Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2013, 5(11):96-101

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Application of LSSVM to logistics demand forecasting based on grey relational analysis and kernel principal component analysis

[1]Xia Zhao and [2]Li-Yan Geng*

[1]Trade Union Committee, Shijiazhuang Tiedao University, Shijiazhuang, China
[2]School of Economics and Management, Shijiazhuang Tiedao University, Shijiazhuang, China

---

ABSTRACT

Precise forecasting of logistics demand is very important for logistics system planning and designing. However, logistics demand is affected by many factors, which will cause the complex model for logistics demand forecasting. To simplify the forecasting model and improve the forecasting precision, this paper proposed a least squares support vector machines (LSSVM) model based on grey relational analysis (GRA) and kernel principal component analysis (KPCA) for forecasting logistics demand. Firstly, the GRA was applied determining the main factors affecting logistics demand. Secondly, the KPCA was applied to eliminating the correlation among the main factors and extracting the nonlinear principal components. Finally, the extracted nonlinear principal components were as the input variables to build LSSVM model for logistics demand forecasting. The parameters in LSSVM were optimized by the adaptive inertia weight particle swarm optimization (AIWPSO) algorithm. The logistics demand in China was used to evaluate the effectiveness of the proposed model. The results indicate that the proposed model greatly reduces the dimensions of the input variables and improves the forecasting precision for logistics demand.

Keywords: Logistics demand forecasting, Least squares support vector machines, Grey relational analysis, Kernel principal component analysis, Adaptive inertia weight particle swarm optimization.

INTRODUCTION

Logistics system is characterized by uncertainty, nonlinearity and dynamics, which increases the difficulty for forecasting logistics demand. Some traditional forecasting methods, such as regression analysis and time series analysis establish forecasting model based on the mathematical theory and hypothesis. They can't describe the inner structure and complexity of logistics demand data. To improve the forecasting accuracy of logistics demand, artificial neural network (ANN) was used to forecast logistics demand [1-2] and has been proven to improve forecasting results to some extent. However, the applications of ANN often trapped into local minima because of its own disadvantages.

Support vector machines (SVM) was first introduced by Vapnik to solve the classification problem [3]. Based on the principle of structural risk minimization, SVM has been successfully applied to other fields such as data mining, regression and logistics demand forecasting [4]. Least squares support vector machines (LSSVM) [5] is a reformulation of SVM. Besides keeping the advantages of SVM, LSSVM has its own merits. That is, LSSVM improves the computation efficiency and reduces the parameters needed to be estimated by solving linear equations. However, the performance of LSSVM relies heavily on the choice of parameters. So, choosing the suitable parameters is the key to improve the performance of LSSVM.

Logistics system is a subsystem of social economic system and multiple factors affect the varying characteristics of logistics demand. These factors have different influence degree and influence law on logistics demand. Moreover, these factors also influence each other, which will lead to the correlation between the factors and then the information overlap. When constructing the forecasting model, if too many factors were taken into account, the structure of the forecasting model would become complex and the forecasting precision would decrease.

Grey relational analysis (GRA), based on grey system theory, is suitable to analyze the relationships among sequences in a system and determine the degree of influence of them. Literatures [6] and [7] have used GRA to get the main factors affecting synthesis characteristics of hydraulic valve and affecting railway freight volumes, respectively.

Kernel principal component analysis (KPCA) is a kernel-based principal component analysis, which can deal with the nonlinear interrelationships among the complex data. By extracting nonlinear feature information, KPCA can eliminate the correlation existed in multiple sequences and then reduce the dimensions of the data [8]. KPCA, coupled with GRA, has been used to determine the input and output variables of LSSVM for evaporation process modeling [9].

Combining GRA with KPCA, this paper proposes a LSSVM based model for forecasting logistics demand. GRA is utilized to gain the main factors affecting the characteristic of logistics demand and KPCA is used to find the nonlinear principal components from the main factors. Then, LSSVM, optimized by adaptive inertia weight particle swarm optimization (AIWPSO) algorithm, is used to build a forecasting model based on the nonlinear principal components.

EXPERIMENTAL SECTION

Grey relational analysis

The core idea of grey relational analysis (GRA) is to ascertain the relationships of different sequences through the geometric proximity between them. The more proximal the geometrical shape, the closer the relationship of the corresponding sequences. The degree of the relationship is described by the grey relational degree. The procedure of GRA is summarized as follows.

Step1 Get two types of sequence. Consider the data series $\{P_0(t)\},(t=1,2,...,N)$ as the reference sequence, the data series $\{P_i(t)\},(i=1,2,...,m,t=1,2,...,N)$ as the comparison sequence. Normalize the original data through mean value:

$$x_0(t)=\frac{P(t)}{\frac{1}{N}\sum_{t=1}^{N}P(t)},\ x_i(t)=\frac{P_i(t)}{\frac{1}{N}\sum_{t=1}^{N}P_i(t)} \tag{1}$$

Step2 Calculate the relational degree coefficient. At time t'(t'=1,2,…,N), the grey relational coefficient between the reference and comparison sequences is defined as:

$$L_{0i}(t')=\frac{\Delta_{\min}+\Delta_{\max}}{\Delta_{0i}(t')+\xi\Delta_{\max}} \tag{2}$$

where $\xi\in[0,1]$ is the distinguish coefficient; $\Delta_{0i}(t')=|x_0(t')-x_i(t')|$; $\Delta_{\max}$ and $\Delta_{\min}$ are defined as $\Delta_{\max}=\max_i\max_{t'}|x_0(t')-x_i(t')|$ and $\Delta_{\min}=\min_i\min_{t'}|x_0(t')-x_i(t')|$, respectively.

Step3 Calculate the relational degree. The relational degree is the average value of the relational coefficients and is given as:

$$r(x_0,x_i)=\frac{1}{N}\sum_{t'=1}^{N}L_{0i}(t') \tag{3}$$

Step4 Rank the relational degree. The relational degree is ranked from a big value to small one. Then we can know the degree of importance of the comparison sequence to the reference sequence.

Kernel principal component analysis

Let $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N)$ be a data matrix, with $\boldsymbol{x}_j \in R^m$ be the column vector at time j, for $j=1,2,\ldots,N$. Using the nonlinear mapping Φ, KPCA transforms $\boldsymbol{x}$ into a high-dimensional feature space F and the data covariance matrix in feature space can be written as:

$$\boldsymbol{C} = \frac{1}{n}\sum_{j=1}^{n} \Phi(\boldsymbol{x}_j)\Phi(\boldsymbol{x}_j)^T \tag{4}$$

where $\Phi(\boldsymbol{x}_j)$ is assumed to being mean centered. By introducing a kernel matrix $\boldsymbol{K}$ with elements $k(\boldsymbol{x}_i,\boldsymbol{x}_j)$, for $i,j=1,2,\ldots,N$, In F, inner products of two vectors can be evaluate as $k(x_i,x_j)=\Phi(\boldsymbol{x}_i)^{\mathrm{T}}\Phi(\boldsymbol{x}_j)$. Centralize $\boldsymbol{K}$ by

$$\bar{\boldsymbol{K}} = \boldsymbol{K} - A_n\boldsymbol{K} - A_n + \boldsymbol{A}_n\boldsymbol{K}\boldsymbol{A}_n \tag{5}$$

where $\boldsymbol{A}_n=(1/N)_{N\times N}$. Let $\boldsymbol{\beta}=(\beta_1,\beta_2,\ldots,\beta_n)^T$ be the orthonormal eigenvector associated with the pth largest positive eigenvalues $\lambda_1,\lambda_2,\ldots.,\lambda_p$, then the eigenvalue decomposition to $\bar{\boldsymbol{K}}$ is expressed as:

$$\bar{K}\boldsymbol{\beta} = N\lambda\boldsymbol{\beta} \tag{6}$$

The p kernel principal components in F are given as:

$$\boldsymbol{v} = \sum_{j=1}^{p}\beta_p\Phi(\boldsymbol{x}_p) \tag{7}$$

Let $\boldsymbol{z}$ be a test vector, the pth kernel principal component s_p corresponding to Φ is obtained by

$$s_z(p) = \boldsymbol{v}_p\Phi(z) = \sum_{j=1}^{N}\beta_{j,p}k(\boldsymbol{x}_i,\boldsymbol{z}) \tag{8}$$

Least squares support vector machines

Suppose a training data $\{(\boldsymbol{x}_l,y_l)\}$ for $l=1$ to n, where $\boldsymbol{x}_l \in R^d$ is the input vector and $y_l \in R$ is the corresponding output variable. The optimization problem for LSSVM is given as:

$$\min_{\omega,e} J(\omega,e) = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}\gamma\sum_{l=1}^{n}e_l^2 \tag{9}$$

$$y_l = w^T j\ \ (\boldsymbol{x}_l) + b + e_l,\ l = 1,\ldots,n. \tag{10}$$

where the mapping function $\varphi(\boldsymbol{x}_l)$ transfers x_l into a high-dimensional feature space. And γ is the regularized parameter. And ω, b are weight vector and bias term, respectively; e_l is the error variable at time l. Combining the Lagrangian of the optimization problem with the Karush-Kuhn-Tucker (KKT) condition, we can get the linear equations expressed in the following matrix:

$$\begin{bmatrix} 0 & \boldsymbol{1}^T \\ \boldsymbol{1} & \boldsymbol{\Omega} + \boldsymbol{I}/\gamma \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{Y} \end{bmatrix} \tag{11}$$

where $\boldsymbol{Y}=[y_1, y_2,\ldots, y_n]^{\mathrm{T}}$, $\boldsymbol{1}=[1,\ldots,1]^{\mathrm{T}}$, and $\boldsymbol{I}$ is an n order unit matrix. And $\boldsymbol{\alpha}=[\alpha_1, \alpha_2,\ldots, \alpha_n]^{\mathrm{T}}$ is the lagrange multipliers matrix with $\alpha_l \in R$. And $\boldsymbol{\Omega}$ is the matrix with elements $\Omega_l=\varphi(\boldsymbol{x}_l)^{\mathrm{T}}\varphi(\boldsymbol{x})$. The final LSSVM for regression is obtained by

$$y(\boldsymbol{x}) = \sum_{l=1}^{n} \alpha_l k(\boldsymbol{x}_l, \boldsymbol{x}) + b \tag{12}$$

where $k(\boldsymbol{x}_l,\boldsymbol{x})=\varphi(\boldsymbol{x}_l)^{\mathrm{T}}\varphi(\boldsymbol{x})$ is the kernel function and is needed to satisfy Mercer's condition.

RESULTS AND DISCUSSION

GRA and KPCA of the influential factors

This paper takes China logistics demand forecasting as an example. Different researchers use different measurement index for logistics demand. In this paper, the total social logistics cost is used as the index to measure the logistics demand.

Analyzing the logistics system comprehensively, it is found that there are 17 elements influencing the characteristic of the logistics demand. GRA is used to determine the key factors affecting the varying behavior of logistics demand. The total social logistics costs and the 17 influential elements are regarded as reference sequence and comparison sequences. The data used comes from the National Bureau of Statistics of China over the period from 1991 to 2011. The results are shown in Table-1.

Table-1 The calculated grey relational degree of influential factors

Influential factors	Grey relational degree	Influential factors	Grey relational degree
Gross domestic product(x1)	0.9677	Total freight traffic(x10)	0.8953
Total investment in fixed assets(x2)	0.8517	Total freight ton-kilometers(x11)	0.9131
Total output value of the primary industry(x3)	0.9296	Number of employed persons in logistics industry(x12)	0.8865
Total output value of the second industry(x4)	0.9763	Possession of civil trucks vehicles(x13)	0.9060
Total output value of tertiary industry(x5)	0.9465	Number of national owned railway freight cars(x14)	0.9084
Business volume of postal and telecommunication services (x6)	0.7780	Possession of civil cargo vessels(x15)	0.8966
Total value of imports and exports(x7)	0.9183	Total population(x16)	0.9044
Total retail sales of the consumer goods(x8)	0.9580	Retail price index(x17)	0.9028
Household consumption expenditure(x9)	0.9860		

From Table-1, there are twelve factors whose grey relational degree is larger than 0.9. It is believed that these twelve elements are the key influential factors for the varying characteristic of the logistics demand.

Using the Gaussian kernel function, the KPCA is utilized to extract the nonlinear feature information from the twelve important influential factors which is listed in Table-2. As is shown in Table-2, the cumulative contribution rate of the first two principal components is larger than 95%.

Table-2 Eigenvalue, Contribution rate, and Cumulative Contribution rate of the components by KPCA

Components	Eigenvalue	Contribution rate/%	Cumulative Contribution rate/%	Components	Eigenvalue	Contribution rate/%	Cumulative Contribution rate/%
1	0.0036	90.80	90.80	7	0.0000	0.03	99.98
2	0.0003	6.85	**97.65**	8	0.0000	0.02	100.00
3	0.0001	1.54	99.19	9	0.0000	0.00	100.00
4	0.0000	0.52	99.71	10	0.0000	0.00	100.00
5	0.0000	0.13	99.84	11	0.0000	0.00	100.00
6	0.0000	0.11	99.95	12	0.0000	0.00	100.00

LSSVM model for logistics demand forecasting

According to the results of the KPCA, the first two nonlinear principal components are chosen to construct LSSVM model. Gaussian kernel function is selected as the kernel function and the LSSVM for logistics demand forecasting is expressed as:

$$x_{0l} = \sum_{l=1}^{n} \alpha_l \exp(-\|\boldsymbol{s}_l - \boldsymbol{s}\|^2 / \sigma^2) + b \tag{13}$$

where the input vector s_l is the nonlinear principal components obtained by KPCA. The output variable x_{0l} is the total social logistics costs. And σ^2 is the kernel parameter. Coupling (11) with (13), two parameters, the regularized parameter γ and the kernel parameter σ^2, should be determined first. Due to LSSVM model can't select the proper parameters itself, cross validation method is commonly used to determine the values of γ and σ^2. However, cross validation method is characterized by randomness. It is difficult for this method to obtain the optimal parameters.

In this paper, AIWPSO algorithm is applied to optimize the parameters γ and σ^2 in LSSVM. The inertia weight w is one of the key factors for the PSO's optimization performance and convergence. Adjusting the inertia weight w according to the fitness function, AIWPSO algorithm balances the global search and local search capacity of the algorithm and improves the convergence speed [10]. The inertia weight w in AIWPSO algorithm is given as:

$$w=\begin{cases} w_{min}-\dfrac{(w_{max}-w_{min})\times(f-f_{min})}{(f_{avg}-f_{min})}, & f\le f_{avg} \\ w_{max}, & f>f_{avg} \end{cases} \tag{14}$$

where f is the current fitness function value. And f_{avg}, f_{min} are the average and minimum fitness function value. And w_{max}, w_{min} are maximum and minimum inertia weight, respectively.

The parameters of AIWPSO algorithm itself are given as follows: the scale of swarm M=10, the acceleration coefficients c_1=c_2=2.0, the maximum and minimum inertia weight w_{max}=0.9 and w_{min}=0.1, the maximum iteration number is set as 30. The input and output data sets are split into two sets. The initial 15 groups are used to train the model and the remainder 6 groups to test the out-of-sample forecasting performance of the model.

To demonstrate the performance of the proposed model (GRA-KPCA-LSSVM), this paper compares it with the other three models, including GRA-LSSVM, KPCA-LSSVM and LSSVM. For the GRA-LSSVM, the twelve important influential factors obtained by GRA are chosen as the input variables of LSSVM. For the KPCA-LSSVM, the three nonlinear principal components directly extracted from the seventeen influential factors using KPCA are chosen as the input variables of the LSSVM. For the LSSVM, all the seventeen influential factors are chosen as the input variables. The two parameters γ and σ^2 of these three models are also optimized by AIWPSO algorithm.

Three evaluation indices, the normalized root mean squared error (*NRMSE*), the normalized mean absolute error (*NMAE*) and the theil statistics (*THEIL*), are used to evaluate the forecasting performance of the four models.

Table-3 Logistics demand forecasts and relative errors by four models

Year	Actual values /100MY	GRA-KPCA-LSSVM		GRA-LSSVM		KPCA-LSSVM		LSSVM	
		Forecasts /100MY	Relative error/%	Forecasts /100MY	Relative error/%	Forecasts /100MY	Relative error/%	Forecasts /100MY	Relative error/%
2006	38414	38325	-0.23	38565	0.39	38927	1.33	38183	-0.60
2007	45406	45177	-0.50	44531	-1.93	45837	0.95	44402	-2.21
2008	54542	52144	-4.40	52097	-4.48	53036	**-2.76**	51113	-6.29
2009	60826	58781	-3.36	56468	-7.16	57932	-4.76	56407	-7.27
2010	70984	73717	3.85	66602	-6.17	67555	-4.83	65797	-7.31
2011	84000	91371	8.78	75966	-9.56	69810	-16.89	74477	-11.34
NMSE		0.3929		0.4852		0.6942		0.5676	
NMAE		0.2965		0.4038		0.4580		0.4745	
THEIL		0.0280		0.0362		0.0520		0.0426	

Note: MY denotes Million Yuan

Table-3 contains the forecasting results of the four models. It is observed that the maximum and minimum relative forecasting errors of the GRA-KPCA-LSSVM model are -0.23% and 8.78%, which are smaller than ones of the other three models. With the exception of the KPCA-LSSVM of 2008, GRA-KPCA-LSSVM model has the smallest relative forecasting error in the four models. As a whole, the forecasting results of the GRA-KPCA-LSSVM model are more accurate than those of the other three models. Meanwhile, the *NMSE*, *NMAE* and *THEIL* of the GRA-KPCA-LSSVM model are smaller than those of the other three models, which means GRA-KPCA-LSSVM model achieve better performance in logistics demand forecasting.

CONCLUSION

In this paper, a hybrid model called GRA-KPCA-LSSVM is proposed to improve the logistics demand forecasting accuracy. GRA is used to get the main factors affecting the logistics demand and KPCA is used to extract the nonlinear principal components. LSSVM is used to construct forecasting model based on the nonlinear principal components as inputs. The forecasting performance of the GRA-KPCA-LSSVM model is investigated through the use of China logistics data. The results show that the GRA-KPCA-LSSVM model has simplified structure. The forecasting performance of the GRA-KPCA-LSSVM model is better than those of the other models. In conclusion, GRA-KPCA-LSSVM model is a powerful tool to forecast logistics demand. Further works can focus on selecting other improved artificial intelligence technology, least squares wavelet support vector machine or relevance vector machine, to forecast logistics demand.

REFERENCES

[1]SY Chen; YC Guo; *Logist. Technol*., **2012**, 31(9), 231-233.
[2]Y Geng; SD Ju; YN Chen; *Logist. Technol*., **2007**, 26(7), 35-37.
[3]VN Vapnik. The Nature of Statistical Learning Theory, 1st Edition, Springer New York, **1995**; 119-166.
[4]YZ Hu; HY Lv; *Logist. Technol*., **2008**, 27(5), 66-68.
[5]JAK Suykens; J Vandewalle; *Neural Process. Lett*., **1999**, 9, 293-300.
[6]ZY Jia; J W Ma; F J Wang; L Wei; *Expert Syst. Appl*., **2010**, 37(2), 1250-1255.
[7]LY Geng; TW Zhang; *Tiedao Xuebao*, **2012**, 34(3), 1-6.
[8]YG Liang; LY Geng; ZF Zhang; *Transp. and Econ*., **2012**, 34(11), 63-67.
[9]XS Qian; CH Yang; *J. Chinese Comput. Syst*., **2012**, 33(8), 1851-1855.
[10]QY Wei; DS Zhang; WP Gao; *Huagong Zidonghua Ji Yibiao*, **2006**, 33(1), 24-37.