y_k = W^T \phi(x_k) + b + e_k, k = 1, 2, \dots, N$$

式中: $\phi(\cdot): R^n \rightarrow R^{nh}$ ——核空间映射函数; $w \in R^{nh}$ ——权矢量; $e_k \in R$ ——误差变量; b ——偏置量; γ ——可调参数。

根据式(1), 可定义Lagrange函数

$$L = J(w,e) - \sum_{k=1}^N \alpha_k \{W^T \phi(x_k) + b + e_k - y_k\} \quad (2)$$

式中 α_k ——拉格朗日乘子。

对式(2)求偏导并经过变换得:

$$b = \frac{1_n^T \left[\Omega + \frac{1}{\gamma} I_n \right]^{-1} y}{1_n^T \left[\Omega + \frac{1}{\gamma} I_n \right]^{-1} 1_n}$$

收稿日期: 2008-08-02 修订日期: 2009-08-05

基金项目: 温州市科技局项目(H20080051); 浙江省教育厅项目(20070533)

作者简介: 万毅(1971—), 江西上饶人, 工学博士, 主要研究方向为智能算法和计算机智能控制等。温州 温州大学物理与电子信息学院, 325035。
Email: yiwan246@126.com

$$\alpha = \left[\Omega + \frac{1}{\gamma} I_n \right]^{-1} (y - 1_n b) \quad (3)$$

其中: $y = [y_1; \dots; y_N]^T$, $I_n = [1; \dots; 1]^T$, $\Omega_{kl} = \phi(x_k)^T \phi(x_l)$
 $= K(x_k, x_l)$, $k, l = 1, \dots, N$, $\alpha = [\alpha_1; \dots; \alpha_N]$ 。

用于非线性预测的LSSVM为

$$y(k) = \sum_{k=1}^N \alpha_k K(x, x_k) + b \quad (4)$$

式中 $K(x, x_k)$ —— 径向基核函数, $K(x, x_k) = \exp\{-\|x - x_k\|^2 / \sigma^2\}$ 。

1.2 模型参数的选择

本文采用网格搜索法先选择参数对 (γ, σ) , 然后用交叉验证法对目标函数进行寻优, 直至找到最佳的参数对, 使交叉验证的精度最高^[8]。

应用网格搜索法选择一个参数对 (γ, σ) , 用该参数对进行交叉验证。其方法是把样本集 D 分为 S 组 $\{G_1, G_2, \dots, G_S\}$, 把任意的 $S-1$ 组作为训练集, 剩余的一组作为验证集。通过选择不同的验证集, 可重复 S 次。然后循环选择参数对进行交叉验证, 计算每个参数对的均方误差 MSE, 直到网格搜索停止。

2 泵性能的 LSSVM 预测模型

2.1 模型结构设计和数据预处理

泵的扬程与流量和叶片角度存在着很强的非线性关系。泵的性能曲线是描述在一定的叶片角度下泵的扬程随流量变化的特性, 这种变化特性关系复杂且存在不确定因素。它们的关系可以用下式非线性函数描述:

$$H = f(Q, \theta) \quad (5)$$

式中: Q, θ —— 泵的流量和叶片角; H —— 泵的扬程。把 Q, θ 作为 LSSVM 网络模型的输入层, H 为输出, 输出节点数为 1。性能的 LSSVM 预测模型结构如图 1 所示。

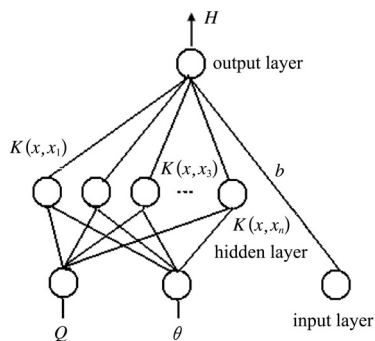


图 1 泵性能的 LSSVM 预测模型结构

Fig.1 Structure of LSSVM model of pump performance prediction

为了提高网络的特性以及给高斯核函数标准差 σ 的选取提供参照系, 先对样本数据进行了预处理, 经过多次仿真试验, 采用归一化处理获得的效果最好。本模型采用高斯核, 即:

$$K(x, x_i) = \exp(-|x - x_i|^2 / \sigma^2)$$

2.2 仿真试验及结果

为验证用交叉验证最优参数选择的最小二乘支持向量机 (LSSVM) 的泵性能曲线预测的精确性和实用性, 开发了泵性能智能预测的通用程序。以 16CJ80 型全调节轴流泵为计算对象^[2,9], 它的扬程与流量、叶片角之间关系的部分试验数据如表 1 所示, 利用这些已知性能曲线的有限的试验数据训练 LSSVM 模型, 再利用剩余数据对训练后模型的预测能力进行验证。采用网格搜索—交叉验证进行 LSSVM 参数寻优, 寻优的遍历图如图 2 所示。

表 1 SVM 和径向基网络 RBFNN 预测值同试验值的比较

Table 1 Comparison between SVM and radial basis function neural network predictive values and experiment values

试验值	SVM 计算值	预测误差	
		SVM	RBFNN
8.55	8.5563	0.0063	0.012
3.55	3.5479	-0.0021	0.032
2.25	2.2467	-0.0033	-0.004
8.80	8.7893	-0.0107	0.004
2.00	1.9979	-0.0021	-0.017
9.00	8.9919	-0.0081	-0.016
2.00	2.0023	0.0023	0.005
9.00	8.9727	-0.0273	0.046
8.70	8.6835	-0.0165	0.029
8.31	8.2969	-0.0131	0.014
4.75	4.792	0.0420	0.035
9.10	9.1055	0.0055	0.006
9.30	9.3079	0.0079	0.008
9.20	9.1996	-0.0004	-0.041
7.10	7.1052	0.0052	0.005
6.05	6.0166	-0.0334	-0.003
2.25	2.2502	0.0002	0.001
9.45	9.4449	-0.0051	0.006
9.50	9.5152	0.0152	0.120
9.40	9.3937	-0.0063	-0.006
9.10	9.1159	0.0159	0.016
8.60	8.5942	-0.0058	0.002
7.93	7.9400	0.0100	-0.005
6.07	6.0753	0.0053	0.007
5.00	5.0011	0.0011	0.002
2.20	2.1998	-0.0002	0.030
9.70	9.7075	0.0075	0.006
9.60	9.5975	-0.0025	-0.005
8.70	8.6991	-0.0009	0.004
7.20	7.2011	0.0011	0.002
6.25	6.2513	0.0013	0.001
2.50	2.5010	0.0010	0.012

最后得到的最佳参数为:

$$\gamma = 39646.9864; \sigma = 0.280660431$$

取出 60 个不同流量和叶片角对应的扬程试验数据与 LSSVM 预测值进行比较, 图 3 为试验值和采用交叉验证优化参数的 LSSVM 方法的计算结果。图 4 为试验值与优化 LSSVM 预测值以及多项式拟合值的相对误差。表 1

为 SVM 和径向基网络 (radial basis function neural network, RBFNN) 预测值同试验值的比较。由图 3、4 以及表 1 可知, 采用交叉验证最优参数选择的最小二乘支持向量机 (LSSVM) 的泵性能预测模型达到了很高的精度, 平均相对误差为 0.02378%。

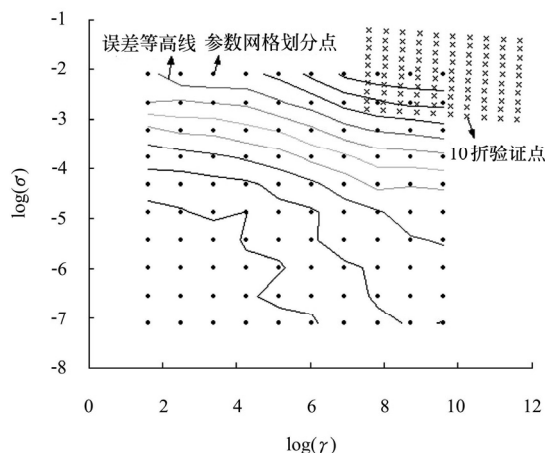


图2 γ, σ 参数网格搜索-交叉验证寻优的遍历图

Fig.2 Optimization traversal process of parameters γ, σ based on grid search-cross validation

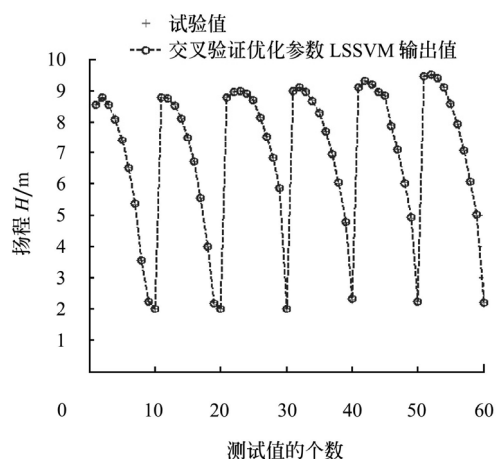


图3 优化参数的LSSVM的泵性能预测结果

Fig.3 Prediction results of pump performance based on LSSVM of parameter optimization

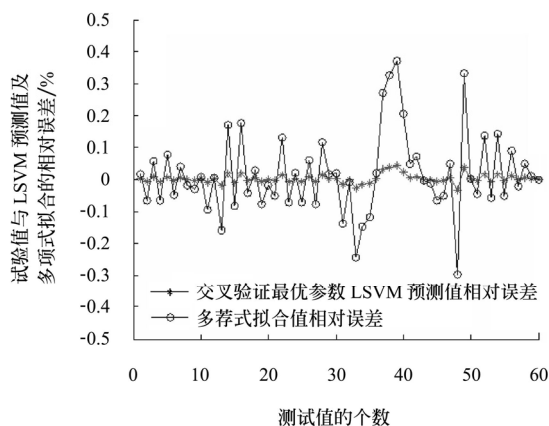


图4 试验值与预测值的相对误差

Fig.4 Relative errors of experiment and predictive values

采用以上设计和训练好的优化模型对叶片角分别为 -5° 和 -7° 在不同的流量下扬程值进行了预测, 它们的性能曲线如图 5 所示。另外应用相同的方法还可以构建“功率-流量”曲线和“效率-流量”曲线预测模型, 从而利用 LSSVM 智能技术可以实现对泵站机组的各种性能进行预测。

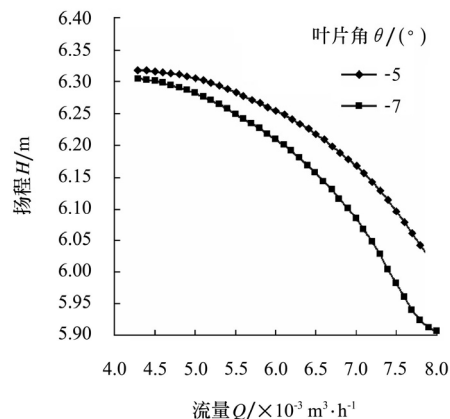


图5 叶片角为 -5° 和 -7° 时的泵的性能曲线

Fig.5 Pump performance curve of the blade angle of -5° and -7°

3 结 语

对改进的支持向量 LSSVM 算法进行了仿真, 并采用网格搜索-交叉验证法对 LSSVM 参数从全局的角度寻找极值点作为最优参数, 它对于小样本学习具有较强的泛化能力, 不过分依赖样本的数量和质量, 具有出色的学习能力和函数逼近能力。

把这种性能优良的交叉验证最优参数选择的最小二乘支持向量机 (LSSVM) 应用于非线性极强的泵的性能分析中, 通过试验和预测值的比较表明该算法在泵性能样本数据较少的情况下, 也表现出较强的推广能力, 使 LSSVM 模型的输出输入关系逼近泵的性能关系函数, 达到了很高的精度。采用这一方法, 可以大大简化泵的性能测定, 选型和优化调度工作, 从而提高工作效率, 这一智能方法适用于各种容积泵的性能分析, 具有广泛的实用价值。

[参 考 文 献]

- [1] 关醒凡. 现代泵技术手册[M]. 北京: 宇航出版社, 1995. Guan Xinghuan. Modern Pump Technology Handbook[M]. Beijing: Space Navigation Press, 1995. (in Chinese)
- [2] 何志霞. 基于神经网络技术的泵站机组预测[J]. 江苏大学学报(自然科学版), 2003, 4(24): 45-48. He Zhixia. Performance Prediction for Pumping Station System Based on Neural Network Technique [J]. Journal of Jiangsu University of Science and Technology, 2003, 4(24): 45-48. (in Chinese with English abstract)
- [3] Jiang Gang, Xiao Jian, Song Chang-lin, A Kind of Peak Load Forecasting Method Based on Support Vector and It'S Parameter Optimization[J]. Control and Decision, 2006,

- 9(21): 1054—1058.
- [4] Nello Cristianini, John S T. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods[M]. Cambridge: University Press, 2000.
- [5] Vapnik V N. Statistical Learning Theory[M]. New York: Springer-Verlag, 2000.
- [6] Hsu C W, Lin C J. A comparison of methods for multi-class support vector machines[J]. IEEE Trans on Neural Networks, 2002, 13(2): 415—425.
- [7] Suykens J A K, Vandewalle J. Least squares support vector machines classifiers[J]. Neural Network Letters, 1999, 9(3): 293—300.
- [8] 郑水波. 支持向量机在汽车动态系统辨识中的应用[J]. 上海交通大学学报, 2005, 39(3): 392—395.
Zheng Shuibao. Application of LSSVMs in the automobile dynamical system identification[J]. Journal of Shanghai Jiaotong University, 2005, 39(3): 392—395. (in Chinese with English abstract)
- [9] 张立群. 泵站综合自动化系统及智能技术的研究与应用[D]. 镇江: 江苏大学, 2001.
Zhang Liqun. Research and Application of Synthesis Automatization System in Pump Station and Intelligent Technology[D]. Zhenjiang: Jiangsu university, 2001. (in Chinese with English abstract)

Pump performance analysis based on least squares support vector machine

Wan Yi

(College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China)

Abstract: Performance curve of pump is important basis of pump types selecting, optimal operation and pump station running. The curve is usually obtained from experiment or by fitting the experimental data of performance graph, but those methods are complex, high expense and imprecision. Based on optimum parameter selection with cross validation, the least squares support vector machine (LSSVM) method was presented for pump performance forecast in the light of the difficulty of the above two methods. Complicated and strong nonlinear pump performance curve was simulated by network design and conformation of LSSVM learning algorithm and the optimized SVM parameters were selected by the method of network searching and cross validation. Compared the errors with output values of the optimized model, test value and output value from polynomial fitting and RBFNN, LSSVM whose parameter was optimized with cross validation had excellent ability of nonlinear modeling and generalization. It gained high precision under limited learning samples (mean relative error is 0.02378%) and provided a simple and feasible intelligent approach for pump performance analysis.

Key words: pumps, support vector machines, nonlinear analysis, optimization