

文章编号: 1006-1576 (2008) 11-0059-02

# 基于神经网络回归分析组合模型的能源消耗预测研究

陈玉金<sup>1</sup>, 刘建永<sup>1</sup>, 李凌<sup>1</sup>, 伍中军<sup>1,2</sup>

(1. 解放军理工大学 工程兵工程学院, 江苏 南京 210007; 2. 防化指挥工程学院 指挥1系, 北京 102205)

**摘要:** 通过对时间序列数据进行处理, 分别建立回归分析预测模型和 BP 神经网络预测模型, 在此基础上建立基于 2 种预测方法的组合预测模型。采用熵值法确定组合预测模型的权系数。结合某省能源消耗总量数据进行仿真, 结果与实际数据的误差较小, 和 2 种单一的预测方法相比, 预测结果更接近于实际情况。

**关键词:** 神经网络; 回归分析; 组合预测; 能源消耗

**中图分类号:** TP393.02; O236 **文献标识码:** A

## Forecast and Research on Energy Consumption Based on Analytical and Combined Model of BP Neural Network and Regression

CHEN Yu-jin<sup>1</sup>, LIU Jian-yong<sup>1</sup>, LI Ling<sup>1</sup>, WU Zhong-jun<sup>1,2</sup>

(1. Engineering College of Engineering Corps, PLA University of Science &amp; Technology, Nanjing 210007, China; 2. No. 1 Department of Command, Command &amp; Engineering Institute of Chemical Defense, Beijing 102205, China)

**Abstract:** By processing time sequence data, regression analysis model and BP neural network model were individually built. Based on the two models, a combined forecasting model was established. Entropy value method was applied to ascertain the weights parameters. The energy consumption data of one province was simulated. Its results have little error. Compare with the two single methods, the hybrid measure results is better fit the actual process.

**Keywords:** BP neural network; Regression analysis; Hybrid forecasting; Energy consumption

### 0 引言

能源消耗预测的方法主要有灰色预测法、趋势外推法、主观推断法、指数平滑法。但上述预测方法的误差比较大。从理论上来说, 人工神经网络模型能以任意精度逼近任意非线性序列, 但容易陷入局部收敛, 在建立回归分析模型、神经网络模型的基础上, 建立组合预测模型, 对能源消耗进行预测。

### 1 神经网络回归分析组合预测模型

#### 1.1 回归分析模型

回归分析中最简单的形式是  $y = \beta_0 + \beta_1 x$ ,  $x$ 、 $y$  均为标量,  $\beta_0$ 、 $\beta_1$  为回归系数, 称为一元线性回归。它的一个推广是  $x$  为多元变量, 形如:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m \quad (1)$$

其中  $x = (x_1, \dots, x_m)$ 。这里,  $y$  对回归系数  $\beta = (\beta_0, \beta_1, \dots, \beta_m)$  是线性的, 称为多元线性回归。在回归分析中, 自变量  $x = (x_1, x_2, \dots, x_m)$  是影响因变量  $y$  的主要因素, 是能够获得的, 但  $y$  还受到随机因素的干扰, 假设这种干扰服从零均值的正态分布, 于是模型记作:

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (2)$$

其中  $\sigma$  未知。现有  $n$  个观测数据  $(y_i, x_{i1}, \dots, x_{im})$ ,  $i=1, \dots, n$ ,  $n > m$ , 代入式 (2), 有:

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_m x_{im} + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2), \quad i=1, \dots, n \end{cases} \quad (3)$$

当回归模型建立后, 可由给定的  $x_0 = (x_{01}, \dots, x_{0m})$  来预测  $y_0$  的值, 其预测值为:

$$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_{01} + \dots + \hat{\beta}_m x_{0m} \quad (4)$$

#### 1.2 神经网络预测

神经网络能逼近任何有理函数, 通过训练神经网络可以模拟序列数据之间、序列之间的复杂关系。在众多的人工神经网络模型中, 多层前馈神经网络模型应用最为广泛。用反向传播学习算法 (简称 BP 算法) 可实现多层前馈神经网络的训练。

BP 网络由输入层、隐层、输出层三层构成。输入层、输出层神经元的个数由样本的维数决定。Kawashima 指出隐含层神经元数可按照公式  $2m+1$  个选取 ( $m$  为输入神经元的个数)。

若输入层有  $n_1$  个神经元  $x \in R^{n_1}$ ,  $x = (x_0, x_1, \dots, x_{n_1-1})^T$ ; 中间层有  $n_2$  个神经元,  $x' \in R^{n_2}$ ,  $x' = (x'_0, x'_1, \dots, x'_{n_2-1})^T$ ; 输出层有  $m$  个神经元,  $y \in R^m$ ,  $y = (y_0, y_1, \dots, y_{m-1})^T$ 。输入层与中间层的权值为

收稿日期: 2008-07-08; 修回日期: 2008-07-23

作者简介: 陈玉金 (1984-), 男, 安徽人, 解放军理工大学在读硕士研究生, 从事战场数字化和指挥自动化研究。

$w_{jk}(j=0,1,\dots,n_1-1; k=0,1,\dots,n_2-1)$ , 阈值为  $\theta_k$ ; 中间层与输出层的权值为  $w_{kl}(l=0,1,\dots,m-1)$ , 阈值为  $\theta_l$ 。则各神经元的输出满足:

$$x_k' = f\left(\sum_{j=0}^{n_1-1} w_{jk} x_j - \theta_k\right) \quad (5)$$

$$y_l = f\left(\sum_{k=0}^{n_2-1} w_{kl} x_k' - \theta_l\right) \quad (6)$$

其中  $f$  为 Sigmoid 传递函数:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

### 1.3 神经网络回归分析组合预测模型

组合预测综合利用各种单项预测方法的预测结果, 以适当的加权平均得出组合预测模型。组合预测的核心问题是如何确定加权平均系数。

设采用  $m$  种预测方法对同一问题预测, 第  $i$  种预测方法在第  $t$  时刻的预测值为  $x_{it}$ , 令相对误差为:

$$e_{it} = \begin{cases} 1 & |(x_t - x_{it})/x_{it}| \geq 1 \\ |(x_t - x_{it})/x_{it}| & 0 \leq |(x_t - x_{it})/x_{it}| < 1 \end{cases} \quad (8)$$

熵值法确定组合预测加权系数的步骤如下:

(1) 将各种预测方法的相对误差序列归一化

$$p_{it} = \frac{e_{it}}{\sum_{i=1}^m e_{it}} \quad (9)$$

(2) 计算第  $i$  种预测方法相对误差的熵值

$$h_i = -k \sum_{t=1}^N p_{it} \ln p_{it}, i=1,2,\dots,m \quad (10)$$

其中,  $k=1/\ln N$ 。

(3) 计算第  $i$  种预测方法的预测相对误差序列的变异程度系数

$$d_i = 1 - h_i \quad (11)$$

(4) 计算各种预测方法的权系数

$$l_i = \frac{1}{m-1} \left( 1 - \frac{d_i}{\sum_{i=1}^m d_i} \right) \quad (12)$$

(5) 计算组合预测值

$$\hat{x}_t = \sum_{i=1}^m l_i x_{it} \quad (13)$$

## 2 实例分析

在回归分析模型、神经网络模型的基础上建立组合预测模型并应用于能源消耗总量的预测上, 时间序列资料为: 0.7311, 0.7484, 0.7664, 0.7815, 0.7975, 0.8233, 0.8371, 0.8564, 0.8727, 0.8817, 0.8798, 0.8880, 0.9024, 0.8967, 0.9144, 0.9269, 0.9364, 0.9425。

以连续  $T$  年的资料来预测第  $T+1$  年的值。取  $T=5$ , 得到的数据如表 1。

表 1 样本输入输出数据

序号	输入样本					输出样本
1	0.7311	0.7484	0.7664	0.7815	0.7975	0.8233
2	0.7484	0.7664	0.7815	0.7975	0.8233	0.8371
3	0.7664	0.7815	0.7975	0.8233	0.8371	0.8564
4	0.7815	0.7975	0.8233	0.8371	0.8564	0.8727
5	0.7975	0.8233	0.8371	0.8564	0.8727	0.8817
6	0.8233	0.8371	0.8564	0.8727	0.8817	0.8798
7	0.8371	0.8564	0.8727	0.8817	0.8798	0.8880
8	0.8564	0.8727	0.8817	0.8798	0.8880	0.9024
9	0.8727	0.8817	0.8798	0.8880	0.9024	0.8967
10	0.8817	0.8798	0.8880	0.9024	0.8967	0.9144
11	0.8798	0.8880	0.9024	0.8967	0.9144	0.9269
12	0.8880	0.9024	0.8967	0.9144	0.9269	0.9364
13	0.9024	0.8967	0.9144	0.9269	0.9364	0.9425

将表 1 中的前 8 组数据用于 BP 网络训练, 对后 5 组数据预测, 结果如表 2。同样将表 1 中的前 8 组数据用于回归分析中的回归系数的计算, 对后 5 组数据预测, 结果如表 2。在此基础上, 用建立的组合预测模型计算组合预测结果及相对误差。从表 2 的结果可知: 组合预测模型的预测相对误差比 2 种单一预测方法的相对误差要小, 说明组合预测模型比单一预测方法能更好地对时间序列数据预测。

表 2 预测结果对比

序号	实际值	神经网络预测结果	相对误差	回归预测结果	相对误差	组合预测结果	相对误差
9	0.8967	0.8907	0.67%	0.9055	0.99%	0.9010	0.48%
10	0.9144	0.9238	1.03%	0.8980	1.80%	0.9059	0.93%
11	0.9269	0.9377	1.17%	0.9179	0.97%	0.9240	0.31%
12	0.9364	0.9457	0.99%	0.9222	1.52%	0.9294	0.74%
13	0.9425	0.9489	0.68%	0.9305	1.27%	0.9329	0.67%

## 3 结论

该模型用熵值法确定了组合预测模型的权系数, 并应用于时间序列预测上, 得到了较好的预测效果。应用实例表明, 组合预测模型比单一的预测方法的预测结果好, 且能适用于各种时间序列资料的预测问题, 有普遍的适应性。

## 参考文献:

- [1] 郭万敏, 赵磊. BP 组合预测方法在维修保障费用预测中的应用[J]. 火力指挥与控制, 2004, 29(3).
- [2] 秦大建, 李志蜀. 基于神经网络的时间序列组合预测模型研究及应用[J]. 计算机应用, 2006, 26(6).
- [3] 钟颖, 汪秉文. 基于遗传算法的 BP 神经网络时间序列预测模型[J]. 系统工程与电子技术, 2002, 24(4).
- [4] 姜晨, 徐宗昌, 肖国军. 用神经网络组合预测法估算反舰导弹研制费用[J]. 系统工程与电子技术, 2004, 26(3).
- [5] 张利萍, 李宏光. 灰色神经网络预测算法在 DMF 回收过程中的应用[J]. 微计算机信息, 2005, 21(1).
- [6] 陈华友. 组合预测方法有效性理论及其应用[M]. 北京: 科学出版社, 2008.
- [7] 李凌, 刘建永, 陈玉金. 基于遗传神经网络的伪装效能评估模型[J]. 兵工自动化, 2007, 26(8): 3-4.