

基于模糊支持向量机方法的短期负荷预测

龚灯才,李训铭,李林峰

(河海大学 电气工程学院,江苏 南京 210098)

摘要: 考虑气象因素对负荷的影响,提出了一种模糊支持向量机 SVM(Support Vector Machine)的短期负荷预测方法。首先选取预测日前 4 星期中差异评价函数小于给定经验值的已知日作为相似日学习样本,然后利用隶属度函数对影响负荷特征因素向量的分量进行模糊处理,得到 SVM 的训练样本集,拟合负荷和影响因素之间的非线性关系。对 24 点每点建立一个 SVM 预测模型,采用改进的序列极小优化算法实现对 SVM 的快速训练。算例数据包括每天的气象数据和 24 点负荷数据,以最大相对误差和平均误差评价预测结果,表明所提方法简便快速且实用有效。

关键词: 短期负荷预测; 支持向量机; 核函数; 隶属度函数

中图分类号: TM 715

文献标识码: A

文章编号: 1006-6047(2005)07-0041-03

短期负荷预测的方法有许多^[1~4],常用的有时间序列法、相似日法和神经网络法等。支持向量机 SVM(Support Vector Machine)^[5]是一种基于结构风险最小化原理的学习技术,也是一种新的具有良好泛化性能的回归方法。近来已有一些将回归 SVM 用于短期电力负荷预测方面的研究^[6,7],均取得了理想的预测结果。本文研究下一日 24 点的日负荷预测,考虑到气象因素对负荷的影响,提出一种基于模糊 SVM 的预测方法。运用该方法对某地区实际负荷数据进行了预测,算例结果表明,该方法具有良好的预测精度和稳定性。

1 回归 SVM 原理^[8]

对于由未知函数随机独立产生的训练样本集 $\{(X_i, y_i)\}_{i=1}^N, X_i \in \mathbf{R}^n, y_i \in \mathbf{R}, N$ 为训练样本个数,回归 SVM 用式(1)估计未知函数:

$$f(X) = \langle \omega, \varphi(X) \rangle + b \quad (1)$$

$\varphi(X)$ 表示由输入空间 X 非线性映射到的高维特征空间。系数 ω 和 b 通过最小化风险泛函(2)得到:

$$\min \omega, b: \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^N |y_i - \langle \omega, \varphi(X_i) \rangle - b|_{\varepsilon} \quad (2)$$

$$|y_i - \langle \omega, \varphi(X_i) \rangle - b|_{\varepsilon} =$$

$$\begin{cases} 0 & |y - \langle \omega, \varphi(X) \rangle - b| \leq \varepsilon \\ |y - \langle \omega, \varphi(X) \rangle - b| - \varepsilon & |y - \langle \omega, \varphi(X) \rangle - b| > \varepsilon \end{cases} \quad (3)$$

式中 第一项 $\|\omega\|^2$ 称为模型复杂性项;第二项是由 ε 不敏感损失函数(3)确定的经验误差项,这个损失函数可带来预测函数(1)的稀疏表达。

通过引入松弛变量 ξ_i 和 ξ_i^* ,将式(2)转换为如下等价形式:

$$\min: \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (4)$$

subject to

$$y_i - \langle \omega, \varphi(X_i) \rangle - b \leq \varepsilon + \xi_i$$

$$\langle \omega, \varphi(X_i) \rangle + b - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i \geq 0, \xi_i^* \geq 0, i = 1, \dots, N$$

由于特征空间的维数很高(甚至无穷)且目标函数不可微,直接求解式(4)几乎是不可行。支撑向量机回归方法的高明之处在于,通过引入点积核函数 $K(X_i, X_j)$ 和利用 Wolfe 对偶技巧避开这些问题,将上述问题转化为下述可有效求解式(4)的对偶问题:

$$\begin{aligned} \max: W(\alpha_i, \alpha_i^*) = & -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(X_i, X_j) - \\ & \varepsilon \sum_{i=1}^N (\alpha_i - \alpha_i^*) + \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) \end{aligned} \quad (5)$$

subject to

$$\sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0, \quad \alpha_i, \alpha_i^* \in [0, c]$$

最后,相应的回归函数式(1)可直接写为

$$f(X) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(X_i, X) + b \quad (6)$$

α_i, α_i^* 为拉格朗日乘子,满足: $\alpha_i \times \alpha_i^* = 0, \alpha_i \geq 0, \alpha_i^* \geq 0$ 。 $K(X_i, X_j)$ 定义为核函数,取为向量 X_i 和 X_j 在特征空间 $\varphi(X_i)$ 和 $\varphi(X_j)$ 中的内积,即 $K(X_i, X_j) = \langle \varphi(X_i), \varphi(X_j) \rangle$ 。任何满足 Mercer 条件的函数均可作为核函数。常用的核函数有多项式核 $K(X_i, X_j) = (\langle X_i, X_j \rangle + 1)^d$ 和高斯核 $K(X_i, X_j) = \exp[-(1/\sigma^2) \times (X_i - X_j)^2]$, 这里的 d 和 σ 是核函数的参数。

基于 KKT 条件,系数 $(\alpha_i - \alpha_i^*)$ 中只有一部分是非零值,并且训练样本的误差 $\geq \varepsilon$, 这些训练样本就是支持向量。式(5)表明,支持向量可以完全确定预测函数,因为其他训练样本对应的系数 $(\alpha_i - \alpha_i^*)$ 均为零。

从应用的角度看, SVM 的训练问题等价于一个线性约束的二次规划问题,其变量数是训练样本数

的 2 倍。由 Scholkopf 和 Smola 提出的序列极小优化 (SMO) 算法可以非常有效地解决这一回归估计问题。本文采用一种改进后的 SMO 算法^[9]训练 SVM。

2 基于模糊 SVM 的短期负荷预测

2.1 训练样本的选择

本文取预测日前 4 星期的已知日作为待选的相似日,当评估函数小于给定小常数 r (经验值) 的已知日作为比较相似的负荷日。评价函数 D 为

$$D = \sqrt{\omega_1(t_{\max} - t_{\max}^p)^2 + \omega_2(t_{\min} - t_{\min}^p)^2} \tag{7}$$

式中 $t_{\max}, t_{\min}$ 为待选日最高和最低气温; $t_{\max}^p, t_{\min}^p$ 为预测日最高和最低气温; ω_1, ω_2 为权值。

考虑前后时段负荷之间的比例,当出现一个超出常规的数值时,可以肯定发生了异常数据现象,修正值可采用正常比例范围内的估计值,对每天的负荷数据逐一加以检查,便可将异常数据选出并校正。

2.2 训练样本的数据处理

为避免出现计算饱和现象,要对负荷数据进行归一化处理,使输入的负荷数据在 $[0, 1]$ 之间, t 时刻负荷数据采用如下归一化公式:

$$\hat{L} = \frac{L_t - L_{\min}}{1.5 L_{\max} - L_{\min}} \quad t = 1, 2, \dots, 24 \tag{8}$$

式中 $L_{\max}, L_{\min}$ 分别为训练样本集中负荷的最大和最小值。

对于影响负荷的因素如平均风速、平均湿度等,也类似上式进行归一化处理。为提高预测精度,使用隶属度函数对输入向量进行模糊化处理,使输入向量的不同分量如温度、天气情况等转化为模糊量。

气温的隶属度函数分别如下:

a. 对低温的隶属度函数采用偏小型梯形分布

$$u_{t1} = \begin{cases} 0 & t > 10 \\ \frac{10 - t}{10 - 0} & 0 \leq t \leq 10 \\ 1 & t < 0 \end{cases} \tag{9}$$

b. 对中温的隶属度函数采用中间型梯形分布

$$u_{t2} = \begin{cases} 0 & t > 25 \text{ 或 } t < 5 \\ \frac{t - 5}{15 - 5} & 5 \leq t \leq 15 \\ \frac{25 - t}{25 - 15} & 15 < t \leq 25 \end{cases} \tag{10}$$

c. 对高温的隶属度函数采用偏大型梯形分布

$$u_{t3} = \begin{cases} 0 & t < 20 \\ \frac{t - 20}{40 - 20} & 20 \leq t \leq 40 \\ 1 & t > 40 \end{cases} \tag{11}$$

将某一相似日的最高气温 t_h 代入式(9)~(11),可分别求出其对于低温、中温、高温 3 个状态的隶属度。根据隶属度最大原则,即可知 t_h 所属的模糊集。

对于最低温度和平均温度,可采用同样的方法求出其 3 个状态隶属度和取值。对天气情况,则采用天气-温度量化系数法。根据 t_h 所属的模糊集,按表 1 量化天气情况。

表 1 天气-温度情况量化系数

Tab.1 The quantified coefficients of weather-temperature

温度状态	量化系数			
	晴	多云	阴	雨
低温	0.05	0.25	0.45	0.95
中温	0.00	0.20	0.40	0.90
高温	0.10	0.30	0.50	1.00

2.3 SVM 模型训练与预测

将这批相似日的影响因素值作为 SVM 的输入样本,每个相似日第 n 点负荷值作为输出样本,对 24 个支持向量机分别进行训练。由于差异评价函数小于 r 的约束,每个相似日的影响向量与预测日影响因素向量在向量空间上都比较接近。利用 SVM 的泛函逼近能力,在向量空间中预测日影响因素向量的附近,拟合负荷和影响因素之间的非线性关系。

将预测日影响因素向量输入到已训练好的 SVM,即可得到预测日 24 点负荷。最后还应修正预测值。由于负荷受多种不确定因素的影响,纯粹根据历史负荷的预测有时难以令人满意,需要运行人员根据经验进行修正。有时预测出的值在某个点可能同负荷变化的规律相违背,如急剧增加或减少,在负荷曲线上表现为突然的尖峰或低谷,这种预测一般不可靠,可以用前后预测值的平均值代替。

3 算例分析

本文采用改进的 SMO 方法实现对 SVM 的快速训练。采用的算例是某地区 2001 年 6 月的负荷资料,包括每天气象数据和每天 24 点的负荷数据。核函数选径向基函数(RBF),取参数 $\sigma = 1$ 。经过实验发现,参数 c 和 ε 对结果影响较大,通过交叉实验,确定相应参数。在本例中取, $c = 0.1, \varepsilon = 0.02$,仿真结果如表 2 所示(δ 为相对误差)。

表 2 计算实例

Tab.2 An example for calculation

时段	L_1/MW	L_2/MW	$\delta/\%$	时段	L_1/MW	L_2/MW	$\delta/\%$
1	5 469	5 557.6	1.62	13	7 262	7 387.3	1.73
2	5 277	5 336.6	1.13	14	7 057	7 273.6	3.07
3	5 107	5 114.7	0.15	15	7 071	7 114.1	0.61
4	5 009	4 978.4	-0.61	16	7 313	7 451.2	1.89
5	5 176	5 284.7	2.10	17	7 384	7 668.3	3.85
6	5 670	5 671.1	0.02	18	7 372	7 525.1	2.08
7	5 846	5 803.3	-0.73	19	7 168	7 355.8	2.62
8	6 018	6 202.2	3.06	20	7 626	7 814.4	2.47
9	6 791	6 780.8	-0.15	21	8 309	8 569.1	3.13
10	7 112	6 989.0	-1.73	22	7 856	7 576.3	3.56
11	7 271	7 260.8	-0.14	23	6 806	7 023.8	3.20
12	7 356	7 517.1	2.19	24	6 003	6 141.0	2.31

由表 2 可以看出,该日 24 个点的最大相对误差绝对值为 3.85%,平均误差为 1.84%,具有非常高的精度。对每个点进行预测,训练程序运行时间都在 1 500 ms 以内,测试时间则小于 20 ms,具有较快的速度。

4 结论

本文提出了一种基于模糊 SVM 的短期负荷预测方法。算法充分利用了多个相似日的信息,并对训练本数据进行归一化和模糊化处理,采用改进的 SMO 算法求解 SVM 训练问题。算例表明,该方法结果合理,运行速度快,精度很高,是一种有应用价值的预测方法。

参考文献:

- [1] 韩民晓,姚蜀军. 短期负荷预测方法的研究及在线应用[J]. 电力系统自动化,1998,22(10):34-37.
HAN Min-xiao,YAO Shu-jun. The research and on-line application of power system short-term load forecasting method[J]. **Automation of Electric Power Systems**, 1998,22(10):34-37.
- [2] 刘晨辉. 电力系统负荷预报理论和方法[M]. 哈尔滨:哈尔滨工业大学出版社,1987.
- [3] 侯志俭. 电力系统短期负荷预报的几种改进手段[J]. 电力系统自动化,1996,20(7):27-31.
HOU Zhi-jian. Several improvements for short-term load forecasting of power systems[J]. **Automation of Electric Power Systems**, 1996,20(7):27-31.
- [4] 程 松. 一种新的基于相似负荷日的短期负荷预计方法[J]. 江苏电机工程,1999,18(4):28-32.
CHENG Song. A new approach to load forecasting based on similar day[J]. **Jiangsu Electrical Engineering**, 1999, 18(4):28-32.

- [5] VAPNIK V N,GOLOWICH S E,SMOLA A J. Support vector machine for function approximation,regression estimation and signal processon[J]. **Adv. Neural Information Processon Syst.**, 1996,(9):281-287.
- [6] 赵登福,王 蒙,张讲社,等. 基于支撑向量机方法的短期负荷预测[J]. 中国电机工程学报,2002,22(4):26-30.
ZHAO Deng-fu,WANG Meng,ZHANG Jiang-she,*et al.* A support vector machine approach for short term load forecasting[J]. **Proceedings of the CSEE**,2002,22(4):26-30.
- [7] 杨镜非,程浩忠. SVM 在电网短期负荷预测中应用研究[J]. 电力自动化设备,2004,24(2):30-32.
YANG Jing-fei,CHENG Hao-zhong. Application of SVM to power system short-term load forecast[J]. **Electric Power Automation Equipment**,2004,24(2):30-32.
- [8] VAPNIK V N. The nature of statistical learning theory[M]. New York:Spring-verlag,1995.
- [9] 张浩然,韩正之. 回归支持向量机的改进序列最小优化学习算法[J]. 软件学报,2003,14(12):2006-2013.
ZHANG Hao-ran,HAN Zheng-zhi. An improved sequential minimal optimization learning algorithm for regression support vector machine[J]. **Journal of Software**,2003, 14(12):2006-2013.

(责任编辑:李育燕)

作者简介:

龚灯才(1981-),男,江苏姜堰人,硕士研究生,研究方向为短期电力负荷预测与负荷建模(E-mail:gdc_hhu@mailsvr.hhu.edu.cn);

李训铭(1949-),男,江苏南京人,教授,主要从事现代控制理论方面的教学与研究工作;

李林峰(1982-),男,江苏如东人,硕士研究生,研究方向为短期电力负荷预测与负荷建模。

Short-term load forecast based on fuzzy support vector machine method

GONG Deng-cai,LI Xun-ming,LI Lin-feng

(College of Electrical Engineering,Hohai University,Nanjing 210098,China)

Abstract: Considering the influence of weather on loads,a method based on fuzzy SVM(Support Vector Machine) is presented for the short-term load forecast. The days during last four weeks, whose difference estimation function values are smaller than the given experiential value,are selected as similar days for learning. The influencing factors of eigenvector are processed by fuzzy membership functions to form the training sample set for SVM and fit the nonlinear relationship between loads and influencing factors. SVM forecast models are established for every point of 24-point loads,and an improved sequential minimal optimization method is used to train SVM. The example provides daily weather data and 24-point load data,and the maximum relative error and mean error are selected to evaluate the forecast. Results show the proposed method simple,fast, practical and effective.

Key words: short-term load forecast; support vector machine; kernel function; membership function