Market Demand Prediction Model Based on Momentum Factor Optimization BP Algorithm

Qiong Sun, Lin Li
Beijing Union University, Beijing 100101, Beijing, China

Abstract
The standard BP algorithm has low precision when being applied in the market demand forecasting, this paper put forward an optimized BP algorithm based on momentum factor. Firstly, in order to accelerate the convergence and avoid oscillation, it introduces a momentum factor. Through adopting new performance index function, the network size is reduced automatically. Then different learning rates are used to dynamically adjust the connection weight among different nodes to improve the adjustment efficiency of weight and the convergence speed of error function. The simulation experiment shows that the proposed algorithm has higher precision in the market demand forecasting.

Key words: Momentum Factor, Oscillation Optimization, Market, Demand Forecasting, Differentiation Learning Rate.

1. INTRODUCTION
Tourism demand forecasting plays an important role in the tourism planning. The common tourism demand forecasting method is a mathematical model based on statistics. However, due to the short tourism demand statistics time and the interference from many unpredictable factors, traditional methods have a lot of human factors with big error and poor fault tolerance, which is difficult to meet current requirement (Gao et al., 2011). Artificial neural network is an information processing system established as mimic biological neurons with the advantages like non-linear approximation, parallel processing, self-learning, self-organization and fault tolerance. It has certain advantages in the practical prediction application (Erdal et al., 2010).

With the rapid growth of global tourism demand, tourism demand forecasting has become the hot topic at home and abroad. Law took neural network model to predict the amount of Japanese to Hong Kong and compared the neural network prediction model with multiple regression model, the naive model, moving average model and exponential smoothing model (Law and Au, 2009). The research shows that neural network prediction is more accurate than the other four kinds of forecasting model. De Mello used vector auto-regression VA model to analyze the relation between and demand, and predicted the amount of British tourist (De and Nell, 2011). The author draws the conclusion that vector auto regression prediction model is effective and accurate to predict the visitors to Britain. Lim and McAleer compared several exponential smoothing models and used them to predict the number of Hong Kong, Malaysia tourists and Singapore tourists in Australia (Lim and McAleer, 2012). The difference between the fitting value and actual value was used to test the accuracy of prediction. The results show that Holt-Winter additive model and multi seasonal prediction model have the highest accuracy. Wang had attempted to use gray and Markov prediction model for Taiwan inbound tourist, and found that traditional grey forecasting model GM (1, 1) model was appropriate for Hong Kong and the United States while Grey-Markov model was especially appropriate for Germany where historical data sequence of inbound tourists had a large change. Oh and Morzuch showed that the combination prediction model of four kinds of time series was more accurate than the single forecast model (Oh and Morzuch, 2014). Wong et al. used advanced combination forecasting model to study the inbound tourists in Hong Kong and came to the same conclusion that combination forecasting model was more accurate than the single forecast model, which can reduce the risk of failure (Wong et al., 2007). Zhu Xiaohua and Yang Xiuchun (Zhu and Yang, 2014) explored a variety of tourist forecasting models, and compared their prediction accuracy. Taking the inbound tourism in China from 1978 to 2001 as an example, the quantitative analysis of linear regression model, moving average forecasting model, exponential smoothing model and grey prediction model were adopted to forecast China inbound tourist. When the original series are relatively short, the grey prediction model prediction has relatively small error, and when the original data sequence is long, the exponential smoothing prediction model is preferred. The linear regression analysis and prediction model is not applicable in this case. Zhang Ming thought the essential to establish the measurement model for tourism demand forecasting was the selection of variables and model form in the analysis of measurement model of tourism demand (Zhang, 2012).
In view of the problems of standard BP algorithm in the market demand prediction, this paper put forward an optimized BP algorithm based on momentum factor. The simulation experiment was conducted to prove the effectiveness of the proposed model.

2. APPLICATION OF BP NEURON NETWORK IN TOURIST MARKET REQUIREMENT PREDICTION

BP neuron network consists of input layer, hidden layer and output layer. Suppose discrete-time series with $N$ sample set $\{(x(t), y(t))|x \in R^m, y \in R^n, t = 1, 2, \cdots, N\}$. The BP network can achieve a highly non-linear mapping from input to output. Namely, a mapping relationship $F: R^m \rightarrow R^n$ is obtained to divide all the samples into training samples $\phi_1$ and test samples $\phi_2$.

$$\phi_1 = \{(x(t), y(t))|x \in R^m, y \in R^n, t = 1, 2, \cdots, N_1, N_1 \leq N\} \quad (1)$$

$$\phi_2 = \{(x(t), y(t))|x \in R^m, y \in R^n, t = N_1 + 1, N_1 + 2, \cdots, N\} \quad (2)$$

The mapping relationship is firstly established by $\phi_1$ to check the applicability in $\phi_2$ from input-output of network. If the input-output is right, this model can be applied in the practical forecasting. It can be realized by a three layer BP neuron network with $m$ input nodes, $n$ output nodes, and $p$ hidden nodes. The relationship between network input and output is written below.

$$\hat{y}_i(t) = \sum_{j=1}^{p} v_{ji} \cdot f \left[ \sum_{k=1}^{m} w_{jk} \cdot x_i(t) + \theta_i \right] + \epsilon_i$$

where $f(x) = \frac{1}{1 + e^{-x}}, k = 1, 2, \cdots, n, t = 1, 2, \cdots, N_1, x_i$ is the input of the network; $\hat{y}_i$ is the output layer; $w_{jk}$ is the weight from input layer $j$ node to output layer $k$ node; $\theta_i$ is the threshold of $j$ in hidden layer; $\epsilon_i$ is the threshold of $k$ node; $f$ is the activation function. If the total error of the network is less than $\epsilon_1$, then,

$$E_1 = \frac{1}{2} \sum_{j=1}^{n} \sum_{t=1}^{N_1} [y_i(t) - \hat{y}_i(t)]^2 \leq \epsilon_1$$

If the average error of the sample is less than $\epsilon_2$, then

$$E_2 = \frac{1}{N - N_1} \sum_{j=1}^{N_1} \sum_{t=1}^{N} [y_i(t) - \hat{y}_i(t)]^2 \leq \epsilon_2$$

As a complex abstract system, tourism demand system is constrained by many factors. This paper determines the main system behavior sequence of the tourism demand is the tourist number time data sequence, denoted as $X_0$. The related influence factors include mainly four parts: taking per capita disposable income as discretionary income, noted as $X_1$; taking travel service person as the service condition, noted as $X_2$; taking domestic general tourism income as the environment and landscape development condition, noted as $X_3$; taking the national highway traffic mileage as the traffic condition, noted as $X_4$. Each index data is from the 2014 China Statistical Yearbook, as shown in table 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>$X_0$</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>11.84</td>
<td>10859.6</td>
<td>220.65</td>
<td>4722.4</td>
<td>186.80</td>
</tr>
<tr>
<td>2005</td>
<td>12.78</td>
<td>11702.8</td>
<td>229.58</td>
<td>5078.4</td>
<td>196.52</td>
</tr>
<tr>
<td>2006</td>
<td>12.70</td>
<td>12472.2</td>
<td>236.05</td>
<td>5642.3</td>
<td>200.98</td>
</tr>
<tr>
<td>2007</td>
<td>15.02</td>
<td>13421.6</td>
<td>247.25</td>
<td>5910.7</td>
<td>207.07</td>
</tr>
<tr>
<td>2008</td>
<td>16.12</td>
<td>14493.0</td>
<td>254.39</td>
<td>6485.8</td>
<td>354.52</td>
</tr>
<tr>
<td>2009</td>
<td>17.34</td>
<td>15759.5</td>
<td>261.43</td>
<td>7429.7</td>
<td>365.70</td>
</tr>
<tr>
<td>2010</td>
<td>20.10</td>
<td>17785.8</td>
<td>264.04</td>
<td>8970.6</td>
<td>378.37</td>
</tr>
<tr>
<td>2011</td>
<td>21.12</td>
<td>19780.8</td>
<td>270.87</td>
<td>9949.3</td>
<td>393.02</td>
</tr>
<tr>
<td>2012</td>
<td>23.02</td>
<td>21174.7</td>
<td>278.57</td>
<td>11383.6</td>
<td>406.08</td>
</tr>
<tr>
<td>2013</td>
<td>25.03</td>
<td>23109.4</td>
<td>283.32</td>
<td>13779.8</td>
<td>420.82</td>
</tr>
</tbody>
</table>

Four influence factors of tourism demand are input into the input layer to establish the multivariable BP neuron network. The three layer BP neuron network has four input nodes and one output node, and the node of
hidden layer is calculated empirically to be 9. After 4295 steps, the network training finished, with the performance curve shown in Figure 1.

![BP network training performance curve](image)

**Figure 1.** BP network training performance curve

The trained network is used to predict the sample. The simulated and predicted structure is listed in Table 2~3.

<table>
<thead>
<tr>
<th>Year</th>
<th>The actual value</th>
<th>Fitted values</th>
<th>Relative error/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>11.84</td>
<td>11.78</td>
<td>2.45</td>
</tr>
<tr>
<td>2005</td>
<td>12.78</td>
<td>12.27</td>
<td>1.11</td>
</tr>
<tr>
<td>2006</td>
<td>12.70</td>
<td>12.49</td>
<td>0.67</td>
</tr>
<tr>
<td>2007</td>
<td>15.02</td>
<td>15.02</td>
<td>2.30</td>
</tr>
<tr>
<td>2008</td>
<td>16.12</td>
<td>16.69</td>
<td>1.03</td>
</tr>
<tr>
<td>2009</td>
<td>17.94</td>
<td>17.19</td>
<td>6.51</td>
</tr>
<tr>
<td>2010</td>
<td>20.10</td>
<td>20.23</td>
<td>1.19</td>
</tr>
</tbody>
</table>

Fit the sample mean relative error(%): 1.813

<table>
<thead>
<tr>
<th>Year</th>
<th>The actual value</th>
<th>Fitted values</th>
<th>Relative error/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>21.12</td>
<td>21.12</td>
<td>0.99</td>
</tr>
<tr>
<td>2012</td>
<td>23.02</td>
<td>23.08</td>
<td>2.12</td>
</tr>
<tr>
<td>2013</td>
<td>25.03</td>
<td>26.13</td>
<td>3.97</td>
</tr>
</tbody>
</table>

Sample mean relative prediction error(%): 2.886

The multivariable BP neuron network model is adopted to predict the tourists number with the average relative error 2.886%. Theoretically, the number of influence factor is the number of neuron of input layer. However, if the network input is too much, the weight between input layer and hidden layer will increase exponentially.

### 3. BP ALGORITHM BASED ON MOMENTUM FACTOR OPTIMIZATION

#### 3.1. Oscillation Optimization Based on Momentum Factor

In order to accelerate the convergence and avoid the oscillation, this paper introduces a momentum factor \( \alpha \):

\[
w(n_t + 1) = w(n_t) + \eta(n_t) d(n_t) + \alpha \Delta w(n_t)
\]

(6)

\( \alpha \) is the momentum item, usually as a constant.

The third item is the weight adjustment of last time and the adjustment direction in time \( (n_t) \) is the combination of direction in \( (n_t - 1) \) and \( (n_t) \).
\[ w(n_0 + 1) = w(n_0) + \eta(n_0) \left[ d(n_0) + \frac{\alpha}{\eta(n_0)} \Delta w(n_0) \right] \]
\[ = w(n_0) + \eta(n_0) \left[ d(n_0) + \frac{\alpha\eta(n_0) - 1}{\eta(n_0)} d(n_0 - 1) \right] \] (7)

The above expression is similar to the conjugate gradient methods, but the \( d(n_0 - 1) \) and \( d(n_0) \) is not conjugate, and \( 0 < \alpha < 1 \). Therefore, it is suggested when \( \eta(n_0) \) is adjusted, \( \Delta E > 0 \) and \( \eta \) is required to decrease, \( \alpha = 0 \); when \( \eta \) increases, \( \alpha \) returns to 0.

The first step of regularization is to modify the error function. The original error function is the sum of square of network to training mode, namely,

\[ F = mse = \frac{1}{N} \sum_{i = 1}^{N} (e_i)^2 = \frac{1}{N} \sum_{i = 1}^{N} (t_i - a_i)^2 \] (8)

Here, \( N \) represents the training mode number, \( t_i \) the expected network output, and \( a_i \) the actual network output. The error function is modified as,

\[ msereg = \gamma \cdot mse + (1 - \gamma)msw \] (9)

\( \gamma \) is the error coefficient, and \( msw \) is the average value of the sum of square of all the network weight, namely

\[ msw = \frac{1}{N} \sum_{i = 1}^{N} w_i^2 \] (10)

It is seen that through the new performance index function, the network training error is kept as small as possible to ensure the small weight of network. That is to say, the effective weight will be much small and the size of the network automatically reduces.

### 3.2 Weight Adjustment Based on Different Learning Rate

This paper adopts the different learning rate to dynamically adjust the connection weight between different nodes so as to improve the adjustment efficiency of weight and the convergence speed of error function. The realization step of the improved BP neuron network is described as follows.

![Figure 2. The propagation mechanism of improved algorithm](image)

The threshold of BP neuron network can be adjusted with the weight and added to the weight matrix. Suppose \( U \) is the weight matrix from input layer to hidden layer, and \( V \) is the weight matrix from hidden layer to output layer. \( X \) is the input vector of the network, \( Y \) the output vector, \( M \) sample amount, \( f (\cdot) \) the activation function of BP neuron network, and \( d_i \) the expected output of network. Then,

\[ X = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1l} \\ 1 & x_{21} & x_{22} & \ldots & x_{2l} \\ 1 & x_{31} & x_{32} & \ldots & x_{3l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{M1} & x_{M2} & \ldots & x_{ML} \end{bmatrix} = (X_m)_{M \times (J + 1)} \] (11)
By using the adaptive learning rate algorithm to improve the BP neuron network, the forward propagation steps remain basically same.

\[ f(x) = (1 + e^{-ax})^{-1} \]

(14)

\[ Y_j = f(\sum_{i=1}^{L} x_{wi} u_i) \]

(15)

\[ Y_p = f(\sum_{j=1}^{P} Y_j v_{jp}) \]

(16)

\[ E(n) = \frac{1}{2} \sum_{p=1}^{P} (d_p - Y_p)^2 \]

(17)

The \( W(n) \) is the weight matrix of BP neuron network after \( n \) times learning. \( W'(n) \) is the inverted sequence weight matrix of \( W(n) \). Then the initial weight matrix and initial inverted sequence weight matrix of the network can be represented as \( W(0) \) and \( W'(0) \). When the weight \( W' \) is adjusted, \( w_i(n) \) is denoted as the value in the \( n \) th learning of \( w_i \) and \( \eta_i(n) \) is denoted as the learning rate of \( w_i \) in the \( n \) th learning.

If \( W(n) \) is in the forward propagation stage, it can meet the learning ending condition.

\[ E(n) \leq e \]

(18)

Here, \( E(n) \) is the error of network in the \( n \) th learning; \( e \) is the preset precision; \( W(n) \) is the optimum weight matrix.

If the learning condition is not satisfied, BP neuron network steps into the invert adjustment stage. In order to clarify this step of the new algorithm, taking any weight \( w_i \) in inverted weight matrix \( W' \) for example, the adjustment method is followed.

If the gradient of \( w_i \) is equal to 0, namely \( \frac{\partial E}{\partial w_i} = 0 \), then next weight \( w_{i+1} \) is adjusted with same method in the adjustment of \( w_i \).

If the gradient of \( w_i \) is not equal to 0, namely \( \frac{\partial E}{\partial w_i} \neq 0 \), then the weight should be recalculated in the \( n \) th learning following the equation (19) and (20).

\[ \Delta w_i(n) = -\eta_i(n) \frac{\partial E}{\partial w_i} \]

(19)

\[ w_i(n + 1) = w_i(n) + \Delta w_i(n) \]

(20)

If \( w_i(n + 1) \) is calculated by forward propagating, and better than \( w_i(n) \), then,

\[ \eta_i(n) \times 2 \Rightarrow \eta_i(n) \]

(21)

Following the negative gradient direction of \( w_i(n) \), the weight \( w_i(n + 1) \) is calculated again with the new learning rate \( \eta_i(n) \). Firstly, a new weight \( w_i^{(i)}(n + 1) \) is obtained. If the error of \( w_i^{(i)}(n + 1) \) calculated by forward propagating is less than \( w_i(n + 1) \), then the learning rate increases continuously with equation (21). Then the
weight $w_i^{(2)}(n+1)$ is calculated by negative propagating of new learning rate. If the $w_i^{(2)}(n+1)$ is still prior to $w_i^{(3)}(n+1)$, the learning rate increases. If the weight $w_i^{(m)}(n+1)$ after $m+1$ times calculation is worse than $w_i^{(m-1)}(n+1)$, then

$$w_i(n+1) = w_i^{(m-1)}(n+1)$$

(22)

$$\eta(n) = \eta_i^{(m-1)}(n)$$

(23)

If the $w_i(n+1)$ is calculated by forward propagating, and worse than $w_i(n)$, namely larger error, then

$$\eta(n)/2 = \eta_i(n)$$

(24)

Following the negative gradient direction of $w_i(n)$, $w_i(n+1)$ is calculated again with the new learning rate $\eta_i(n)$. If $w_i(n+1)$ is not better than $w_i(n)$, then the learning rate $\eta_i(n)$ decreases by equation (24). Only if the corresponding gradient of $w_i(n)$ is not equal to 0, this calculation continuously. There must be a $w_i(n+1)$ different from but better than $w_i(n)$.

By this way, the connection weights among all the nodes are adjusted. This is called as a complete learning. A new weight matrix $W(n+1)$ and inverted sequence weight matrix $W^*(n+1)$ are obtained. To judge whether the iteration ending condition is satisfied, if it is, $W(n+1)$ is the optimal weight of BP neuron network. If not, starting from $W^*(n+1)$, the calculation steps are repeated until the condition is satisfied.

4. SIMULATION EXPERIMENT

In order to test the performance of the proposed algorithm, this paper conducted the simulation experiment. Firstly, following test function is adopted to test the convergence performance of the two algorithms.

$$f(x) = 0.2 + 0.8 \cdot |x + 0.5 \cdot \sin(2\pi x)|$$

(25)

The error difference of standard BP algorithm and improved BP algorithm in test function is shown in Figure 3.

![Figure 3. Error variance of improved BP algorithm in test function](image)

From figure 3, it is seen that the proposed BP algorithm has far less error than standard BP algorithm. Then according to the test data in table 1, the tourism market demand is simulated and predicted, as shown below.

<table>
<thead>
<tr>
<th>Year</th>
<th>The actual value</th>
<th>Fitted values</th>
<th>Relative error/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>11.84</td>
<td>11.81</td>
<td>0.25</td>
</tr>
<tr>
<td>2005</td>
<td>12.78</td>
<td>12.74</td>
<td>0.15</td>
</tr>
<tr>
<td>2006</td>
<td>12.70</td>
<td>12.63</td>
<td>0.55</td>
</tr>
<tr>
<td>2007</td>
<td>15.02</td>
<td>15.02</td>
<td>0.00</td>
</tr>
<tr>
<td>2008</td>
<td>16.12</td>
<td>16.11</td>
<td>0.06</td>
</tr>
<tr>
<td>2009</td>
<td>17.94</td>
<td>17.89</td>
<td>0.28</td>
</tr>
<tr>
<td>2010</td>
<td>20.10</td>
<td>20.12</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Fit the sample mean relative error(%): 0.019
Table 5. The prediction result of improved algorithm

<table>
<thead>
<tr>
<th>Year</th>
<th>The actual value</th>
<th>Fitted values</th>
<th>Relative error/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>21.12</td>
<td>21.12</td>
<td>0.00</td>
</tr>
<tr>
<td>2012</td>
<td>23.02</td>
<td>23.01</td>
<td>0.04</td>
</tr>
<tr>
<td>2013</td>
<td>25.03</td>
<td>25.02</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Sample mean relative prediction error(%): 0.023

Figure 4. The fitting results comparison of two algorithms

Figure 5. The prediction results comparison of two algorithms

The above results show that the improved algorithm has higher precision and low error of 0.023% in the tourism market demand prediction.

5. CONCLUSIONS

The in-depth study of domestic tourism demand and accurate analysis and prediction of current demand and future changes in the domestic tourism market, is quite important not only for the development and utilization of the tourism resources and environment construction, but also for the government to expand domestic demand and formulate relevant supporting policies so as to ensure the rapid development of national economy. In view of the problems of standard BP algorithm in the tourism market demand prediction, this paper put forward an optimized BP algorithm based on momentum factor. The simulation experiment proved the higher precision of the proposed model in practical application.

Acknowledgements

This study was supported by a grant from the the Science and Technology Project of Beijing Municipal Education Commission (to Sun Qiong, No. KM201511417009)

REFERENCES


