http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (17): 4153-4157, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.4153.4157

Application Study on Improved Grey Theory Model in the Pork Price Prediction

Du Zhi-Guo and Hu Da-Hui
Department of Information Management,
Southwest University, Rongchang, Chongqing, 402460, China

Abstract: The fluctuations in pork prices not only affect our daily life but also the development of social economy. Therefore, it is particularly important to reasonably predict the pork prices within a period in the future. The common price prediction models have different defects and application limitation. Summarizing advantages and disadvantages of each prediction model, this study, based on the grey theory, improves the unreasonable solution procedures of the original methods and proposes an improved grey theory model. And it makes use of simulation instruments to validate the model. The simulation experiments show that the error between the predicted pork price and the actual pork price is within 10% and the improved model can reasonably predict the pork price trend within a period in the future.

Key words: Grey theory, support vector machine, bp neural network, pork price prediction

INTRODUCTION

Animal husbandry is a vital part of agriculture and one of the main development directions of agriculture. The rational and orderly development of animal husbandry will actively promote the healthy progress of agriculture and lays a solid foundation for the rapid advancement of national economy. Pork is an important kind of animal product and residents' consumption of pork is much higher than that of other animal products in our daily life. In May 2012, the national average price of pork is 7.3% higher than that in May 2011 and 69% higher compared with that in May 2010. The rise of pork price will directly affect our life quality and influence the decision making of government. The national pork prices generally increased in 2007, directly promoting the rise of CPI (Consumer Price Index) and affecting the life quality of residents in the country.

Pork price has a huge impact on the national economy. The pork price is determined by the market which brings multiple influencing factors to it. The correct prediction of pork price is not only very important for farmers but also plays a significant role in the forecast of market economy. At present, some researchers have put forward some price prediction models, including combination prediction model, BPNN prediction model, time series analysis model and Support Vector Machine (SVM) model. These prediction models have different price prediction advantages and disadvantages. Analyzing the merits and demerits of the above-mentioned prediction models, this study presents a prediction method on a basis of the improved grey theory model and utilizes the grey theory model to simulate changes of prices to conduct the pork price prediction.

EXISTING PREDICTION MODELS

The theoretical basis of pork price prediction is the quantitative economic prediction model. Currently, some related models have been proposed. From the inspection of the actual conditions, each model should be improved.

VAR model: VAR (Vector Auto-Regression) model is the simplified form of VARMA model, whose application scope is limited because of its excessive parameters. The regression model requires every influencing factor of pork price to ensure its accuracy (Liu *et al*., 2001). The simplest regression model is the linear model, as shown in Eq. 1:

$$y_t = \beta_0+\beta_1 x_1+\varepsilon_t \quad (1)$$

where, β_0, β_1 are undetermined parameters and ε_t is the random error.

The biggest advantage of regression prediction model is that it can clearly reflect the effects of each external factor on predicted results. In consideration of uncertain factors that may influence the pork price, it is comparatively difficult to build a reasonable pork price prediction system with the regression model. VAR model can be represented by Eq. 2:

Corresponding Author: Du Zhi-Guo, Department of Information Management, Southwest University, Rongchang, Chongqing, 402460, China

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + B_1 x_t + \cdots B_r x_{t-r} + \varepsilon_t \quad (2)$$

where, $A_1 \ldots A_p$, $B_1 \ldots B_r$ are undetermined parameters and ε_t is the random error.

BPNN model: Neural network is a new type of information processing system imitating the functions and structures of biological nervous system, having the features of self-learning, self-organizing, self-adaption and fault tolerance.

Back Propagation Neural Network, (BPNN) is the most widely used neural network. It is also called as BP algorithm because of its application of error back propagation. Analyzing the input and output data, BPNN algorithm trains the BPNN according to the principle of error back propagation; meanwhile, it repeatedly applies the error back propagation and forward propagation processes to reduce errors and achieve the expected effects. The simplest BPNN at least consists of one input layer, one hidden layer and one output layer (Suykens and Vandewalle, 1999).

BPNN model algorithm is divided into two stages. First, the input information reaches the output layer through the hidden layer with each layer calculating its own output value. Then this process is called as the forward propagation stage. Second, the output error moves forward and calculates the error layer by layer, propagating back this error to revise the parameters of the former layer. This is called as the back propagation stage.

X_i is the input signal; M_j is the output of the hidden layer; Y_k is the output of the output node; Z_k is the expected output of the target signal; W_{ij} is the weight between the hidden node j and the output node k. N_i, N_y and N_0 respectively show the node number of input layer, the node number of hidden layer and the node number of output layer; f() is the transfer function; n is the sample number 1, 2, ...P; P is the total number of sample; α_j is the threshold value of the hidden layer node j; β_k is the threshold value of the output layer node and η is the learning step(Huang *et al*., 2012).

At the forward propagation stage, the output value of the hidden layer can be worked out by Eq. 3 and the output value of the output layer can be worked out by Eq. 4:

$$M_j^n = f\left(\sum_{i=1}^{N_j} W_{ij}^n x_i^n + \alpha_j\right) \quad (3)$$

$$Y_k^n = f\left(\sum_{j=1}^{N_y} W_{jk}^n M_j^n + \beta_k\right) \quad (4)$$

Substitute Eq. 3 into Eq. 4 and then the value of output layer can be obtained through Eq. 5:

$$Y_k^n = f\left[\sum_{j=1}^{N_0} W_{jk}^n f\left(\sum_{i=1}^{N_1} W_{ij}^n X_i^n + \alpha_j\right) + \beta_k\right] \quad (5)$$

At the back propagation stage, the error between the output value and the expected target can be obtained by Eq. 6. The definition of error function is shown in Eq. 7. Substitute Eq. 6 into Eq. 7 and then you can get Eq. 8:

$$e_k^n = Z_k^n - Y_k^n \quad (6)$$

$$E = \frac{1}{2}\sum_{\substack{n=1\\k=1}}^{PN_0} (e_k^n)^2 \quad (7)$$

$$E = \frac{1}{2}\sum_{\substack{n=1\\k=1}}^{PN_0} (Z_k^n - Y_k^n)^2 \quad (8)$$

In BPNN, the modified value of each layer is shown in Eq. 9, so the modified value of the next layer can be determined by Eq. 10.

$$\Delta W = -\eta \frac{\partial E}{\partial W} \quad (9)$$

$$W(k+1) = W(k) + \Delta W = W(k) - \eta \frac{\partial E}{\partial W} \quad (10)$$

The number of neuron is closely related to the number of the hidden layer. Obviously, when the number of hidden layer is large, the number of neuron will increase correspondingly. When there is only a hidden layer, the number of neuron can be determined by Eq. 11 and when there are two hidden layers, the number of neuron can be determined by Eq. 12; when the number of hidden layer is more than 3, the number of neuron will become a complex function based on n_1,n_2:

$$n_1 = \frac{(P-1)N_0}{N_i + 1 + N_0} \quad (11)$$

$$n_2 = \frac{(p-1)N_0 - (N_i + 1)n_1}{n_1 + 1 + N_0} \quad (12)$$

If the neural network model has only one hidden layer, the momentum gradient descent algorithm can be used to train the neural network. The calculation of the modified value is just shown in Eq. (13):

$$W_{ij}(n+1) = W_{ij}(n) + \eta((1-\alpha)D(n) + \alpha D(n-1)] \quad (13)$$

where, D(n), D(n-1) refer to the negative gradients of n, n-1.

SVM model: SVM (Support Vector Machine, SVM) model, first put forward by Cortes and Vapnik, is a machine learning method based on statistical learning theory, showing many unique advantages on the identification of small sample, nonlinear and high dimensional pattern identification.

The basic working principle of SVM is to use the nonlinear mapping method selected previously to map the input vector into the high-dimensional space and construct the optimal classification hyperplane in this high-dimensional space. Then make use of this hyperplane to conduct fitting or classification (Wu *et al.*, 2004).

If the training sample set is $\{(x_i, y_i), i = 1, 2,\ldots, n\}$, $x_i \in R^n$, $y_i \in R$, x_i, y_i, are corresponding target values, n is the number of sample and ε is the insensitive loss function, then SVM model can be represented by Eq. 14:

$$|y-f(x)|-\varepsilon=\begin{cases}0 & |y-f(x)|\le\varepsilon \\ |y-f(x)|-\varepsilon & |y-f(x)|>\varepsilon\end{cases} \quad (14)$$

The machine learning aims at constructing f(x) function and ensuring that the gap between it and the target value is less than ε.

IMPROVED GREY THEORY MODEL

Grey theory model: The grey theory has been widely used in the agricultural and industrial economic prediction since it emerged, especially in the conditions of uncertain data and the lack of mass data.

The working processes of grey theory can be divided into 4 steps (Deng, 2002).

Step 1: Given that the original data sequence $x_0 = \{x_0(1), x_0(2),\ldots, x_0(n)\}$, we conduct accumulated generating operation on x_0 and then get $x_0 = \{x_1(1), x_1(2),\ldots, x_1(n)\}$; where

$$x_1(j)=\sum_{i=1}^{j} x_0(i), j=1,2,\cdots n$$

Step 2: Construct sequence z according to sequence x_1, $z_1 = \{z_1(2), z_1(3), z_1(4)\ldots, z_1(n)\}$; where, $z_1(j) = \alpha x_1(j-1)+(1-\alpha)x_1(j)$ j = 2, 3,...n; generally, $\alpha = 0.5$

If x_1 follows the approximation exponential changes rule, then its shadow equation can be represented by Eq. 15:

$$\frac{dx_1}{dt}+ax_1=\mu \quad (15)$$

Equation 15 can be transformed into the grey differential equation of grey theory, just as shown in Eq. 16:

$$x_1(j)+\alpha z_1(j) = \mu \quad (16)$$

Step 3: Adopt the least square method to work out the parameters α and μ; where α is the development coefficient reflecting the growth rate of x_0 and μ is the grey action

Step 4: Establish the prediction equation

The prediction equation of x_1 is:

$$\wedge x_1(j+1)=[x_0(1)-\frac{\mu}{\partial}]e^{-\partial k}+\frac{\mu}{\partial}$$

where, j = 0, 1, 2, 3,…

The prediction equation of x_0 is $\wedge x_0(j+1\_ = x_1(j+)-x_1(j) = (1-e^{\partial})[x_0(1)-\mu/\partial]e^{-\partial k}$, where j = 1, 2, 3… and $\wedge x_0 = x_1$.

Improved grey theory model: Similar to other prediction models, the grey theory model also has certain defects. When solving the differential equation, it takes $\wedge x_1(j) = x_1(j)-x_0(1)$ as the known condition (Deng, 2004).

Solve the differential equation $dx_i/dt+\partial x_1 = \mu$ and obtain the Eq. 17:

$$x_1=\frac{\mu}{\partial}-\frac{c}{\partial}e^{-\partial t} \quad (17)$$

Discretize Eq. 17 and obtain the Eq. 18:

$$\wedge x_1(j+1)=-\frac{c}{\partial}e^{-\partial k}+\frac{\mu}{\partial} \quad (18)$$

It requires a condition to work out the constant. If $\wedge x_1(1) = x_1(1) = x_0(1)$, then obtain the Eq 19 and 20 according to Eq. 18:

$$\wedge x_1(1)=-\frac{c}{\partial}+\frac{\mu}{\partial}=x_0(1)-\frac{c}{\partial}=x_0(1)-\frac{\mu}{\partial} \quad (19)$$

$$\wedge x_1(j+1)=[x_0(1)-\frac{\mu}{\partial}]e^{-\partial k}+\frac{\mu}{\partial} \quad (20)$$

Because $x_1(1)$ is the oldest data, it is necessary to choose new data; for instance, deduce a new equation according to $\wedge x_1(m) = x_1(m)$. Substituting the new condition into the equation, we can obtain Eq. 21:

$$\begin{aligned}x_1(k+1)&=(x_1(m)+\frac{\mu}{\partial})e^{-\partial k+\partial(m-1)}\\ &+\frac{\mu}{\partial}=[x_1(m)-\frac{\mu}{\partial}]e^{-\partial(k-m+1)}+\frac{\mu}{\partial} \quad m=1,2,3\cdots,n\end{aligned} \quad (21)$$

Table 1: Pork prices in Chongqing city (2000-2012)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2000	6.93	7.43	7.64	6.58	5.62	5.58	5.87	6.48	6.44	6.68	6.67	7.11
2001	6.01	6.92	7.68	8.15	7.45	7.06	6.98	6.78	7.32	7.22	7.72	7.71
2002	7.96	7.79	8.34	7.82	7.36	6.83	6.29	6.81	7.27	7.52	8.04	7.71
2003	7.32	7.86	7.54	7.39	5.95	5.39	5.58	5.69	6.16	6.67	7.42	7.54
2004	7.71	7.68	8.60	9.13	8.65	8.67	9.94	11.66	13.30	12.32	11.79	12.15
2005	12.52	12.29	12.97	12.50	11.09	10.16	9.21	8.88	9.04	7.66	6.77	6.75
2006	7.56	7.70	7.22	5.53	4.62	4.46	4.89	6.04	6.89	6.70	7.19	8.14
2007	8.55	9.72	10.16	10.19	11.63	12.87	15.72	20.54	21.83	22.44	23.35	26.72
2008	30.26	30.92	32.10	34.22	33.70	30.55	28.25	27.88	27.96	24.08	19.23	18.30
2009	18.32	19.29	18.96	15.88	12.47	11.70	10.94	13.45	14.06	13.17	11.88	10.99
2010	11.08	11.20	10.77	10.11	9.94	9.33	9.66	11.51	12.54	12.19	11.79	12.30
2011	13.25	14.55	17.57	20.50	20.94	23.19	27.53	29.59	30.15	28.47	25.22	23.78
2012	23.44	24.31	25.75	25.50	23.60	22.39	20.91	19.87	20.94	20.64	19.12	18.93

Therefore, the improved grey theory prediction model can be roughly divided into five steps:

- Accumulate and generate
- Establish the model
- Solve ∂, μ
- Establish the prediction equation according to m = 1, 2,…n and calculate the error
- Select the m value of the minimum error to establish the best prediction equation through comparison

SIMULATION PREDICTION

Data collection: The pork price curve is an unstable dynamic nonlinear curve. It is necessary to select the sample data under normal circumstances to correctly predict pork prices; otherwise, the accuracy of prediction will decrease. Selecting the average pork prices in Chongqing city from January 2000 to December 2012, it takes MATLAB 2009a as the simulation instrument and improved grey theory as the algorithm model to predict the pork price trend of Chongqing city from January 2013 to June 2013. Table 1 is the pork market prices in Chongqing city from 2000 to 2012 (Unit: RMB/Kg).

Simulation steps: The establishment of grey theory model can be divided in to the following steps:

Step 1: Accumulate and generate the original data sequences x_0 and x_1

Step 2: The mean value generates x_1 and obtains z_1

Step 3: Obtain the differential model

Step 4: Use the data again and adopt the least square method to work out the estimated value α and μ

Step 5: Substitute the values of α and μ into the equation to work out the white differential equation

Step 6: Regenerate $x_0(k) = x_1(k+1)-x_1(k)$

The steps to adopt the above-mentioned model to establish the grey theory prediction model are as follows:

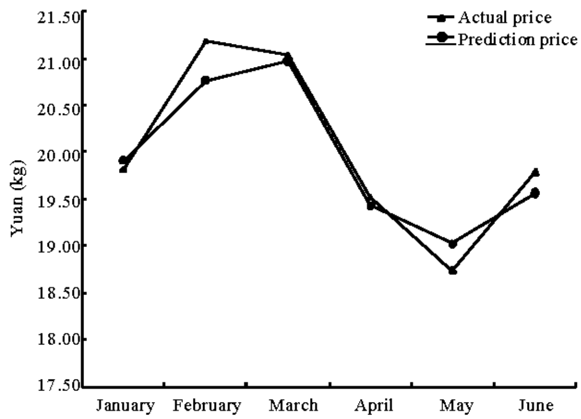

Fig. 1: Comparison between the predicted prices and the actual prices

Step 1: Validate the data

Step 2: Establish the GM (1, 1) model

Step 3: Improve the model

Step 4: Conduct the residual test

Step 6: Conduct the deviation value test

Step 6: Conduct the prediction forecast

Take advantage of MATLAB 2009a to simulate the above-mentioned algorithm. Use the historical data to check the model and finally predict the average pork price of Chongqing city from January 2013 to June 2013. The difference between the predicted price and the actual price is within 10% which indicates that the improved grey theory model can better conform to the actual condition. Fig. 1 is the comparison between the predicted prices and the actual prices

CONCLUSION

The changes of pork price have a huge impact on our daily life, so the correct prediction of pork price can help farmers to reduce risks and assist the government in making policies. Therefore, the improved grey theory model is applied to predict pork prices, taking advantage of the inner links of the historical pork price data to make the reasonable prediction. After many simulation experiments comparing the output pork price prediction results with the actual values, it proves that the prediction data of improved grey theory model are comparatively accurate. With the error between the prediction data and the actual data no more than 10%, it basically achieves the purpose of accurate prediction.

REFERENCES

Deng, J.L., 2002. Common operation and concept number in grey theory. J. Grey Syst., 6: 184-186.

Deng, J.L., 2004. The Primary Methods of Grey System Theory. Huazhong University of Science and Technology Publishers, China.

Huang, G.B., H. Zhou, X. Ding and R. Zhang, 2012. Extreme learning machine for regression and musicals classification. Trans. Syst., 42: 513-529.

Liu, W.Q. and X.Q. He, 2001. Applied Regression Analysis. China Renmin University Press, Beijin, China.

Suykens, J.A.K. and J. Vandewalle, 1999. Least squares support vector machine classifiers. Neural Process. Lett., 9: 293-300.

Wu, C.H., J.M. Ho and D.T. Lee, 2004. Travel-time prediction with support vector regression. Trans. Intell. Trans., 5: 276-281.