The Application of Modern Optimization Algorithm in Time Series Prediction

Zhaoyue Hu[1, a *],Yanping Bai[2,b]

[1] Department of Mathematics North University of China, Taiyuan, Shanxi CHINA

[2] Department of Mathematics North University of China, Taiyuan, Shanxi CHINA

[a]huzhaoyuee@163.com, [b]baiyp666@163.com

Keywords: Stochastic chaos (SC), Principal components analysis (PCA), Stochastic volatility jump diffusion (SVJD), Artificial neural network (ANN), Genetic algorithm (GA).

Abstract. This paper applies the SC model and SVJD model to artificially generated data, and we put forward a forecast model that hybridizes genetic algorithm, principal component analysis and artificial neural network methods. This article utilizing genetic algorithm to search for the initial weights of the BP neural network could guarantee a relatively high probability to obtain the global optima, and we include principle component analysis (PCA) to extract contribution rate to meet 85% of the principal component as the input variables. The experiment results demonstrate that the combination methods PCA-BP and PCA-GA-BP model is adopted to overcome the fitting compared with the traditional forecasting method.

Introduction

Artificial neural network (ANN) has been popular in time series prediction, because of their advantages in handling nonlinear systems. As a large of scale parallel nonlinear information processing systems and neural network possesses excellent fault tolerance, robustness, self-learning, self-organization, and adaptability properties [1]. In recent years many scholars put forward kinds of the digital simulation prediction models based on neural network [2,3]. But the researchers are not aware of the impact of their data selection on the results, their test sets are small, and the selection of the data sets is not supported by any previous data analysis. The primary motive for using artificial data is that there are no limits to the size of the sample. We can estimate the parameters from a training set with sufficiently large degrees of freedom, and then forecast with a relatively ample test set.

BP neural network

In 1986, Rumelhart, Hinton and Williams proposed error back propagation algorithm applied to feed forward neural network learning and training, it is composed of input layer, hidden layer and output layer. The learning process of BP network is made up of two parts: forward propagation of information and back propagation of error. During forward propagation, information is transferred from input layer to output layer after processing in the hidden layer, and the neural state of each layer only affects the neural state of the next layer [4]. If the expected output is not obtained at the output layer, back propagation will be initiated. The error signal will be transported back along the original neuronal pathway through the network, and weights of neurons at different layers will be modified. Iteration happens to this process continuously till the expected target is reached.

GA-BP algorithm

Conventional BP neural networks suffer from a number of drawbacks, such as the low convergence rate, sensitivity to weight initialization and high probability of being entrapped in local optimum. Genetic algorithm is a parallel random search optimization method proposed by Professor Holland from University of Michigan in 1962 to simulate genetic mechanism and biological evolution in natural world. Its basic thought is to solve complicated optimization problems by simulating

biological and human evolution [5]. Genetic algorithm are a global optimization algorithm, being capable of finding the globally optimal solution in complex networks, utilizing genetic algorithm to search for the initial weights of the BP neural network could guarantee a relatively high probability to obtain the global optima.

Principal component analysis

Principal component analysis was introduced by Pearson in 1901 and further developed by Hotelling in 1933. The main idea of principal component analysis is to recombine numerous indexes with correlation into a new group of irrelevant comprehensive indexes which will replace the original indexes [6]. Its mathematical model is as follows:

For sample data, the information matrix of observing p variables and $x_1, x_2, \mathbf{L}\ x_p$, samples is:

$$X = \begin{pmatrix} x_{11} & x_{12} & \mathbf{L} & x_{1p} \\ x_{21} & x_{22} & \mathbf{L} & x_{2p} \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ x_{n1} & x_{n2} & \mathbf{L} & x_{np} \end{pmatrix} = \left(x_1, x_2, \mathbf{L}\ x_p \right), \ x_j = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \mathbf{M} \\ x_{nj} \end{pmatrix}, j = 1, 2, \mathbf{L}, p \tag{1}$$

Principal component analysis is integrated p observation variable to become a new variable p,

$$\begin{cases} F_1 = a_{11}x_1 + a_{12}x_2 + \mathbf{L}\ a_{1p}x_p \\ F_2 = a_{21}x_1 + a_{22}x_2 + \mathbf{L}\ a_{2p}x_p \\ \mathbf{LL} \\ F_p = a_{p1}x_1 + a_{p2}x_2 + \mathbf{L}\ a_{pp}x_p \end{cases} \tag{2}$$

Experimental results

In this paper, applies the stochastic chaos (SC) model and Stochastic volatility/jump diffusion (SVJD)model to artificially generated data, and the experiment was conducted in the MATLAB (R2014b). The SC model and the SVJD model were presented by Equation (4) and Equation (5). The total number of samples is 2000, since the artificial data in this paper are intended to replicate properties of higher-frequency data, we select a lag length of ten, among the data, the data of the first 1970 are for training and those of the late 20 are for testing.

$$\begin{cases} y_n = 4 \cdot a_n \cdot y_{n-1} \cdot \left(1 - y_{n-1}\right) \\ y_0 = 0.5 \\ a_n \sim U\left(0,1\right) \end{cases} \tag{4}$$

The stochastic term a_n is a draw from a random uniform distribution. The variable y_n depends on its own lag y_{n-1}, multiplied by a factor of 4.

$$\begin{cases} dS(t) = (r - d\overline{k}) \cdot S(t)dt + \sqrt{V(t)} \cdot S(t)dZ(t) \\ + k \cdot S(t)dq(t), S(0) = S; \\ dV(t) = (q - eV(t))dt + e\sqrt{V(t)} \cdot dZ_v(t), V(0) = V; \\ In(1 + k): \ j\ (In\left[1 + \overline{k}\right] - 0.5d, d^2) \end{cases} \tag{5}$$

Where $S(t)$ is the asset price, r is the risk free interest rate , d is the annual frequency of jumps, k is the jump size, $V(t)$ is the diffusion component of return variance, $dq(t)$ is a Poisson counter with intensity d , $prob(dq = 1) = d \cdot dt$, d is standard deviation, q is the mean of the conditional variance, e is the mean-reversion coefficient, e is the variance of the volatility process, $dZ(t)$ and $dZ_v(t)$ are correlated wiener processes with constant correlation r .

To validate the effectiveness of the proposed algorithm in this article, we select typical benchmark test function mean squared error (MSE) and squared correlation coefficient (R^2) to validate the optimization performance of algorithms, and squared correlation coefficient (R^2) is the better for more close to 1. The experimental results are shown in Table 1 Table 2, Figure 1 and Figure 2.

Table 1. The test results of SC model

Test Index	MSE	R^2	Time(s)	CCR	PCN
BP	0.1451	89.84%	97		
PCA-BP	0.0196	86.38%	86	87.12%	7
GA-BP	0.0023	90.18%	117		
PCA-GA-BP	0.0002	98.01%	110	87.47%	7

Table 2. The test results of SVJD model

Test Index	MSE	R^2	Time(s)	CCR	PCN
BP	0.2069	85.12%	96		
PCA-BP	0.0121	87.35%	88	88.62%	6
GA-BP	0.0009	95.28%	136		
PCA-GA-BP	0.0002	99.89%	114	93.37%	7

Where PCN represents principal component number, CCR represents cumulative contribution rate. Through the comparison among Table 1 and Table 2, Figure 1 and Figure 2, we can find that the testing errors of making combination predictions are more desirable than making mere BP prediction. The combination predictions model are more better by combining PCA with BP neural networks, and run time are more shorter.
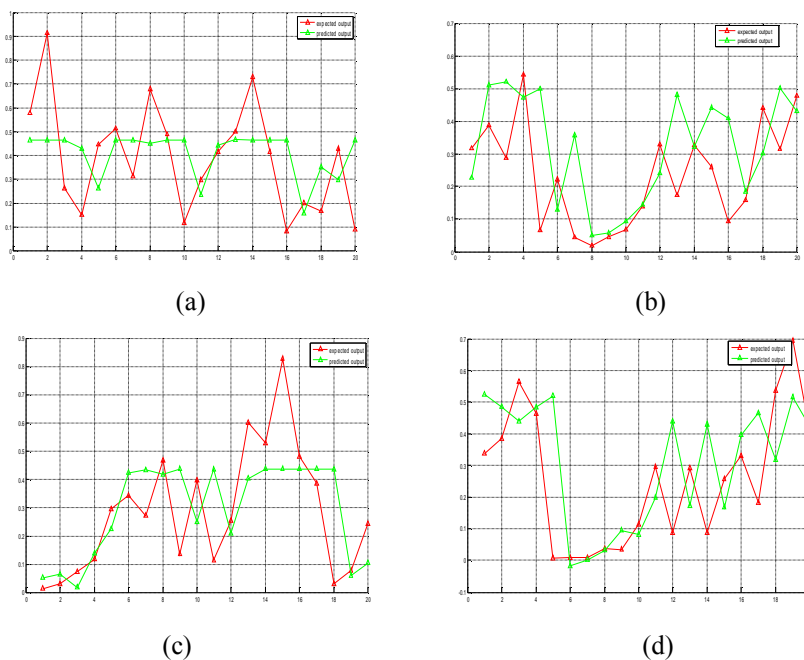

(a)

(b)

(c)

(d)

Figure 1. The test result of the SC model: (a) Test results of the SC model in BP; (b) Test results of the SC model in GA-BP; (c) Test results of the SC model in PCA-BP; (d) Test results of the SC model in PCA-GA-BP.
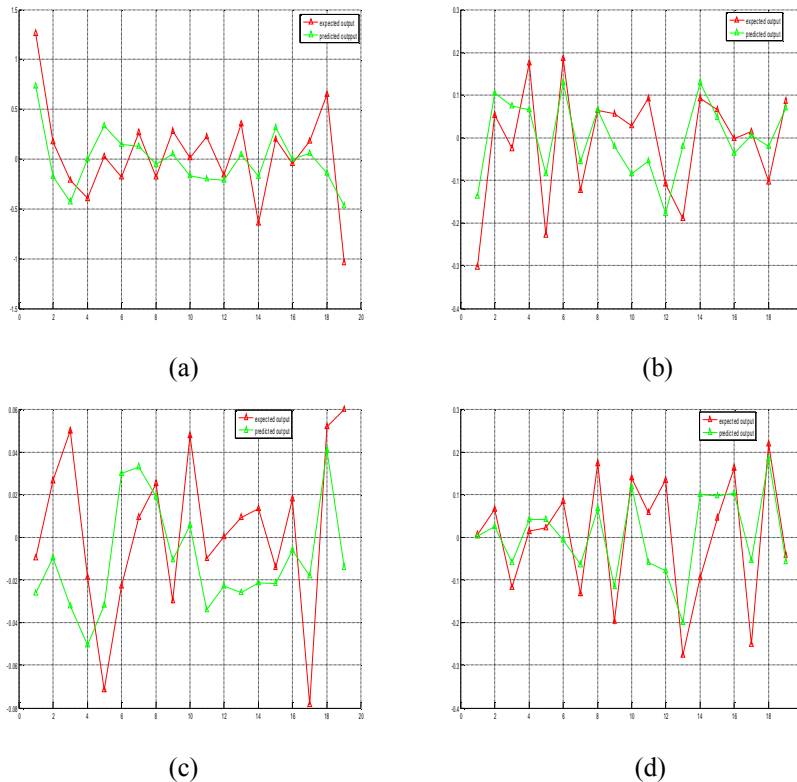
Figure 2. The test result of the SVJD model: (a) Test results of the SVJD model in BP; (c) Test results of the SVJD model in GA-BP; (e) Test results of the SVJD model in PCA-BP; (g) Test results of the SVJD model in PCA-GA-BP.

Conclusions

In this article, we put forward a forecast model that hybridizes genetic algorithm, principal component analysis and artificial neural network methods. Experimental results demonstrate that combination predictions model produces highly accurate ensembles that outperform mere BP predictions, so that the model can be used for forewarning applications.

References

[1] Principe, Jose C., and Craig L. Fancourt. "Artificial neural networks."Handbook of Global Optimization 2 (2013): 363.

[2] Islam, Sadeka, et al. "Empirical prediction models for adaptive resource provisioning in the cloud." Future Generation Computer Systems 28.1 (2012): 155-162.

[3] Han, Hong-Gui, Jun-Fei Qiao, and Qi-Li Chen. "Model predictive control of dissolved oxygen concentration based on a self-organizing RBF neural network." Control Engineering Practice 20.4 (2012): 465-476.

[4] Xiaochuan Wang, Feng Shi, Yang Li. The analysis of 43 cases in MATLAB Neural Network[M]. Beijing: Beijing University of Aeronautics and Astronautics Press,2013.8.

[5] Zhang, Yilong, Yajun Cai, and Huanhuan Wu. "Research on Cost Estimation of Highway Project Based on the GA-BP Algorithm." Proceedings of the 19th International Symposium on Advancement of Construction Management and Real Estate. Springer Berlin Heidelberg, 2015.

[6] Gao, HuiXuan. Principal Components Analysis. Applied Multivariate Statistical Analysis; Beijing: Beijing University Press, CHINA, 2005; pp. 265-290.