

Application of the Grey System in Commercial Housing Price Prediction

Zhuoshi Li^{1,a}, Yangyang Ye^{2,b*}, Ming Fu^{3,c}, and Fugui Zhu^{4,d}

^{1,2,3,4} Jilin Agricultural University, Changchun, Jilin, 130118, China

^aleezs643@sina.com, ^bls036@163.com, ^c499360221@qq.com, ^dleezs036@foxmail.com

Keywords: commercial housing price prediction; grey system; improved GM(1,1) model; Matlab

Abstract. In this paper, basing on the grey system theory with "partial information be known, partial information be unknown" of "small sample", "poor information" uncertainty system as researching object. From the basis of historical data to build models, the price of house and from the background value of improvement is given and the primary value of the two aspects from the selection of optimization of the improved model, using the residual value, poor development coefficient, posterior ratio and the small error is probable to the improved model.

Introduction

Commodity house price is closely related to people's lives, whether from the Angle of driving the development of national economy, or from the perspective of satisfy people's basic needs, the changes in prices of commercial housing and the development trend are crucial, mainly based on the grey system theory "part" of the known information generation and development, to extract valuable information, achieve a correct description of system behavior, evolution rule and effective monitoring. At present, many famous scholars in different regions and international organizations to be engaged in the research and application of grey system. The application range of the grey system theory has expanded to many fields such as industry, agriculture, social, economic, and successfully solved the scientific research and production of a large number of practical problems in life [1]. Prediction is based on the available information to a certain rule and method estimates the unknown things, development process and the results in order to understand things. Commercial housing price forecast is based on the existing house prices and consider the various factors affecting the development of housing prices, according to the scientific prediction method, measure a certain time of commercial housing prices in the future. Prediction of practical significance for commercial housing prices, the country can be based on the social and economic development planning. China's commodity house prices reflected the national economic level and the macro policy, may be possible in prediction results can indicate the economic development problems, is the government to formulate the basis of correct policies. To reasonably predict prices, this paper use greysystem theory, grey prediction GM(1,1) model was established based on the raw data first, with the help of Matlab software to solve the established model, according to house prices problem better optimization improvement GM(1,1) model is presented.

The Basic Theory of Grey System

Different information principle: "differences" are informat, all the informations will be different. The greater information content, There will be more information with the differences.

Definition 1 Let $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$, $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$,

$$x^{(0)}(k) + ax^{(1)}(k) = b \quad (1)$$

Definition 2 Let $X^{(0)}$ $X^{(1)}$ such as definition, as shown in the (1),

$$Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)\}, z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)), \quad k = 2, 3, \dots, n$$

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (2)$$

For the basic form of the GM(1,1) model. The parameter $-a$ for the development of coefficient, b grey action.

Theorem 1 Set a negative sequence, $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$, Among them $x^{(0)}(k) \geq 0, k = 1, 2, \dots, n$; $X^{(1)}$ is $X^{(0)}$ 1-AGO; $x^{(0)}(k) \geq 0, k = 1, 2, \dots, n$; $X^{(1)}$ is $X^{(0)}$ 1-AGO for the sequence $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$. Among them $x^{(1)}(k) = n \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n$; $X^{(1)}$ As close to the mean of generating, $Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)\}$,

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)), k = 2, 3, \dots, n \quad (3)$$

If $\hat{a} = (a, b)^T$,

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (4)$$

The GM(1,1) model of least squares estimate parameters,

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

$$\hat{a} = (B^T B)^{-1} B^T Y \quad (5)$$

Definition 3 Let $X^{(0)}$ to a negative sequence, $X^{(1)}$ for the sequence of $X^{(0)}$ generate $X^{(1)}$ 1-AGO sequence, $Z^{(1)}$ for close to the mean of,

$$(a, b)^T = (B^T B)^{-1} B^T Y$$

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (6)$$

$X^{(1)}$ for the sequence of $X^{(0)}$ generate $X^{(1)}$ 1-AGO sequence, $Z^{(1)}$ for close to the mean, Says $(a, b)^T = (B^T B)^{-1} B^T Y$ As the model equation of bleaching be also called shadow equation.

Theorem 2 Let B, Y, \hat{a} as described above theorem 1, $\hat{a} = (a, b)^T = (B^T B)^{-1} B^T Y$,

- 1) An albino equation solution into time response function is also An albino equation solution into the time response function is braking.

$$2) \quad \frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (7)$$

- 3) Model $x^{(0)}(k) + az^{(1)}(k) = b$ of time sequence for accordingly

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a} \quad k = 1, 2, \dots, n \quad (8)$$

- 4) Reduced value

$$\hat{x}^{(0)}(k+1) = x^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^a)(x^{(0)}(1) - \frac{b}{a})e^{-ak}, k = 1, 2, \dots, n \quad (9)$$

Commercial Housing Price Forecast Model Based on Improved Model Prediction

Grey prediction model is the core and basis of the model. The original modeling data has a great influence on the model prediction accuracy, the prediction precision and the fitting precision depends on the initial data has good smoothness. Accuracy of accurate or not has nothing to do with the merits of the model, is due to external disturbance and system behavior data distortion. The current priority is to exclude system behavior data of external disturbance, weaken the random error, restore the original data, improve the fitting and forecasting precision.

Improved model from the background value of improvement and the two aspects of the selection of the initial value optimization, improve the fitting precision and prediction precision.

Improve Commercial House Price Forecast Model is Constructed

Table 1 The Raw Data

k	1	2	3	4	5
$x^{(0)}(k)$	4.68	5.03	5.36	5.79	6.24

Step 1: According to the above theorem background values:

$$z^{(1)}(k) = x^{(1)}(k) + \frac{x^{(0)}(k)}{\ln x^{(0)}(k) - \ln x^{(0)}(k-1)} - \frac{[x^{(0)}(k)]^2}{x^{(0)}(k) - x^{(0)}(k-1)}$$

Step 2: Make sure the data matrix B , y_N

$$B = \begin{bmatrix} -7.186 & 1 \\ -12.376 & 1 \\ -17.927 & 1 \\ -23.994 & 1 \end{bmatrix} \quad y_N = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \end{pmatrix} = \begin{pmatrix} 5.03 \\ 5.36 \\ 5.79 \\ 6.24 \end{pmatrix}$$

Step 3: Calculate $(B^T B)^{-1} = \begin{pmatrix} 0.0064 & 0.0979 \\ 0.0979 & 1.7553 \end{pmatrix}$,

Step 4: Parameter columns $\hat{a} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, to least squares \hat{a} estimate of parameter are listed.

$$\hat{a} = (B^T B)^{-1} B^T y_N = \begin{pmatrix} -0.0830 \\ 4.5061 \end{pmatrix} \quad \alpha = -0.0830, \beta = 4.5061$$

Step 5: Determine the model. Time response is

$$\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = \beta \quad \frac{dx^{(1)}}{dt} - 0.0830x^{(1)} = 4.5061$$

$$\begin{cases} \hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{\beta}{\alpha})e^{-\alpha k} + \frac{\beta}{\alpha} \\ x^{(0)}(1) = 4.68, \quad \frac{\beta}{\alpha} = \frac{4.5061}{-0.0830} = -54.2904 \end{cases} \quad \hat{x}^{(1)}(k+1) = 58.9704e^{0.0789k} - 54.2904$$

Here and get the predicted $\hat{x}^{(1)}$ value, and then determine $k=1,2,\dots,n$ the value of $\hat{x}^{(1)}(2), \dots, \hat{x}^{(1)}(n)$ the reduction sequence can be obtained.

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) \quad (k=1,2,\dots,n)$$

Step 6: Accuracy test

Table 2 Restore Data

Generation Model calculations	TheActual Value	Restore Data	The Actual Value
$\hat{x}^{(1)}(2) = 9.719$	$x^{(1)}(2) = 9.71$	$\hat{x}^{(0)}(2) = 5.039$	$x^{(0)}(2) = 5.03$
$\hat{x}^{(1)}(3) = 15.102$	$x^{(1)}(3) = 15.07$	$\hat{x}^{(0)}(3) = 5.378$	$x^{(0)}(3) = 5.36$
$\hat{x}^{(1)}(4) = 20.954$	$x^{(1)}(4) = 20.86$	$\hat{x}^{(0)}(4) = 5.853$	$x^{(0)}(4) = 5.79$
$\hat{x}^{(1)}(5) = 27.193$	$x^{(1)}(5) = 27.10$	$\hat{x}^{(0)}(5) = 6.239$	$x^{(0)}(5) = 6.24$

From the Table 2 ,we can get the residual values,

Absolute residuals: $q(k) = [x^{(0)}(k) - \hat{x}^{(0)}(k)]$, $q(2) = -0.001$, $q(3) = -0.018$, $q(4) = -0.063$, $q(5) = 0.001$

Relative residuals : $e(2) = -0.0002$, $e(3) = -0.0034$, $e(4) = -0.0101$, $e(5) = 0.0002$

Discrete differential equations for solve:

$$\hat{x}^{(1)}(k+1) = 61.7421e^{0.0830k} - 54.2904$$

The improved GM(1,1) model as follows:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) \quad (k=1,2,\dots,n)$$

By above $-a = 0.0830 \leq 0.3$ and step 1 prediction accuracy is 98%, steps 2 and 5 over 97%,Improved GM(1,1) model available medium-term prediction.

Accuracy of inspection, the model can be obtained:

$$S_1 = \sqrt{\frac{1}{n} \sum_{k=1}^n (x^{(0)}(k) - \bar{x}^{(0)}(k))^2} = 0.5501 \quad , \quad S_2 = \sqrt{\frac{1}{n} \sum_{k=1}^n (q(k) - \bar{q}(k))^2} = 0.0245$$

$$C = \frac{S_2}{S_1} = 0.04454 \quad , \quad p = \{|q(k) - \bar{q}(k)| < 0.6745S_1\} = 1$$

When the effective after C the smaller the difference ratio and small error probability of the value, p the greater the model prediction accuracy will be more ideal.General requirements and not more than 0.65, biggest and not less than 0.70;Control inspection level "precision control standard table"known. $C = 0.04454 < 0.35$, $p = 1 > 0.95$, it is the highest accuracy of this model.

Acknowledgements

Jilin Province Innovation and Entrepreneurship Students Training Program Project.

Jilin Agricultural University Undergraduate Science and Technology Innovation Fund Project.

References

- [1]Li,H.B,Shuai.B,Application of grey-combinations of linear regression model in the prediction.Journal of shanxi institute of technology.19(2003) 64-68 .
- [2]Xie.N.M,The grey system model research. Nanjing university of aeronautics and astronautics.(2008)9-17.
- [3]Bao.Y.K,Li.N,Matlab Mathematical Statistics and Data Processing,Northeastern University Press,Beijing,2008.
- [4]Song,Y.C,The problems existing in the real estate industry in our country and the development countermeasures.Gansu Agricultura.l8(2010)37-40.