

RBF Model Based on the Improved KELE Algorithm

Chen Xiu-rong, Tian Yi-xiang

School of Management and Economics, University of Electronic Science and Technology of China, Chengdu, China

Email address:

465123149@qq.com (Chen Xiu-rong), 15982859240@139.com (Tian Yi-xiang)

To cite this article:

Chen Xiu-rong, Tian Yi-xiang. RBF Model Based on the Improved KELE Algorithm. *Science Journal of Business and Management*. Vol. 5, No. 3, 2017, pp. 101-104. doi: 10.11648/j.sjbm.20170503.12

Received: March 24, 2017; Accepted: April 8, 2017; Published: May 4, 2017

Abstract: Firstly, we use the idea of mapping by kernel function of KECA to transfer original global nonlinear problem into global linear one under the high-dimensional kernel feature space to improve the manifold learning dimension reduction algorithm LLE, then put the results obtained from KELE into RBF, constructing RBF model based on KELE. And we choose the foreign exchange rate time series to verify the improved RBF model, and the results show that the improved KELE can effectively reduce the dimension of samples and the prediction accuracy of the RBF model based on KELE is increased obviously.

Keywords: Locally Linear Embedding, Kernel Entropy Component Analysis, Kernel Entropy Linear Embedding

1. Introduction

Exchanges play an important role in determining the success of many businesses and financial institutions around the globe. Many researchers at home and abroad have conducted studies over financial time series forecast, such as daily exchange rates forecast study of the Indian Rupee against the United States Dollar, British Pound, Euro and Japanese Yen utilizing ARIMA model, neural network model and Fuzzy Neuron separately by Babu (2015) [1], Luo (2016) [2] investigated the realised volatility forecasts for the short, mid, and long term by developing the HAR models with Bayesian approaches and employing the high-frequency data of the China Stock Index 300 future for the period from 16 April 2010 to 21 May 2014; Hafezi (2015) [3] and Zhang (2015) [4] separately applied bat-neural network multi-agent system (BNNMAS) and a new dynamic business network energy cascading mode (ECM) into the forecast of stock price. Thus, how to accurately predict exchange rate will largely benefit them, however, which is considered as a challenging work for researchers. Several forecasting techniques have been proposed in order to gain some advantages. Zorzi et al. took a different route and assessed the Dornbusch model on the basis of its forecasting performance [5]. Li et al. mentioned that the only unanimous opinion is monetary fundamental approach have a growing effect on exchange rate movements as the time horizon increases [6]. Babu et al. used ARIMA model and Neural

Network to predict exchange rates [7]. Though all these models are available in the application of forecasting exchange rates, there usually exist problems such as high dimension of input samples and algorithm convergence, which will inevitably bring disadvantages to the performance for exchange rates forecasting.

As a relatively new non-linear dimension-reduction algorithm, LLE (Locally Linear Embedding) integrates the linear characteristics and the non-linear advantages and is widely used in dimension reduction, image data clustering and text recognition and other fields. However, how to determine the domain parameter is a difficult challenge. Whether the domain is set proper or not has a straight influence on the dimension-reduction effects. Based on this, in this paper, we firstly use the idea of KECA (Kernel Entropy Component Analysis) to improve LLE as the KELE (Kernel Entropy Linear Embedding) algorithm, transforming the non-linear problems into the linear ones by kernel function mapping, and avoided the risk caused by the improper domain. Then we combine the KELE with RBF (Radial Basis Function) network to predict the foreign exchange time series.

2. KELE Algorithm

2.1. LLE

The algorithms commonly used in manifold learning are

LLE, Isomap and Laplacian eigenmaps (LE), etc, among which the most representative is LLE. Compared to traditional linear dimension-reduction algorithms, it is more consistent with the original nonlinear characteristics of data, which makes it more effectively to mine the low-dimensional embedding representation of data. Assume the input data is $X=\{X_1, X_2, \dots, X_n\}$, $X_i \in R^d$, the dimension-reduced data is $Y=\{Y_1, Y_2, \dots, Y_n\}$, $Y_i \in R^m$ ($m < d$). The process of LLE algorithm is as follows:

(1) Determine the domains

Use the k-nearest algorithm, choosing the k points whose Euclidean distances are nearest to the sample points as the domains of this sample point.

(2) Determine the locally optimal reconstruction weight matrix

Find the linear combination of the sample point and the points in its domain for each domain to construct the reconstruction weight matrix w. In order to approach the wholly non-linear topological features maximumly, the cost function represented by equation 1 should be minimal:

$$\varepsilon(W) = \sum_{i=1}^n \left\| X_i - \sum_{j=1}^k W_{ij} X_{ij} \right\|^2 \quad (1)$$

Where $\sum_{j=1}^k W_{ij} = 1$, and X_{ij} denotes the jth neighborhood point of the sample point X_i . The cost function of X_i is:

$$\varepsilon_i(W) = \left\| X_i - \sum_{j=1}^k W_{ij} X_{ij} \right\|^2 = \sum_{j=1}^k \sum_{m=1}^k W_{ij} Q_{im} Q_{jm}^{(i)} \quad (2)$$

(3) calculate the dimension-reduced vectors

Each dimension-reduced vector should satisfy the local topology represented by reconstruction weight matrix of equation 2, so how to calculate the dimension-reduced vectors is transformed into figure out each low-dimensional vector Y_i in equation 3 when it reaches the minimum value:

$$E(Y) = \sum_{i=1}^n \left\| Y_i - \sum_{j=1}^k W_{ij} Y_{ij} \right\|^2 \quad (3)$$

Where the restrictions is: $YY^T = I$, then the problem is finally changed into the optimization problem of conditional constraints in equation 4:

$$\begin{cases} \min \sum_{i=1}^n \left\| Y_i - \sum_{j=1}^k W_{ij} Y_{ij} \right\|^2 = \min(YMY^T) \\ S.t: YY^T = I \end{cases} \quad (4)$$

Where $M=(I-W)^T(I-W)$, so the final problem is calculate the eigenvectors corresponded to the first minimum m eigenvalues, and they are the dimension-reduced vectors we need to work out.

How to select the domain parameters in step 1 determines the whole dimension-reduction effects. In this paper we by the KELE algorithm proposed, transforming the global non-linear problems into global linear one in the

high-dimensional kernel space, we avoid the risk caused by the unproper sets of domain parameters. The specified improved process and the flow chart of KELE are introduced in section 2.2.

2.2. KELE Algorithm

First, we upgrade LLE algorithm with KECA-based ideas [8, 9], then turn the global nonlinear problem into global linear problem in high dimensional kernel space through kernel function mapping. Assume the dimension of sample X is N, and the probability density is p(x), then the Reny entropy is:

$$H(p) = -\lg \int p^2(x) dx \quad (5)$$

Set $V(p) = \int p^2(x) dx$, then how to estimate the Reny entropy could be transformed into the estimation of $V(p)$. By use of

Parzen window $\hat{p}(x) = \frac{1}{N} \sum_{x_i \in D} K_\sigma(x, x_i)$ to estimate $V(p)$, we can get:

$$\hat{v} = \frac{1}{N} \sum_{x_i \in D} \hat{p}(x_i) = \frac{1}{N} \sum_{x_i \in D} \frac{1}{N} \sum_{x_j \in D} K_\sigma(x, x_j) = \frac{1}{N^2} \mathbf{1}^T \mathbf{K} \mathbf{1} \quad (6)$$

Where K is the kernel matrix, 1 is n-dimensional column vector.

From equation 6, we can know that the kernel matrix consisted by the vector inner product of the mapped high-dimensional kernel space can be used to estimate the amount of the original sample data information, each value of the element can be seen as a contribution to the total amount of information. So it can be considered that the inner product of the vector determines the local information characteristics of the samples, and the sum of all inner products represents the global information characteristics of samples. So in this paper, we use the kernel matrix H to replace the reconstructed weight matrix W between the original sample points and the neighborhood points, thus the reconstructed weights between the sample points and the neighborhood points are transformed into those between the sample points and the global points.

Compared to the data set X before reducing the dimension, the data set dimension-reduced Y should be not only consistent with it in the local information features of the high-dimensional kernel space, but also the Renyi information entropy. Based on the idea, we add p balance lags in the process of obtaining the dimension-reduced vector, i.e. transform the problem of figuring out the eigenvectors corresponding to the first minimum m eigenvalues into figuring out the eigenvectors corresponding to the first minimum m+p eigenvalues, and then from which choose the m eigenvectors with the least loss of Renyi entropy information. The KELE algorithm is as follows:

(1) Enter the d-dimensional data set $X = \{X_1, X_2, \dots, X_n\}$, $X_i \in R^d$, and determine the reduction dimension m and the number of balance lags;

- (2) Standardize the data;
- (3) Construct a kernel matrix using the Parzen window function and denote it as $H^{n \times n}$, where

$$h_{ij} = \exp(-\|X_i - X_j\|^2 / \sigma^2) [10] \quad (7)$$

(4) Use the kernel matrix H in (3) to replace the previous reconstructed matrix W between the sample points and its neighborhood points, the new $M = (I-H) T (I-H)$, where I is a $n \times n$ unit matrix;

(5) Calculate the eigenvectors corresponding to the first smallest $m+p$ eigenvalues of square matrix M.

(6) From the eigenvectors obtained from (5), choose the first m eigenvectors with the least loss of Renyi information entropy after reducing the dimension of data set X.

3. The Empirical Study

We choose the daily data of three foreign exchange time series of GBP / USD, USD / JPY and AUD / USD from June 30, 2010 to June 30, 2015 to study. All of them contain 1302 trading days, of which the former 1200 is defined as the training set and the last 102 is defined as the test set.

3.1. Determine the Parameters of the Models

3.1.1. Determine the Number of RBF Hidden Layer Nodes

Because the determination of the reduced dimension of KELE is based on the network structure of RBF, we need to first determine the RBF structure. Given that the ratio of the training set and test set is almost 10:1, so we adopt the way of 10 times cross validation to choose the parameters and select the mean of average absolute errors obtained from the 10 times cross validation prediction as evaluation criteria. In our paper, we set the biggest number of hidden layer nodes is 30, then evaluate from 1 to 30. The evaluation results of the RBF under different parameters are shown as Figure 1:

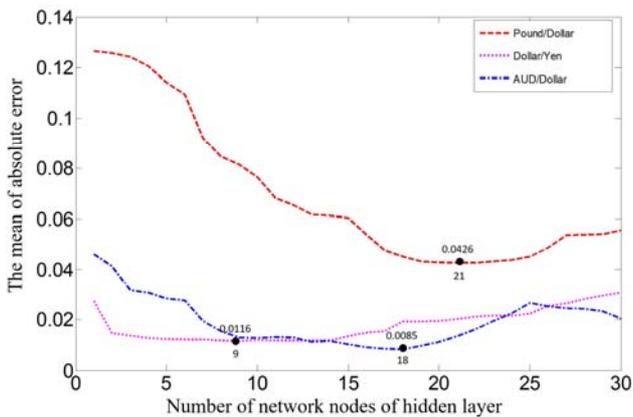


Figure 1. The evaluation results of RBF under different parameters for the three foreign exchanges.

From Figure 1, we can see that the average absolute errors of GBP / USD, USD / JPY and AUD / USD are minimal when the RBF hidden layer nodes are 21, 9 and 18 respectively.

3.1.2. Determine the Reduced Dimension of KELE

We set the reduced dimension of KELE as ten choices from 1 to 10, likewise, we adopt the way of 10 times cross validation to evaluate for each choice. The evaluation results according to the 10 choices is shown as Figure 2:

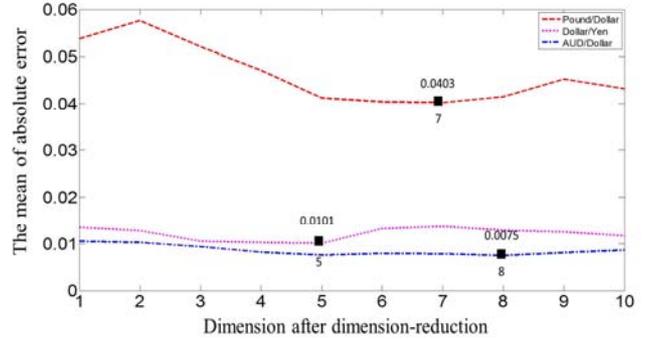


Figure 2. The evaluation results of the three foreign exchanges according to different reduced dimensions.

From Figure 2 we can see that the average absolute errors of GBP / USD, USD / JPY and AUD / USD are minimal when the reduced dimension are 7, 5 and 8 respectively.

3.2. The Empirical Results

To demonstrate the validity of the model constructed in this paper, we compare the prediction results of KELE-RBF with RBF. And we use the mean absolute error (MAE) indicator to evaluate the predictive effects of the two models:

$$MAE = \frac{\sum_{i=1}^T |y_i - y'_i|}{T} \quad (8)$$

Where $Y = (y_1, y_2, \dots, y_T)^T$ denotes the actual exchanges, $Y' = (y'_1, y'_2, \dots, y'_T)^T$, denotes the forecasted exchanges. T is the number of test samples. The evaluation results of the two models are shown in Table 1:

Table 1. The evaluation results of the two models.

MAE	RBF	KELE+RBF
GBP / USD	0.0361	0.0337
USD / JPY	0.0094	0.0083
AUD / USD	0.0082	0.0074

From Table 1 we can see that for the forecasting results of the three foreign exchanges, the accuracy of combination model of the RBF and KELE is improved in some degree in comparison with the single RBF model. It can be seen that the KELE algorithm can effectively reduce the dimension of data, remove the redundant information, and give the more optimized expression of data, and thus improve the prediction accuracy.

4. Conclusions

For the deficiency of LLE algorithm, we use the idea of

KECA to improve it and construct the KELE algorithm, then combine RBF to predict the three foreign exchanges. Compared to the single RBF, the accuracy of the KELE combination with RBF has been improved to a certain extent. The KELE algorithm proposed in this paper can effectively reduce the dimension of data, give the more optimized expression of data, and thus improve the prediction accuracy.

Acknowledgements

The authors acknowledge the National Natural Science Foundation of China (nos. 71171025) and National Social Science Foundation of China (nos. 14BJY174).

References

- [1] A. Babu and S. Reddy, "Exchange Rate Forecasting using ARIMA, Neural Network and Fuzzy Neuron," *Journal of Stock & Forex Trading*, vol. 2015, 2015.
- [2] J. Luo and L. Chen, "Realised Volatility Forecasts for Stock Index Futures Using the HAR Models with Bayesian Approaches," *China Accounting and Finance Review*, vol. 18, pp. 1-29, 2016.
- [3] R. Hafezi, J. Shahrabi, and E. Hadavandi, "A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price," *Applied Soft Computing*, vol. 29, pp. 196-210, 2015.
- [4] W. Zhang, C. Li, Y. Ye, W. Li, and E. W. Ngai, "Dynamic business network analysis for correlated stock price movement prediction," *IEEE Intelligent Systems*, vol. 30, pp. 26-33, 2015.
- [5] Ca'Zorzi M, Kocięcki A, Rubaszek M. Bayesian forecasting of real exchange rates with a Dornbusch prior [J]. *Economic Modelling*, 2015, 46: 53-60.
- [6] Li Q. How to forecast exchange rate, an unanswered puzzle [J]. 2010.
- [7] A. Babu and S. Reddy, "Exchange Rate Forecasting using ARIMA, Neural Network and Fuzzy Neuron," *Journal of Stock & Forex Trading*, vol. 2015, 2015.
- [8] R. Jenssen, "Kernel entropy component analysis," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, pp. 847-860, 2010.
- [9] B. Shekar, M. S. Kumari, L. M. Mestetskiy, and N. F. Dyshkant, "Face recognition using kernel entropy component analysis," *Neurocomputing*, vol. 74, pp. 1053-1057, 2011.
- [10] Y. Huang and G. Kou, "A kernel entropy manifold learning approach for financial data analysis," *Decision Support Systems*, vol. 64, pp. 31-42, 2014.