Study of BP neural network based on improved particle swarm algorithm

Xiaoqian Ma and Liyuan Li*
School of Data Science, Anhui University of Finance and Economics, Bengbu, China
*Corresponding author e-mail: 20181883@aufe.edu.cn

Abstract. This paper uses first-order difference to transform non-smooth data into smooth time series data, determines the p and q parameters in the model by judging the trailing and truncated nature of ACF, PACF, and finally establishes the ARIMA model after ACI, BCI detection. According to the parameters of the neural network randomly selected similar to the initial spatial position of the particles in the particle swarm algorithm, the improved particle swarm algorithm is used instead of the gradient correction method to precisely adjust the parameters and establish the BP neural network, which improves the robustness and accuracy of the prediction model.

Key words: ARIMA models, particle swarm optimization, neural networks.

1. Forecasting changes in regional economic dynamics based on ARIMA models

1.1. Research ideas
The long-term impact of economic transformation policies on the region's economic dynamism during the period 1990-2017 is studied using the city of Huaibei in Anhui Province as an example. Short-term impacts are predicted by identifying the city's gross domestic product and the value of production of primary, secondary and tertiary industries before and during the economic transformation period for comparison, and an ARIMA model is developed to predict the long-term impacts over the next two decades following the economic transformation policy using time series analysis.

1.2. Research methodology
(1) Differentiation: Difference operations are performed on non-stationary time series to turn them into a stationary time series.
First order differential: \[ \nabla X_t = X_t - X_{t-1} \]
Second order differential: \[ \nabla^d x_t = \nabla (\nabla^{d-1} x_t) = \sum_{i=0}^{d} (-1)^i C_d^i x_{t-i} \]
(2) ARIMA model
Order p autoregressive model AR (p):
\[ Z_t = \phi_1 z_1 + \phi_2 z_2 + \cdots + \phi_p z_{t-p} + \varepsilon_t; \]
\[ Z_t = y_t - \mu \]
Q order sliding average model MA (q):

\[ Z_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_t \]

The ARMA (p, q) model containing q autoregressive items and m moving average items is as follows:

\[ x_t = \phi_1 z_t + \phi_2 z_{t-1} + \cdots + \phi_p z_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_t + \epsilon_t \]

ARIMA (p, d, q) model can be expressed as:

\[
\left(1 - \sum_{i=1}^{p} \phi_i L^i \right) \left(1 - L \right)^d x_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i \right) \epsilon_t
\]

(3) model recognition and order

ACF: autocorrelation coefficient

\[ ACF(k) = \frac{\text{Cov}(y_t, y_{t-k})}{\text{Var}(y_t)} \]

PACF: partial autocorrelation coefficient

\[ \rho_{X_t X_{t-k} \mid X_{t-1}, \ldots, X_{t-k+1}} = \frac{E[(X_t - \bar{X}_t)(X_{t-k} - E\bar{X}_{t-k})]}{E[(X_{t-k} - \bar{X}_{t-k})^2]} \]

1.3. Analysis of results

(1). The differential processing of the data

This paper studies the time series of the gross product of Huaibei City and the gross product of primary, secondary and tertiary industries from 1990-2017 are studied, and the data are obtained from the China Statistical Yearbook. From Figure 1, it can be seen that the data are non-stationary series, while t-test and p-test are used to determine the stability of the series. The first-order difference of the data is performed, and Figure 2 shows the data of the first-order difference, and it is known that the data of the first-order difference is smooth according to the statistical test.

(2). Model recognition and order

The following figure gives the autocorrelation and partial autocorrelation plots of this first-order difference series, which can be fitted by building an ARIMA model to estimate p, q according to the broken tails of ACF, PCF. For the data of primary industry, the ACF plot (Fig. 3) breaks the tail at q=3, and the PACF plot (Fig. 4) truncates the tail at p=0, so ARMA (0,3) is chosen; for the data of secondary industry, the ACF plot (Fig. 5) is broken at q=2 and the PACF plot (Fig. 6) is truncated at p=1, so ARMA (1,2) is chosen; for the data of the tertiary industry, the ACF plot (Fig. 7) is broken at q=2 and the PACF plot (Fig. 8) is truncated at p=2, so ARMA (2,2) is chosen.
The data were tested by ACI, BCI, ARIMA(0,1,4), ARIMA(2,1,4), and ARIMA(2,1,4) fit the best.

(3). model prediction

The sequence prediction results of the next 20 years through ARIMA are shown in the figure below:
2. Study of economic dynamics based on neural networks

2.1. Research ideas
By quantitatively analyzing regional economic vitality, which was originally an illusory concept, into concrete values, 19 cities are ranked in terms of regional economic vitality. After layers of screening, regional economic vitality can be evaluated in 6 aspects, and 2 indicators are screened out for each aspect. A complete regional economic vitality system is established with these 12 indicators, and a neural network model is applied to derive the optimal solution among the 19 cities within 2015 to 2017, and finally the regional economic vitality ranking of the cities is obtained according to the optimal solution.

2.2. Research Process
(1). Identifying indicators
First, six aspects of regional economic vitality were measured and analyzed to start with, and two indicators were picked as representatives of each aspect. The final established regional economic vitality indicator system is shown in the following table.

| Level indicators             | The secondary indicators                                      |
|------------------------------|----------------------------------------------------------------|
| Comprehensive economic strength | GDP per capita            |
|                               | Average wage             |
| Financial strength           | Per capita disposable income |
|                               | Total fixed asset investment |
| openness                     | Total imports             |
|                              | Total exports             |
| Talent and technology level  | Science and technology activity personnel |
|                              | Local government expenditure on science and technology |
| The management level         | Revenue from the general public budget |
| Infrastructure and housing   | General public budget expenditures |
|                              | Afforestation rate        |
|                              | Total investment in infrastructure |
(2). Establishment of neural network model based on particle swarm algorithm optimization

① The initial data is normalized, and then the data is accumulated. The data is normalized according to the formula: 
\[ x'_i = \frac{x_i - x_{i\min}}{x_{i\max} - x_{i\min}} \]

② Network initialization, the initial setting of each coefficient: the acceleration coefficient c1 and c2 of the particle swarm optimization algorithm are both 0.8, the maximum number of cycles maxgen is 800, the population size sizepop is 39, the weight factor is reduced from 0.9 to 0.4, and the upper and lower limits of the individual range are respectively 100 and 0.02.

③ Determine the gray neural network structure according to the input and output sample pairs, determine the network parameters to be optimized according to the network structure, and the number of network parameters is the population length. a b.

④ Population initialization randomly generated sizepop initial population, by calculating the fitness value. \(X_i, f_{it}, f_{it}\). The fitness USES the average absolute error of the output of the gray god through the network. The average absolute error function is:
\[ f_a = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{m} |y_{ik} - t_{ik}| \]

In the above equation: the predicted value of the test set; \(y_{ik}, t_{ik}\) is the true value of the test set; \(n\) is the number of test sets; \(m\) is the number of output contacts. According to the calculated fitness value, find the individual corresponding to the minimum fitness value. \(f_{it,\min} X_{i\min}\) To:
\[ ZX_{i\min} = X_{i\min}, Zf_{it,\min} = f_{it,\min} \]

\(ZX_{i\min}\) is the ultimate optimal individual and is the ultimate optimal fitness value \(Zf_{it,\min}\).

⑤ Start the cycle, update the position and velocity of particles according to equations \(X_i\) and \(V_i\), and then initialize the population \(X_i\) to the updated \(X'_i\) population with a certain probability. Its fitness value \(f'_{it}\) is calculated from the updated \(X'_i\).

⑥ Compare \(f_{it}\) and \(f'_{it}\) size, if \(f_{it} > f'_{it}\), then \(X_i = X'_i\), to preserve the excellent individual.

⑦ \(X_i\) after the steps ④⑤ is the new generation of individuals. Find the smallest \(f_{it,\min}\), and its corresponding \(X_{i\min}\). If \(f_{it,\min} < Zf_{it,\min}\), then \(ZX_{i\min} = X_{i\min}\), otherwise no operation.

⑧ Loop steps ④ to ⑥, end the loop when the maximum number of loop iterations is reached, and output the calculation result

⑨ Assign the parameters of the gray neural network to, and carry out the next step of calculation and result output. \(ZX_{i\min} a b\)

### Table 2. Regional Economic Vitality Ranking

| City      | Score   | Ranking | City      | Score   | Ranking |
|-----------|---------|---------|-----------|---------|---------|
| Shanghai  | 1.174629734 | 1       | Chongqing | 0.392448598 | 11      |
| Shenzhen  | 0.796603289  | 2       | Ningbo    | 0.34195382  | 12      |
| Beijing   | 0.775780164  | 3       | Kunming   | 0.300374678 | 13      |
| Guangzhou | 0.596547381  | 4       | Nanjing   | 0.300374678 | 14      |
| Shenyang  | 0.567440463  | 5       | Changsha  | 0.261439406 | 15      |
| Qingdao   | 0.52559593   | 6       | Dongguan  | 0.259814989 | 16      |
| Tianjin   | 0.451387483  | 7       | Wuhan     | 0.236210664 | 17      |
| Suzhou    | 0.432112122  | 8       | Chengdu   | 0.222913027 | 18      |
| Hangzhou  | 0.423652637  | 9       | Zhengzhou | 0.222384234 | 19      |
3. Conclusion
For the prediction of the economic vitality of the target region, HuaiBei City, Anhui Province was taken as an example, and the selected non-stationary data was transformed into stationary data by differencing, and then an ARIMA model was built to fit. The error is roughly linearly distributed along the prediction series, and the residuals follow a Gaussian distribution in theory, with 95% of the numbers falling within the confidence interval, which shows that the model fits well and the error is small.

For the city economic vitality ranking, this paper uses the 12 selected indicators to combine the gradient descent method and neural network by improving the particle swarm algorithm, adjusting the weights of the network, and using the mean square error of the grey neural network output as the particle fitness value to establish the regional economic evaluation system, the established model has higher prediction accuracy and faster convergence speed, improving the convergence speed of traditional algorithms. The model established has higher prediction accuracy and faster convergence, which improves the convergence speed of traditional algorithms and avoids falling into local optimum. The final ranking of cities in terms of regional economic vitality is obtained based on the optimal solution of the neural network model optimised by the particle swarm algorithm.

4. References
[1] Yi wei. Construction of evaluation index system for regional economic vitality in sichuan province [A]. Journal of sichuan vocational and technical college. 2015.
[2] Ma junjie, you jianxin, Chen zhen. Grey neural network model based on improved particle swarm optimization algorithm [A]. Journal of tongji university. 2012.
[3] Zhao jiwu, zou changwu. Application of neural network based on particle swarm optimization algorithm in water resources evaluation [A]. Journal of chengdu university of information engineering. 2010.
[4] he ruqun. Evaluation of urban economic vitality in the pearl river - xijiang economic belt [D]. 2019.
[5] precipitation literature in the collection of papers: gu cuiling, wang ning. Application of ARIMA model in population sequence prediction [A]. 2016.
[6] precipitation literature in the collection of papers: an xin, jia jin-zhang. ARIMA prediction model for time series of mine water inflow [A]. 2015.