

# 带扩展记忆的粒子群优化最小二乘支持向量机在中长期 电力负荷预测中的应用

段其昌 周华鑫 曾 勇 张广峰  
(重庆大学自动化学院 重庆 400044)

**摘 要** 针对中长期电力负荷预测,考虑了影响中长期电力负荷的各种因素,提出了一种带扩展记忆的粒子群优化算法(PSOEM)与最小二乘支持向量机(LSSVM)相结合的中长期负荷预测方法。PSOEM比传统PSO收敛速度更快,精度更高,具有更强的寻优能力,因此利用 PSOEM算法对 LSSVM 参数进行优化选择,获得了较优的 PSOEM-LSSVM 预测模型。通过实例仿真表明,该方法与其他几种方法相比具有更高的预测精度和速度。

**关键词** 中长期负荷预测,带扩展记忆粒子群,最小二乘支持向量机

**中图分类号** TP391 **文献标识码** A

## Application of PSOEM-LSSVM in Medium and Long Term Power Load Forecasting

DUAN Qi-chang ZHOU Hua-xin ZENG Yong ZHANG Guang-feng  
(College of Automation, Chongqing University, Chongqing 400044, China)

**Abstract** For medium and long term power load forecasting, considering the factors affecting medium and long-term power load, a method based on the particle swarm optimization with Extended Memory (PSOEM) and Least squares support vector machines (LSSVM) was proposed for medium and long-term load forecasting. PSOEM has more extensive capability of global optimization than PSO owing to higher accuracy and convergence rate, so it is used to optimize the parameters of the LSSVM, and obtains an optimum PSOEM-LSSVM model to forecast the load. An example simulation shows that this method can offer a higher precision and speed of forecasting than several other methods.

**Keywords** Medium and long term load forecasting, PSOEM, LSSVM

电力负荷预测中经常按时间期限进行分类,通常分为长期、中期、短期和超短期负荷预测。中长期负荷预测一般指5年左右到10年以上并以年为单位的预测,其意义在于帮助决定新的发电机组的安装与电网的规划、增容和改建,是电力规划部门的重要工作之一<sup>[1]</sup>。

中长期电力负荷受经济、社会、气候等不确定因素影响,传统的中长期负荷预测方法难以模拟这些复杂因素对中长期负荷变化的影响<sup>[2]</sup>。长期以来,国内外学者对负荷预测的理论和方法做了大量的研究,如灰色理论法<sup>[3,4]</sup>、神经网络方法<sup>[5,6]</sup>、支持向量机方法<sup>[7,8]</sup>等,以及综合各单一模型的组合预测法。

支持向量机(Support Vector Machine, SVM)是由 V. Vapnik 等提出的基于统计学习理论的实用化机器学习方法,能够较好地解决小样本、非线性、高维数和局部最小点等实际应用问题<sup>[9]</sup>。支持向量机克服了神经网络方法存在的泛化能力较差、易陷入局部极小、收敛慢等缺陷,被认为是神经网络的替代方法<sup>[10,11]</sup>。最小二乘支持向量机(Least Squares Support Vector Machines, LSSVM)则是标准支持向量机的改进和扩展, LSSVM 预测的效果在很大程度上受到其关键参数选择的影响<sup>[12]</sup>。因此,本文提出了一种采用带扩展记忆的粒子群优化算法对支持向量回归参数寻优选择的方法,构建了

基于 PSOEM-LSSVM 的中长期负荷预测模型,以提高预测精度和速度。

### 1 最小二乘支持向量机

支持向量机是一种基于统计学习理论模式识别的方法,它是由 Vapnik 等在 20 世纪 90 年代中期提出的一种新的机器学习方法<sup>[13]</sup>。最小二乘支持向量机是由学者 Suykens<sup>[14]</sup>提出的,是标准支持向量机的改进和扩展,它用等式约束代替 SVM 的不等式约束,将 SVM 的二次规划问题转化为线性方程组求解,降低了计算复杂度,提高了求解速度和收敛速度,能够较好地解决小样本、非线性、高维数和局部极小等实际问题。

假设训练集  $T = \{(x_i, y_i) | i = 1, 2, \dots, n\}$ , 其中  $x_i \in R^d$  为输入变量,  $y_i \in R$  为相应的输出量,  $n$  为样本个数, 采用一个非线性映射  $\varphi: R^d \rightarrow R^k (k \geq d)$  将输入空间映射到高维的特征空间, 再在特征空间中进行线性回归。设线性回归函数为:

$$f(x) = w^T \varphi(x) + b \tag{1}$$

式中,  $w \in R^k$  为权向量,  $b \in R$  为常数。

为使实际风险最小, 根据 SRM 准则, 回归问题的优化目标和约束条件分别为:

段其昌(1953—),男,教授,博士生导师,主要研究方向为智能控制算法、电力市场技术等;周华鑫(1987—),男,硕士生,主要研究方向为智能控制算法;曾 勇(1985—),男,硕士生,主要研究方向为电网规划;张广峰(1987—),男,硕士生,主要研究方向为智能控制算法。

$$\min \frac{1}{2} \|w\|^2 + \frac{1}{2} c \sum_{i=1}^n \xi_i^2 \quad (2)$$

$$\text{s. t. } y_i - w^T \varphi(x_i) - b = \xi_i, i=1, 2, \dots, n$$

式中,  $\xi_i \geq 0$  为松弛变量,  $c > 0$  为惩罚因子。为求解上述优化问题, 引入拉格朗日乘子  $\alpha_i$  得到拉格朗日函数, 即:

$$L = \frac{1}{2} \|w\|^2 + \frac{1}{2} c \sum_{i=1}^n \xi_i^2 - \sum_{i=1}^n \alpha_i [w^T \varphi(x_i) + b + \xi_i - y_i] \quad (3)$$

根据 KKT 条件, 对于  $i=1, 2, \dots, n$ , 消去  $\xi_i$  和  $w$  后, 得线性方程组:

$$\begin{bmatrix} 0 & E^T \\ E & ZZ^T + C^{-1}e \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (4)$$

式中,  $E$  为元素为 1 的  $n \times 1$  向量,  $e$  为  $n \times n$  阶单位矩阵,  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ ,  $y = [y_1, y_2, \dots, y_n]^T$ ,  $Z = [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_n)]^T$ 。

根据 Mercer 条件, 定义核函数:

$$k(x, y) = \varphi^T(x) \varphi(y) \quad (5)$$

上述线性方程组可用最小二乘法求出  $b$  和  $\alpha$ , 得到的 LSSVM 函数为

$$f(x) = \sum_{i=1}^n \alpha_i k(x_i, x_j) + b \quad (6)$$

常用的核函数有线性核、多项式核、径向基核等, 根据中长期负荷预测样本值特点, 本文采用具有较好泛化能力的径向基核函数, 其表达式如下:

$$k(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\delta^2}\right) \quad (7)$$

## 2 带扩展记忆的粒子群优化算法

粒子群优化 (Particle Swarm Optimization, PSO) 算法是 1995 年由美国博士 Kennedy 和 Eberhart 受鸟群觅食行为启发提出的基于自然系统的一种进化计算方法<sup>[15]</sup>。标准粒子群算法中第  $i$  个粒子看成是在  $d$  维搜索空间里的向量  $X_i = (x_{i1}, \dots, x_{id})$ , 即第  $i$  个粒子的位置, 每个粒子的位置就是一个潜在的解。将  $X_i$  代入一个目标函数就可以计算出其适应度值, 根据适应度值的大小衡量  $X_i$  的优劣。第  $i$  个粒子的速度也是一个  $d$  维的向量, 记为  $V_i = (v_{i1}, \dots, v_{id})$ 。在迭代过程中, 记第  $i$  个粒子搜索到的最优位置为个体极值, 用  $P_{id} = (p_{i1}, p_{i2}, \dots, p_{id})$  表示。而所有粒子搜索到的最优位置成为全局极值, 用  $P_{gd} = (p_{g1}, p_{g2}, \dots, p_{gd})$  表示。每一次迭代中, 每一个粒子都会调整其在每一维空间的速度并计算出新的位置, 算法表示如下:

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (p_{gd}^k - x_{id}^k) \quad (8)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (9)$$

式中,  $\omega$  称为惯性权重因子; 非负常数  $c_1$  和  $c_2$  称为加速度因子;  $r_1$  和  $r_2$  是介于  $(0, 1)$  之间的随机数;  $v_{id}^k$  表示粒子在第  $k$  次迭代的速度,  $x_{id}^k$  表示粒子第  $k$  次迭代的位置,  $p_{id}^k$  表示粒子在  $k$  次迭代过程中的个体极值,  $p_{gd}^k$  表示种群在  $k$  次迭代过程中目前的全局极值; 一般事先设置速度的上下边界  $v_{\max}$  和  $v_{\min}$ , 防止粒子远离搜索空间。

本文应用带扩展记忆的粒子优化算法 (PSO with Extended Memory, PSOEM), 通过扩展记忆 (Extended Memo-

ry) 来存储粒子近来的位置、个体极值点和全局极值点。PSOEM 算法可用如下公式来表示:

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 [\xi_i (p_{id}^k - x_{id}^k) + \xi_{i-1} (p_{id}^{k-1} - x_{id}^{k-1})] + c_2 r_2 [\xi_i (p_{gd}^k - x_{id}^k) + \xi_{i-1} (p_{gd}^{k-1} - x_{id}^{k-1})] \quad (10)$$

式中,  $\xi_i, \xi_{i-1} \in R^+$ ,  $\sum_{i=0}^k \xi_{i-1} = 1$ ,  $\xi_i$  称为当前有效因子,  $\xi_{i-1}$  称为扩展记忆有效因子。特别地, 当  $\xi_{i-1} = 0$  时  $\xi_i = 1$ , PSO 是 PSOEM 的一种特殊情况。

仿真分析比较证明, PSOEM 比 PSO 的搜索方向性、目的性更强, 且收敛速度更快, 精度更高, 更适合优化多变量函数。

## 3 基于 PSOEM-LSSVM 的中长期负荷预测方法

中长期电力负荷预测以年为单位进行, 负荷受到经济、人口、气候等不确定因素的影响。该方法应用电力历史年负荷数据, 考虑 GDP、人口和气候因素, 建立基于 PSOEM-LSSVM 的中长期负荷预测模型。

基于 PSOEM-LSSVM 的短期负荷预测流程如下:

(1) 对历史负荷数据进行预处理, 包括对缺失数据的补充、异常数据的修正舍弃和数据的归一化处理, 形成训练样本数据集。

(2) 训练一次 LSSVM, 初始化设置 LSSVM 参数  $c = 10$ ,  $\delta^2 = 2.5$ , 通过模型输出预测结果, 再根据如下公式计算适应度值。

$$I = \frac{1}{L} \sum_{i=1}^L |y_i - y_i'| \quad (11)$$

(3) 对 PSOEM 参数进行初始化设置, 本文采用线性递减的惯性权重方法进行仿真,  $\omega$  从 0.9 到 0.4 线性递减, 种群规模  $d = 30$ , 迭代次数  $k = 300$ , 加速度因子  $c_1 = 2$ ,  $c_2 = 2$ , 当前有效因子  $\xi_i = 0.6$ , 扩展记忆有效因子  $\xi_{i-1} = 0.4$ 。

(4) 把寻优目标参数  $c, \delta$  作为粒子, 它们的初始值分别为 (2) 中的初始化值, 并将它们作为当前个体最优解。

(5) 将 (2) 中的适应度值作为粒子的初始适应度值, 搜索其中的最优值作为全局最优适应度值, 并将其相对的粒子作为当前全局最优解。

(6) 迭代寻优过程开始, 根据 PSOEM 式 (9) 和式 (10) 搜索更新粒子的速度和位置。

(7) 将更新后的粒子值重新带入 LSSVM 模型, 重新进行 LSSVM 训练, 保存其输出结果, 计算粒子的适应度值。

(8) 将 (7) 中的适应度值与当前的粒子适应度值作比较, 如果优于当前的粒子适应度值, 则更新当前的粒子适应度值, 并将当前的粒子相对应的个体最优值更新为与 (7) 中适应度值相对应的粒子值。

(9) 如果当前的粒子适应度值优于全局最优适应度值, 则更新全局最优适应度值为当前的粒子适应度值, 并更新全局最优解为当前粒子的粒子适应度值所对应的粒子值。

(10) 判断是否达到迭代次数, 若是则终止迭代, 否则返回 (6)。

(11) 输出最优解。

## 4 实例分析

本文对某地区中长期负荷预测考虑了将经济、人口、气候

作为影响因素,将该地区 1997—2008 年电网的年负荷原始数据作为样本,具体数据如表 1 所列。

表 1 某地区 1997—2008 年负荷样本数据

年份	GDP / 亿元	人口 / 万	冬季平均气温 / °C	夏季平均气温 / °C	用电量 / 亿 KWh
1997	316.1760	424.8	7.05	23.95	27.1880
1998	374.4048	426.0	7.10	23.83	30.5533
1999	405.1348	427.6	7.37	23.47	35.6680
2000	436.1784	427.8	7.43	23.33	39.8876
2001	465.0548	430.0	6.93	23.27	43.5769
2002	509.9332	431.7	7.23	23.60	44.3338
2003	569.0944	431.4	7.97	23.20	44.5311
2004	628.9136	431.0	8.13	23.77	54.4839
2005	706.1040	431.6	7.67	23.30	87.4739
2006	828.2988	432.0	6.80	23.60	85.3555
2007	1013.6760	432.2	7.00	25.43	104.7661
2008	1197.0340	432.3	7.15	23.90	123.6139

利用 Matlab 软件对样本进行训练建模,利用训练后的模型来预测 2005—2008 年的负荷,将预测结果与其他几种预测方法进行比较,包括 LSSVM、BP 神经网络和灰色预测模型 GM(1,1)方法。

本文采用平均相对误差(MAPE)作为评价指标,MAPE 定义为:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i' - y_i}{y_i} \right| \times 100\% \quad (12)$$

式中,  $N$  为预测个数,  $y_i'$  为预测值,  $y_i$  为实际值。

表 2 为基于 PSOEM-LSSVM 的预测模型对 2005—2008 年进行负荷预测的结果和平均相对误差与其他几种预测方法的比较分析。

表 2 某地区 2005—2008 年负荷预测误差分析

年份	实际值 / 亿 KWh	PSOEM-LSSVM / 亿 KWh	LSSVM / 亿 KWh	BP 神经网络 / 亿 KWh	GM(1,1) / 亿 KWh
2005	87.4739	88.8509	89.7683	90.7683	80.8028
2006	85.3555	84.4042	81.9989	80.9989	90.5563
2007	104.7661	105.9887	110.0081	111.0081	98.3240
2008	123.6139	122.5241	126.2346	127.2346	117.4928
平均相对误差 / %		1.18	3.42	4.44	6.21

图 1 为本文方法与其他几种方法的负荷预测结果曲线。

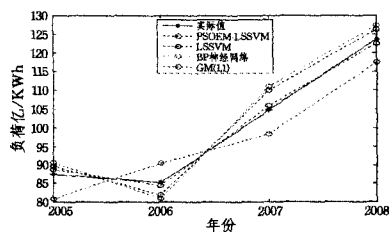


图 1 2005—2008 年负荷预测结果曲线

通过图 1 和表 2 分析对比结果可以看出,基于 PSOEM-LSSVM 预测方法的平均相对误差为 1.18%,远远小于其他几种方法,原因在于带扩展记忆的粒子群优化算法搜索方向性、目的性较强。因此,基于 PSOEM-LSSVM 的预测方法能

够有效地提高中长期电力负荷的预测精度,其稳定性较强,收敛速度较快,适合于电力投资规划的要求。

**结束语** 中长期电力负荷受到多种复杂因素的影响,其预测精度的高低直接影响到城市电网规划质量的优劣。本文提出的基于 PSOEM-SVR 的中长期电力负荷预测方法,通过 PSOEM 算法准确快速地优化组合核函数 LSSVM 参数,克服了传统 LSSVM 参数设置的盲目低效性。通过实例分析验证,该方法可以利用较少的历史数据对中长期电力负荷进行预测,并且能够极大地提高负荷预测的精度和速度。

## 参 考 文 献

- [1] 牛东晓,曹树华,等. 电力负荷预测技术及其应用[M]. 北京:中国电力出版社,1998
- [2] 卜虎正,姚建刚,李文杰,等. 中长期电力负荷预测的改进免疫粒子群算法[J]. 电力系统及其自动化学报,2011,23(3):139-144
- [3] 王成山,杨军,张崇见. 灰色系统理论在城市年用电量预测中的应用:不同预测方法的分析比较[J]. 电网技术,1999,23(2):15-18
- [4] 王翠娟,孙辰军,杨静,等. 改进残差灰色预测模型在负荷预测中的应用[J]. 电力系统及其自动化学报,2006,18(1):86-89
- [5] 胡越黎,计惠杰. 短期电力负荷预测的模糊神经网络方法[J]. 应用科学学报,2009,27(1):79-83
- [6] 高海龙,张国立. 改进遗传神经网络及其在负荷预测中的应用[J]. 华北电力大学学报,2009,36(5):37-40
- [7] 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
- [8] 李元诚,方廷健,于尔铿. 短期负荷预测的支持向量机方法研究[J]. 中国电机工程学报,2003,23(6):55-59
- [9] 章健. 电力系统复合模型与辨识[M]. 北京:中国电力出版社,2007
- [10] Zhang Ming-guang. Short-term load forecasting based on support vector machines regression[C]//Proceedings of 2005 International Conference on Machine Learning and Cybernetics. Lanzhou, China, 2005,7:4310-4314
- [11] Chen Bo-juen, Chang Ming-wei, Lin Chih-Jen. Load forecasting using support vector machines; a study on EUNITE competition 2001[J]. IEEE Transactions on Power Systems, 2004, 19(4): 1821-1830
- [12] 孙翠娟. 基于 K 型核函数的支持向量机[J]. 淮海工学院学报:自然科学版,2006,15(4):4-7
- [13] Vapnik N. The Nature of Statistical Learning Theory[M]. New-york: Springer-Verlag, 1995
- [14] Suykens J A K, Vandewalle J. Least squares support vector machines classifiers[J]. Neural Network Letters, 1999, 19(3): 293-300
- [15] Kennedy J, Eberhart R. Particle swarm optimization [C]//IEEE International Conference on Neural Networks. Perth, Australia, 1995,4:1942-1948