

doi:10.3969/j.issn.1672-4348.2017.04.012

短期电力负荷时间序列预测的极限学习机方法

张宁

(闽江学院 物理学与电子信息工程系, 福建 福州 350108)

摘要: 提出将一种进化的神经网络模型——极限学习机应用于短期电力负荷时间序列预测中, 该方法具有模型参数设置少、训练速度快和良好的泛化能力等明显优点。通过实例分析表明该模型的预测精度要优于 BP 神经网络模型, 同时也验证了该模型应用于短期负荷预测的有效性和可行性。

关键词: 极限学习机; BP 神经网络模型; 短期电力负荷; 时间序列

中图分类号: TM715 **文献标志码:** A **文章编号:** 1672-4348(2017)04-0367-04

Extreme learning machine method for short-term power load time series forecasting

Zhang Ning

(Physics & Electronic Information Engineering Department, Minjiang University, Fuzhou 350108, China)

Abstract: The utilization of extreme learning machine—an evolution model of neural networks in short-term power load time series forecasting was proposed, which has the advantages of less parameter setting, fast training speed and better generalization ability. The validity and feasibility of the model were verified. The simulation results indicate that the forecasting accuracy of the model is better than that of BP neural network model.

Keywords: extreme learning machine; BP neural network model; short-term power load; time series forecasting

对短期电力负荷进行准确预测是电力系统安全稳定运行的重要前提。预测短期电力负荷经常根据气温、气压、相对湿度、降雨量、节假日等多种影响因素来进行建模。如果这些影响因素和电力负荷之间有良好的相关性, 建立一定的数学模型可以获得较好的预测效果。但是, 影响因素的取法众多, 气温、气压、相对湿度等因素之间的组合也有多种形式, 往往需根据经验来进行确定。为消除这些不确定组合方式的影响, 另一种预测短期负荷的处理方法是将历史负荷值作为一时间序列进行建模分析。在此基础上, 人们提出了灰色模型预报法、时间序列分析法、Kalman 滤波法等^[1-3]多种负荷预测方法, 并取得了较好的应用

效果。

随着机器学习理论不断发展, 以神经网络为代表的非线性方法在负荷预测中获得了较为成功的应用^[4-5]。但神经网络方法存在参数设置较多、训练速度较慢、容易陷入局部极小等缺陷。为此, 提出将极限学习机(extreme learning machine, ELM)模型应用于短期负荷预测中。该方法是由南洋理工大学 Huang G B 博士在 2006 年提出, 为一种进化的神经网络方法, 具有较好的泛化能力, 并能避免传统神经网络方法存在的诸多问题^[6]。应用该方法对实际负荷数据进行预测实验分析, 结果验证了所提方法的可行性, 并具有较高的预测精度。

收稿日期: 2017-07-20
基金项目: 福建省中青年教师教育科研项目(JAT170437)
通讯作者: 张宁(1984-), 女, 福建罗源人, 讲师, 硕士, 研究方向: 信号处理、智能优化算法。

1 极限学习机的基本原理

极限学习机是一种新型神经网络训练算法^[7-9],对于 N 个训练样本 $(x_i, y_i), i = 1, 2, \dots, N$, 其中 $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$ 为输入变量, $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T \in \mathbf{R}^m$ 为输出变量, 其 ELM 数学模型可以表示为:

$$\sum_{i=1}^N \beta_i g_i(x_j) = \sum_{i=1}^N \beta_i g(w_i \cdot x_j + b_i) = o_j, j = 1, \dots, N \quad (1)$$

式中, \tilde{N} 为隐层节点个数, $g(x)$ 为激活函数, $o_j = [o_{j1}, o_{j2}, \dots, o_{jm}]^T \in \mathbf{R}^m$ 为第 j 个样本输出值, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ 是连接 i 个隐层节点的输出权值, $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ 为连接第 i 个隐含层节点的输入权值; b 是 i 个隐层节点的偏置。

设 ELM 模型可以零误差逼近 N 个训练样本, 即 $\sum_{j=1}^N \|o_j - y_j\| = 0$ 。将存在 β_i, w_i, b_i 可得:

$$\sum_{i=1}^N \beta_i g(w_i \cdot x_j + b_i) = y_j, j = 1, \dots, N \quad (2)$$

上式可表示成矩阵形式为:

$$\mathbf{H}\mathbf{\beta} = \mathbf{Y} \quad (3)$$

其中:

$$\mathbf{H} =$$

$$\begin{pmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_N + b_{\tilde{N}}) \end{pmatrix}_{N \times \tilde{N}},$$

$$\mathbf{\beta} = \begin{pmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{pmatrix}_{\tilde{N} \times m}, \mathbf{Y} = \begin{pmatrix} y_1^T \\ \vdots \\ y_N^T \end{pmatrix}_{N \times m} \quad (4)$$

ELM 模型的训练目标是寻找最优权值 $\mathbf{W} = (w, b, \beta)$, 以使得下式成立:

$$\min E(\mathbf{W}) = \min \|\mathbf{H}\mathbf{\beta} - \mathbf{Y}\| \quad (5)$$

ELM 模型与传统 BP 神经网络模型的差异在于: BP 神经网络模型采用梯度下降方法对 \mathbf{W} 进行调整; 而 ELM 模型是将 (3) 式作为一线性方程组进行求解, 其输出权值 β 的最小二乘解为:

$$\mathbf{\beta} = \mathbf{H}^+ \mathbf{Y} \quad (6)$$

其中, \mathbf{H}^+ 为隐层输出矩阵 \mathbf{H} 的 Moore-Penrose 广义逆^[7]。则 ELM 模型的网络输出 \mathbf{F} 为:

$$\mathbf{F} = \mathbf{H}\mathbf{\beta} \quad (7)$$

从上述原理可以看出, ELM 模型具有训练误差最小、可得到权值的最小范式等显著的优越性。

2 基于 ELM 的短期电力负荷预测模型

2.1 样本数据集的构建

对于一实际短期负荷时间序列, 我们假定 t 时刻的负荷值 $X(t)$ 可以由 $(t-1, t-2, \dots, t-m)$ 时刻的历史负荷值 $X(t-1), X(t-2), \dots, X(t-m)$ 来进行预测, 则预测模型可以表示成:

$$X(t) = f[X(t-1), X(t-2), \dots, X(t-m)] \quad (8)$$

其中, $f(\cdot)$ 为映射函数, m 为嵌入维数。

由此, 在选择一定的嵌入维数, 即确定 ELM 网络输入个数的情况下, 可以构造出一个多输入单输出的 ELM 模型。

2.2 短期负荷时间序列数据预处理

为了减少数据量纲对建模的影响, 对原始短期负荷时间序列数据进行归一化处理, 即使用下式将数据归一至 $[-1, 1]$:

$$X'(t) = 2(X(t) - X_{\min}) / (X_{\max} - X_{\min}) - 1 \quad (9)$$

其中, X_{\min} 、 X_{\max} 为原序列中的最小值、最大值, $X'(t)$ 为经归一化后的负荷序列。在建模完成之后, 对于输出的训练、预测数据, 再进行反归一化处理, 将其还原至原始区间。

2.3 模型的评价指标

为了定量评价模型的负荷预测精度, 选取两种常用的评价指标: (1) 平均相对误差: $\text{MAPE} =$

$$\frac{1}{n} \sum_{t=1}^n \frac{|X_t - \hat{X}_t|}{|X_t|}; (2) \text{均方根误差: } \text{RMSE} =$$

$$\sqrt{\frac{1}{n} \sum_{t=1}^n (X_t - \hat{X}_t)^2}, \text{ 其中 } X_t, \hat{X}_t \text{ 分别为实际负荷与预测负荷, } n \text{ 为预测样本数目。}$$

3 实例分析

以国家电网浙江省台州某地区 2010-02-14 ~ 2010-03-27 共 42 d 的电力负荷数据为研究对象。采集的数据为每天整点时刻的负荷值, 共计 1 008 个, 形成短期负荷时间序列。如图 1 所示。

对短期负荷时间序列构建样本集时, 在采用不同嵌入维数的情况下, 根据式 (8) 构建的样本数目是不相同的。为了方便数据间的对比, 将嵌入维数分别设置为 48、24、12, 即由前两天、前一

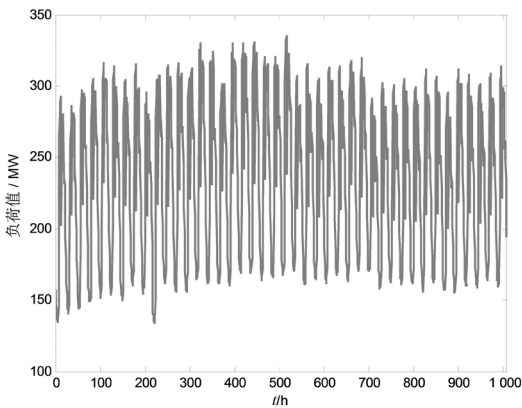


图 1 原始短期负荷数据序列

Fig.1 Original power load data series

天、前半天的负荷数据预测下一整点时刻的负荷值。由此,总样本数目可构建为 960、984、996 个。

表 1 两种模型平均预测精度和 CPU 时间

Tab.1 Average predicted precision and CPU time for two models

嵌入 维数	ELM 预测精度		BP 预测精度		ELM CPU t/s	BP CPU t/s
	RMSE/MW	MAPE	RMSE/MW	MAPE		
48	4.902 4	0.016 6	6.019 2	0.017 2	0.011 5	54.990 4
24	3.313 3	0.012 3	4.682 7	0.015 8	0.011 3	24.133 4
12	5.479 1	0.017 5	5.977 3	0.016 9	0.011 2	5.725 2

从表 1 中可以看出:(1)在设置不同嵌入维数下,ELM 模型的预测精度都要优于 BP 神经网络模型,表明 ELM 模型具有良好的学习能力和泛化能力;(2)从 CPU 运行时间上对比,两种模型不在一个数量级上,ELM 模型的运行速度要远快于 BP 神经网络模型,这是由于 ELM 直接解算线性方程组,而 BP 需要不断进行反向迭代运算造成的;(3)在嵌入维数取值为 24,即以前一天的数据进行预测下一时刻负荷值时其预测精度最高。

以嵌入维数为 24 为例,预测最后一天 24 h 的负荷值,其预测值与实际负荷值的对比情况如图 2 所示,预测负荷的绝对误差如图 3 所示。

从图 1 中可以看出,所采用的 ELM 预测模型的预测结果曲线与真实负荷值曲线非常吻合。从图 2 中可以看出,ELM 模型的预测绝对误差值分布在±5 MW 范围之内,表明了 ELM 模型具有较高的预测精度。

统一以最后一天 24 个负荷值为预测对象,则训练样本数目设置为 936、960、972 个,预测样本数目均为 24 个。以此为基础,建立 ELM 模型和常用的 BP 神经网络模型。其中 BP 模型采用稳健的 Levenberg-Marquardt 算法。由于两种模型原理的不同,ELM、BP 的隐层神经元数对结果的影响也不同,一般 ELM 需设置较多的隐层神经元数。在本研究中,两种模型的隐层神经元数分别设置为 30、8。激活函数统一选择为“Sig”函数。由于 ELM 和 BP 均采用随机权值,会使得模型预测单次结果出现不稳定现象,为此在进行预测时,各循环操作了 50 次,并取 50 次预测均值作为最终预测结果。

表 1 给出了不同嵌入维数下 ELM 模型、BP 模型进行预测的精度统计及消耗的 CPU 时间。

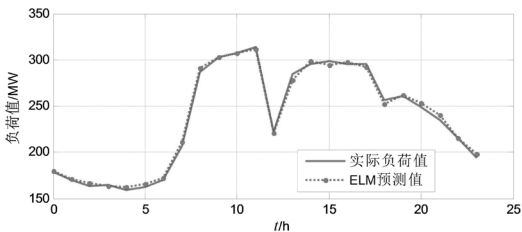


图 2 实际负荷值与预测负荷值对比

Fig.2 Comparison between practical load and predicted load

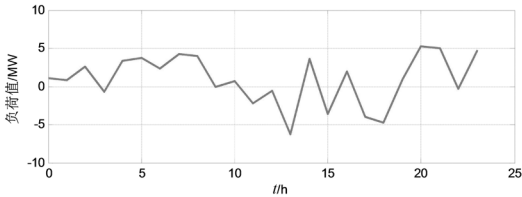


图 3 负荷预测绝对误差

Fig.3 Absolute error of load forecasting

4 结论

(1)将 ELM 模型应用于短期负荷时间序列预测中,取得了较高的预测精度,表明 ELM 模型在短期负荷预测中应用的可行性和有效性。

(2)本研究在对负荷时间序列进行样本构建时,其嵌入维数是根据经验来进行选取的,在实际中由于负荷时间序列具有混沌特性,可进一步通

过相空间重构计算来选取最佳嵌入维数,这将在下一步的研究中予以实现。

(3)所采用的 ELM 模型预测精度虽较 BP 神经网络模型为高,但其本质仍然是一种基于经验风险最小化原则的进化神经网络,为此需利用结构风险最小化原则对其进行进一步改进,以提升其预测精度和泛化能力。

参考文献:

[1] 康重庆,夏清,刘梅.电力系统负荷预测[M].北京:中国电力出版社,2007.

[2] 张思远,何光宇,梅生伟,等.基于相似时间序列检索的超短期负荷预测[J].电网技术,2008,32(12):56-59.

[3] 李明干,孙健利,刘沛.基于卡尔曼滤波的电力系统短期负荷预测[J].继电器,2004,32(4):9-12.

[4] 向峥嵘,王学平.基于小波-神经网络的电力系统短期负荷预测[J].系统仿真学报,2008,20(18):5018-5020.

[5] 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.

[6] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications[J]. Neurocomputing,2006,70(1/3): 489-501.

[7] 邓万宇,郑庆华,陈琳.神经网络极速学习方法[J].计算机学报,2010,33(2):279-287.

(责任编辑:肖锡湘)

(上接第 366 页)

[3] Hochbruck M, Ostermann A. Exponential multistep methods of Adams-type[J]. Numerical Mathematics, 2011, 51(4): 889-908.

[4] Xu Y, Zhao J J, Sui Z N. Exponential Runge-Kutta methods for delay differential equations[J]. Mathematics and Computers in Simulation,2010,80(12):2350-2361.

[5] Hout K J.On the convergence of waveform relaxation methods for stiff nonlinear ordinary differential equations[J].Applied Numerical Mathematics,1995,18(1):175-190.

[6] Zhou S, Huang T Z. Convergence of waveform relaxation methods for Hermitian positive definite linear systems[J]. Applied Mathematics and Computation,2008,203(2):943-952.

[7] 张平文,李铁军.数值分析[M].北京:北京大学出版社,2007.

(责任编辑:陈雯)