Research of chaotic SVM with Incorporated Intelligence Algorithm forecasting model in SCM

Zhifang Liu[1], Guoqiao Ren[2], and Xiaoting Shi[2]

[1] Center of Computer and Electric Network of Tangshan Teacher's College, Tangshan, 063000, Hebei Province, China

†† Department of Business Administration of North China Electric Power University . Baoding, Hebei Province 071003, China

Summary

Forecasting activities are widely performed in the various areas of supply chains for predicting important supply chain management(SCM) measurements such as demand volume in order management, product quality in manufacturing processes, capacity usage in production management, traffic costs in transportation management, and so on. The accuracy of forecasting has a great influence on the efficiency of SCM. According to the chaotic and non-linear characters analyze of SCM data, the model of support vector machines (SVM) based on Lyapunov exponents was established. The time series matrix was established according to the theory of phase-space reconstruction, and then Lyapunov exponents was computed to determine time delay and embedding dimension. A new incorporated intelligence algorithm is proposed and used to determine free parameters of support vector machines. Subsequently, examples of demand glass panels for CRT TV data. The empirical results reveal that the proposed model outperforms the SVM model. BP algorithm was used to compare with the result of SVM. The results show that the presented method is feasible and effective.

Key words:

Forecasting activities, supply chain management, Lyapunov exponents, support vector machines, embedding dimension

1. Introduction

Forecasting is one of the most important activities for enhancing productivity and improving quality in the whole organizational function because most forecasting results are very influential in making managerial decisions and evaluating performance of the company. In many cases, a forecasting system has been developed that can serve as a diagnostic tool for identifying potential distortions in organizational activities. For example, predicting the product defects during manufacturing processes is the beginning of quality improvement activities and the forecasting results are used to correctly control the future process for producing higher-quality products. This paper is concerned with the development of a computerized causal forecasting system for efficiently forecasting the various performance indices involved in the supply chain, which encompasses the whole processes of inbound and outbound logistics of the company, ranging from material supply to delivery to customers.[7]

The supply chain has three stages: purchase, production and distribution. These stages also involve three business entities shown in Fig. 1: suppliers, manufacturers, and customers who initiate and exert efforts on each corresponding stage. Many researchers also consider distributors and retailers as key business entities instead of customers. Whatever supply chain is defined, supply chain management (SCM) is nothing but the management of material and information flows required in the supply chain both in and between facilities, such as vendors, manufacturing and assembly plants and distribution centers. Although many diverse functions and activities are included in each stage, this paper is only focused on the forecasting activities. Recognizing the importance of forecasting activity will determine how effective the company's supply chain is in dealing with the environmental uncertainty it faces. The forecasting activities in a supply chain are different from each other according to its forecast variables and the involved stages; these forecast values are the essential information for the supply chain to function well.

Scholars who come from many countries have focused their studies on forecasting, and they have proposed many methods which can be classified into two sorts: one is a conventional method which takes advantage of time series, the other is the novel method which makes use of Artificial Neural Network. The activities time series of SCM is influenced by various factors so the change of activities time series presents complicated and non-linear characteristics. As a result, when traditional forecasting models are applied to such a complicated time series, the forecasting accuracy is always not satisfied. In order to meet the practical requirements, researchers have presented a lot of new approaches to forecast, using grad model, nonlinear model and artificial intelligence. Among AI there are algorithms like expert system, artificial neural network (ANN), logic, genetic algorithm. But the systemic and practical studies are still inadequate.

In the last twenty years, chaotic theory and statistic learning theory have been well-developed and the usage of

Manuscript received October 10, 2006.
Manuscript revised October 25, 2006.
.

activities required in SCM prediction has become more and more acceptable. Support Vector Machines (SVM) is based on the statistic theory which is presented by Vapnik. It is a machine learning algorithm which began to be used in the middle of ninetieth century. Statistic theory employs the criterion that minimizes the structure risk, meanwhile it can lower the global error of model. It raises the generalization capability of the model, which is more prominent in the small-sample learning.[5]-[6]

The method of phrase space reconstruction is used in this paper, it can disclose the complex the dynamic character of system in phase space which can't be disclosed by traditional method. It tests that the demand data has chaotic character, The time series matrix was established according to the theory of phase-space reconstruction, and then Lyapunov exponents was computed to determine time delay and embedding dimension. a new immunogenetic particle swarm incorporated intelligence algorithm is proposed in this paper in order to hunt the free parameter of the SVM model, The method is tested by the actual data, the result prove that the model has high training speed and forecasting accuracy.

2. Chaotic Time Series AND Lyapunov Exponents

Base on Chaos theory, the drive factors have influenced each other in chaotic system. Therefore the digital points which are got according to time are relative. At present, people are employing the phase space delay coordinate reconstruction method in general to analyze the factors of serial dynamics. Generally, the dimension is very great even infinite in the phase space in system, but the dimension can't be known in many situation. In fact, the phase space delay coordinate reconstruction method can expand the given time series $x_{1,}x_2,\cdots,x_{n-1},x_n,\cdots$ to three-dimensional and even higher dimensional space and the information which exposed sufficiently from time series can be classified and extracted. [1]- [2]

2.1 Reconstruction of Phase Space

The technology of reconstruction phase space is the premise to calculate Lyapunov exponents. In electric power system, actual activities required in SCM series of single argument { $x(t_j)$ =1, 2, 3… n} can be got with the gap Δt. The structural character of system attractors is contained in this time series. The specific method, which can estimate the information of phase space reconstruction in single argument time series, is:

$$\begin{array}{lllll}
x(t_1) & x(t_2) & \ldots\ x(t_j) & \ldots & x(t_n-(m-1)\tau) \\
x(t_1+\tau) & x(t_2+\tau) & \ldots\ x(t_j+\tau) & \ldots & x(t_n-(m-2)\tau) \\
x(t_1+2\tau) & x(t_2+2\tau) & \ldots\ x(t_j+2\tau) & \ldots & x(t_n-(m-3)\tau) \\
\ldots & \ldots & \ldots & & \ldots \\
x(t_1+(m-1)) & x(t_2+(m-1)) & \ldots\ x(t_j+(m-1)) & \ldots & x(t_n) \\
y(t_1) & y(t_2) & \ldots\ y(t_j) & \ldots & y(t_n-(m-1)\tau)
\end{array}$$

In this method, the time series can be continued to m-dimensional phase space. The time delay is $\tau = k\Delta t$ (Δt =1, 2….). In previous permutation, every column makes up a phase point of m -dimensional phase space. And each phase point has m components. These $n_p = n-(m-1)\tau$ phase points { $x(t_j)$, j =1, 2… n_p } make up a facies pattern in m -dimensional phase space, and the continuation of these phase points describes the evolutionary trace of system in the phase space.

2.2 Calculation of Lyapunov Exponents

A. Wolf submitted a method which is to extract maximal Lyapunov exponents in single argument time series [2]-[3]. The process is:

(i) Reconstructing m-dimensional phase space with time series.

(ii) Choosing minimal τ which marks the correlation among phase space.

(iii) In the continuation m-dimensional phase space, the initial phase point $A(t_1)$ is chosen as a reference point. There are m components in the phase space, they are: $x(t_1), x(t_1+\tau), x(t_1+2\tau),\cdots x(t_1+(m-1))$.

According to the following formula:

$$L_{nbt} = Min\left[\left\|Y_i - Y_j\right\|\right] \qquad i \neq j \qquad (1)$$

$B(t_1)$ which is the nearest neighborhood point to $A(t_1)$ can be impetrated. L_{nbt} which is assumed $L(t_1)$ means the distance between $A(t_1)$ and its nearest neighborhood point in Euclidean meaning. Suppose $t_2 = t_1 + k\Delta t$ with $k\Delta t$ as the step length and $A(t_1)$ evolves into $A(t_2)$, meanwhile $B(t_1)$ evolves into $B(t_2)$, then the distance $A(t_2)B(t_2) = l(t_2)$ is got. Let λ_1 represent the rate of exponential growth and $l(t_2) = L(t_1)2^{\lambda_2}$, then the following equation can be got.

$$\lambda_1 = \frac{1}{k(t_2 - t_1)} \log_2 (l(t_2) / L(t_1)) \qquad (\Delta t = 1) \qquad (2)$$

λ is the Lyapunov exponent. It can take advantage of λ to judge the stability of time behavior of the system.

(iv) Search a small neighborhood point $C(t_2)$ which subjects to the angle θ_1 in the nearest neighborhood points to $A(t_2)$ (If it can't meet the two conditions: small θ_1 and neighborhood, it should still choose $B(t_1)$). Supposing $t_3 = t_2 + k\Delta t$, $A(t_2)$ evolves into $A(t_3)$ and $C(t_2)$ evolves into $C(t_3)$ $A(t_2)C(t_2) = L(t_2)$ and $A(t_2)B(t_2) = l(t_2)$, then:

$$\lambda_2 = \frac{1}{k} \log_2 (l(t_3) / L(t_2)) \qquad (3)$$

It can not stop carrying out the previous steps until it reaches the end of point-group $\{X(t_j), j = 1, 2, \cdots, n_p\}$. Then choose the average of the calculated rates of exponential growth as the maximal estimated value of Lyapunov exponent. That is:

$$LE_1(m) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{k} \log_2 \frac{l(t_i - 1)}{L(t_i - 1)} \qquad (4)$$

$N = n_p / k$ means total steps of step length.

(v) It can not stop increasing embedding dimension m in turn and carrying out the steps（3）～（4）until the estimated value $LE_1(m)$ of the exponent keeps stable and $LE_1(m_0) = LE_1(m_0 + 1) + LE_1(m_0 + 2) = \cdots = LE_1$. LE_1 is just the maximal Lyapunov exponent.

3 SVM Regression Theory [4]-[6]

Suppose a set of data (x_i, y_i), $i = 1, 2 \cdots n$, $x_i \in \boldsymbol{R}^n$ are given as input , $y_i \in \boldsymbol{R}$ are the corresponding output. SVM regression theory is to find a nonlinear map from input space to output space and map the data to a higher dimensional feature space through the map, then the following estimate function is used to make linear regression.

$$f(x) = [\omega \cdot \phi(x)] + b \qquad \phi : \boldsymbol{R}^m \to \boldsymbol{F}, \quad \omega \in \boldsymbol{F} \qquad (5)$$

b is the threshold value. The problem of the function approximate is equivalent with the minimizing the following problem.

$$\boldsymbol{R}_{reg}[f] = \boldsymbol{R}_{emp}[f] + \lambda \|\omega\|^2 = \sum_{i=1}^{s} \boldsymbol{C}(e_i) + \lambda \|\omega\|^2 \qquad (6)$$

$\boldsymbol{R}_{reg}[f]$ is objective function and s is the number of the sample. $e(\bullet)$ is loss function and λ is adjusting constant meter. The following loss function can be gained concerning with the rarefaction character of the linear insensitive loss function ε.

$$|y - f(x)|_\varepsilon = \max\{0, |y - f(x) - \varepsilon|\} \qquad (7)$$

Empirical risk function is：

$$\boldsymbol{R}_{emp}^{\varepsilon}[f] = \frac{1}{n} \sum_{i=1}^{n} |y - f(x)|_\varepsilon \qquad (8)$$

According to statistic theory, the regression function is determined by minimizing the following functions.

$$\min \left\{ \frac{1}{2} \|\omega\|^2 + \boldsymbol{C} \sum_{i=1}^{n} (\xi_i^* + \xi_i) \right\} \qquad (9)$$

$$y_i - (\omega \cdot \phi(x) - b \le \varepsilon + \xi_i^* \qquad (10)$$

$$(\omega, \phi(x)) + b - y_i \le \varepsilon + \xi_i \qquad (11)$$

$$\xi_i, \xi_i^* \ge 0 \qquad (12)$$

$\boldsymbol{C}$ is used to equalize the complicated item of the model and the parameters of the item of training error. ξ_i^* and ξ_i are relaxation factors and ε is insensitive loss function. The problem can be converted into the dual problem：

$$\max[-\frac{1}{2} \sum_{i,j=1}^{n} (a_i^* - a_i)(a_j^* - a_j) K(X_i, X_j) + \sum_{i}^{l} a_i^* (y_i - \varepsilon) - \sum_{i=1}^{n} a_i (y_i - \varepsilon)] \qquad (13)$$

$$\text{s.t.} \qquad \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_i^* \qquad (14)$$

$$0 \le a_i^* \le \boldsymbol{C} \tag{15}$$

$$0 \le a_i \le \boldsymbol{C} \tag{16}$$

Solve the problem, then the regression equation is:

$$f(x) = \sum_{i=1}^{n} \left(a_i - a_i^*\right) k(X_i, X) + b \tag{17}$$

where $\bar{b}$ is calculated as follow functions: Choose $\bar{a}_j$ or $\bar{a}_j^*$ in open interval $(0, \frac{C}{l})$, if $\bar{a}_j$ is selected then:

$$\bar{b} = y_j - \sum_{i=1}^{l} (\bar{a}_i^* - \bar{a}_i)(x, x) + \varepsilon \tag{18}$$

if $\bar{a}_j^*$ is selected then:

$$\bar{b} = y_k - \sum_{i=1}^{l} (\bar{a}_i^* - \bar{a}_i)(x_i, x_k) - \varepsilon \tag{19}$$

4 Immunogenetic particle swarm incorporated intelligence algorithm

The parameters σ penalty factor C and $\varepsilon-$insensitive loss function are important to the efficiency and general forecasting performance of the algorithm, they are important contents in building the SVM model..The genetic algorithm has the competence of fast global hunt, but it does not make use of the feedback information so that the efficiency of solution is low; it can't prevent the appearance of the degeneracy in some degree although it assurance the evolution of the population. The optimum ratio of the particle swarm algorithm, its solution swarm can convergence to the optimum solution in less steps, but it has the demerit of slow speed, low accuracy and easy to divergence result from lack initial information.

In order to overcome the demerit of the two algorithms and utilize the merit each other, we associate the principle of immun0logy and the genetic algorithm to improve the whole performance of the algorithm, utilize selective and purposive the information of the solution procedure so that suppress the appearance of the degeneracy in optimization procedure. The paper utilizes the high ratio, hunting random and the initial information created by the global convergence, then utilizing the character of the easy to realize, brevity, high performance of solution and high convergence speed. The incorporated intelligence algorithm has the higher performance of solution than the genetic algorithm and higher performance of time, so that a heuristic algorithm which high performance in time and solution. The new immunogenetic particle swarm incorporated intelligence algorithm is used to hunt the parameters σ penalty factor C and $\varepsilon-$insensitive loss function.The steps of the incorporated intelligence algorithm are illustrated as follows:

Step 1 Input the goal function as antigen: $F = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{L-\bar{L}}{L}\right| \times 100\%$ where L is actual load data, $\bar{L}$ is the forecasting load data;

Step 2 (Initialization). Generate randomly an initial population of chromosomes. The three free parameters σ , penalty factor C and ε are encoded in a binary format; and represented by a chromosome;

Step 3 Calculate every compatible character of every antigen and antibody: classify the data into training set and test set, train the network and then test network to calculate the F ;

Step 4 Create memory cell in order to record the excellent antibody in evolution procedure, if reach the condition of end or the maximum iteration step go to step 7;

Step 5 Calculate the density and existence of every antibody, survival select and refill antibody;

Step 6 Cross, inherit and variance, go to Step 3;

Step 7 Initiate the parameter, create the initial information from the optimum solution of the immunogenetic algorithm, initiate m particles as initial population, calculate the sufficiency and record the code number of the best particle, set the position of the best particle as g_{best} ;

Step 8 Evaluate every particle by calculating its sufficiency, set the position of the particle as p_{best} and renew individual extremeness if the sufficiency is better than the current extremeness; set the best particle as g_{best} , record the position and renew the global extremeness if the best particle of all the particle is better than the current global extremeness;

Step 9 Renew the position and speed of the particle,

$$v_{id}^{k+1} = v_{id}^k + c_1 rand_1^k ((p_{best})_{id}^k - x_{id}^k) + c_2 rand_2^k ((g_{best})_d^k + x_{id}^k) , \tag{20}$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} , \tag{21}$$

where v_{id}^k is the speed in dimension of the particle i in time k , c_1, c_2 is the coefficient of acceleration in order to adjust maximum step length of the particle fly to the best particle;

Step 10 Check the particle, if reach the end condition then output the optimum solution (the parameter) else go to Step 8.

5 Application and Analysis

5.1 Choosing Samples

For experiments, we used a group of real process data and quality (‘‘blister’’) data that were collected every day for 10 months in three different glass manufacturing lines where the demand glass panels for CRT TV.

5.2 Chaos Analysis

For the training sample, τ =1 is chosen and Wolf method is used to compute Lyapunov exponents and embedding dimension. According to the theory, Lyapunov exponents λ begin to show stationary trend when the embedding dimension is 15. The required time series shows chaotic character because λ >0. The embedding dimension is 15 and the number of phase points is 305. The above parameters are used to reconstruct the phase-space. λ (Lyapunov exponent) changes with m (embedding dimension). When embedding dimension is 15, Lyapunov exponents begin to show stable tendency.

5.3 Prediction Process

The incorporated intelligence algorithm is used to hunt the optimum solution and exponents. The variation of the fitness can be shown in Fig-1. the exponent is:

C =59.205908, ε =0.000946, $\sigma^2 = 0.617695$

$$K(x,x') = \frac{1-q^2}{2(1-2q\cos(x-x')+q^2}, \forall x, x' \in R \qquad (21)$$

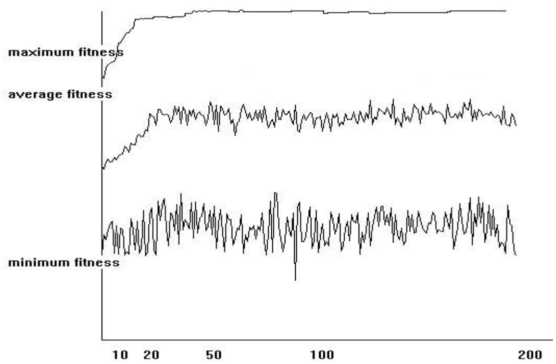

Fig. 1. The variation of the fitness

Table 1: Comparison of the predicted values and evaluating indicators

Date	E_r SVM(15)	E_r SVM(14)	E_r SVM(16)	E_r BP(15)
16	1.98%	2.81%	3.12%	5.31%
17	1.17%	3.06%	2.95%	4.74%
18	-2.13%	-4.15%	-3.99%	-6.82%
19	3.42%	4.83%	4.75%	5.47%
20	-1.76%	-2.03%	-2.35%	-3.24%
21	-0.83%	-0.99%	-1.05%	-1.84%
22	1.39%	2.57%	3.04%	4.87%
23	3.28%	0.98%	4.77%	-4.05%
24	0.11%	0.83%	1.39%	2.31%
25	-2.84%	-3.57%	-3.98%	-5.04%
RMSRE	0.0163%	0.22%	0.29%	1.36%

BP algorithm is used to make prediction with sigmoid function. The parameters are chosen as the following: the node number of input layer is 14 and the node number of output layer is 1. The node number of interlayer is 8 according to the experience. The system error is 0.001 and the maximal interactive time is 5000. The results are shown in table Ⅰ.

6 Conclusions

The results show that SVM based on Lyapunov exponents has great effectiveness for activities required data forecasting. And the conclusions are shown as the following:

(i) The required data show apparent chaotic characteristic. Chaotic time series is established and chaotic parameters are computed, then SVM prediction model is established to make prediction. The real required data prediction shows that the model is effective in activities required forecasting.

(ii) The embedding dimension is chosen through Lyapunov method. The predicted results with chosen dimension and other random dimension are compared. The comparison shows that the approach is scientific and rational. The results show there is a suitable embedding dimension which is used to predict the demand data effectively. The predicted values by the model with chosen dimension are highly accurate.

(iii) In the condition of the same dimension, SVM is much higher than BP in accuracy. The reason is the difference between BP and SVM. BP network is based on empirical risk minimization and easy to trap in local optimum. SVM is based on structure risk minimization. Local optimum is globe optimum. SVM exhibits better generalization ability than BP network.

Because the political and other information are hard to be acquired while the data of activities required are easy to be acquired in fact, the model of SVM based on lyapunov exponents is more significant in the application than the models which need more activities required data or the models which need political and other information.

References

[1] Jinghu Lv, Junan Lu and Shihua Chen, *Analysis and Application of Chaotic Time Series* (Book style). Wuhan: Wuhan University Press, 2002.
[2] Kehui Sun, Guoqiang Tan and Liyuan Sheng, "Design and Implementation of Lyapunov Exponents Calculating Algorithm (Translation Journals style)," *Computer Engineering and Application*, Vol.35, 2004, pp.12-14.
[3] Guohui Li,Shiping Zhou and Deming Xu, "Computing the Largest Lyapunov Exponents form Time Series (Translation Journals style)," *Journal of Applied Sciences*, Vol.21, 2003,pp.127-131.
[4] Vladimir N.Vapnik and Xuegong Zhang, *Nature of Statistics Theory* (Book style). Beijing:Tsinghua University Press, 2000.
[5] Lin Zhang, Xianshan Liu and Hejun Yin, "Application of Support Vector Machines Based on Time Sequence in Power System Load Forecasting. Power System Technology (Translation Journals style)," Vol.29, 2004,pp. 38-41.
[6] Shuxing Du and Tiejun Wu, "Support Vector Machines for Regression (Translation Journals style)," Journal of System Simulation, Vol.15, 2003, pp.1580-1586.
[7] Bongju Jeong, Ho-Sang Jung, Nam-Kyu Park. A computerized causal forecasting system using genetic algorithms in supply chain management. The Journal of Systems and Software 60 (2002) 223–237.
[8] LIU H, BAI D, DING W, An explicit routing optimization algorithm for internet traffic engineering[A]. ICCT 2003[C]. Beijing, China, 2003, 1(4): 186—192.
[9] Toyoo Fukuda, Kazuyudi Mori. Makoto Tsukiyama. Dasgupta Dipankar. Artificial Immune Systems and Their Application[C]. New York: Springer Verlag Berlin Heidelberg, 1999, 210— 220.
[10] X.H. Shi, Y.C. Liang, H.P. Leeb, C. Lu, L.M. Wang. An improved GA and a novel PSO-GA-based hybrid algorithm[C]. Information Processing Letters 93 (2005) 255–261.
[11] Emad Elbeltagi, Tarek Hegazy, Donald Grierson. Comparison among five evolutionary-based optimization algorithms[J]. Advanced Engineering Informatics 19 (2005) 43–53.
[12] Bo Liu, Ling Wang, Yi-Hui Jin, Fang Tang, De-Xian Huang. Improved particle swarm optimization combined with chaos[J]. Chaos, Solitons and Fractals 25 (2005) 1261–1271.