## Prediction model of stock market based on Chaos Theory

Oiao Wenhan<sup>1.,</sup>

<sup>1</sup>Jiangxi University of Finance and Economics, Nanchang, 330013

**Abstract:** After more than 30 years of development, China's stock exchange market has already had a considerable scale. Modeling and forecasting stock prices is always a problem. Based on the data of three stocks given in the title, this paper analyzes the characteristics and development trend of each stock. Based on the foothold of chaos theory<sup>[1]</sup> this paper studies the nonlinear dynamic characteristics and predictability of China's stock market, and finally tests it.

### 1 Introduction

### 1.1 Background

For a listed company, the change of its stock price can directly reflect the operation status of the listed company and the recognition degree of the market. Stock modeling and forecasting is always a difficult problem. Because the stock has both trend factors and random factors. For the stock market, it is a typical nonlinear complex model, and the chaotic model<sup>[2]</sup> is a very effective theory, which has achieved certain theoretical and effective results in the fields of power and communication. This is very similar to the situation of stocks. We can establish the model of stock price<sup>[3]</sup> through chaos thought and theory, and use the data of three stocks to solve the following problems

### 2 Problem analysis

For Question 1: Analyse the daily, weekly and monthly trends of the three stocks, analyze the trend of the stock qualitatively and quantitatively (including chaos test), and give the relevant analysis results.

Aim at Question 2: Try to establish stock price by using appropriate method or model to evaluate the model in Question 1

Mathematical models of trends and randomness, and their predictions are evaluated using data from the last 20 days.

Aim at Question 3: Due to the reverse phenomenon of stock price trend, that is, excessive rise will fall, excessive fall will rise, showing some cyclical changes. Use the results of Question 2 to build a stock reversal judgment model and test it with data.

### 2.1 Analysis of question one

Question one requires that we first analyze what chaos

theory is. By consulting relevant data, a phenomenon between randomness and certainty is chaos. We analyze the change of wealth through related data given in the topic. By selecting related data, the chaotic system has a high sensitivity to predict things, which results in a huge response of the whole system. Therefore, the state of the later stock market is very sensitive to small changes in initial conditions. Therefore, it is necessary to choose the dominant stock input to maximize the benefits. The advantage of an advantageous stock is that it has a high profit margin and a low risk.By building a regression model, we can get the most profitable and least risky three stocks to get the best stock investment plan. For a long time, the stock market chaos system is unpredictable. We study whether the stock price can be predicted by determining the chaotic characteristics [4]of the stock price.

### 2.2 Analysis of question two

In Question 2, we find the correlation dimension and Lyapunov index on the basis of Question 1, and analyze the mechanical characteristics of chaos of stock. For the establishment of the model, we digest the trend of data from the machine to get the residual series of linear regression, and then calculate the correlation dimension. After determining the dimension of time delay, we can get the largest Layp.The UNOV exponent, based on the data given in the figure, obtains the correlation dimension of the data of the three stocks, and finally determines the dimension of the time delay.

### 2.3 Analysis of question three

Question 3 For the stock market, due to the inverse phenomenon, when the rise is obvious, it will fall sharply but when the fall is too large, it will rise. On the basis of Question 1 and 2, we build a model of flip-over transaction and cut the period of time. We will determine the time of the selected stocks as multiple Sundays, and

<sup>,</sup> qiaowenhan20000626@163.com

then judge the profitability level and stability of their side rates in different periods, and determine The return of the stock is tested using the data obtained in question 2. If it has a good reversal in a short period of time, the strategy is correct and the strategy is used to adjust the prediction analysis. Its effective test on the method yields the final results and the test results.

### 3 Symbol and Assumptions

### 3.1 Symbol Description

Table 1: symbol Description

Name	Symbol
Chaotic time series Time delay coefficient Crosscorrelationinformation coefficient	$\begin{cases} x \end{cases}^{N} \\ n = 1 \\ \tau$
Reference phase point	$I( au) \ X(t_i)$
Number of delays Embedding dimension	K m
Distance of track  Detail signal	$L_{o}$ $D_{i}$
Time series of original income Cycle time	X(t) $T$

### 3.2 Fundamental assumptions

- 1. Assume that the data given in the question is true and valid
- 2. Assume there are no closing days except Saturday and legal holidays
- 3. Assume that three stocks do not fluctuate due to force majeure
- 4. Assume the future is influenced by the current market
- 5. Assume that stock prices are correlated with volume over time

### 4 Model

# 4.1 Model establishment and solution of problem one

Chaotic time series prediction is a new nonlinear prediction<sup>[5]</sup> method developed in the late 1980s. The trajectory of the state point also tends to stay in the attractor to maintain its fractal characteristics. To eliminate this trend, the relationship between the critical state of the state point and the subsequent points can be used as the basic model for prediction, and then the prediction of chaotic time <sup>[6]</sup>series can be formed.

The observed chaotic time series is set as  $(x)^N$ 

, assuming N = T + L The first Tnumber is taken as the sample needed in the follow-up, and the state vector of reconstruction in phase space is obtained by delay reconstruction method.

$$x_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau})$$
  $n = T_0 + 1, T_0 + 2, \dots, T$  (1)

For delay time  $\tau$  In terms of the time series method, the time series method is more suitable for the analysis of the

given time series<sup>[7]</sup>.

$$I(\tau) = \sum_{n=1}^{N} P(x_n, x_{n+\tau}) \log_2 \left| \frac{P(x_n, x_{n+\tau})}{P(x_n) P(x_{n+\tau})} \right|$$
(2)

Among among  $I(\tau)$  The information that can be obtained from a sequence is pointed out. When it reaches the first minimum point, the corresponding action  $\tau$  Embedding time delay.

For embedded dimensions m Where dimension m become m+1 . Investigation of the trajectory  $X_n$  What are the real, the real and the false.

$$\left\| X_{\eta(n)} - X_n \right\|_{\mathfrak{m}} = \min_{\substack{j = J_0, \dots, N, j \\ \neq n}} \left\| X_j - X_n \right\|$$
(3)

Let the reference phase point be  $X(t_i)$  After K times of delay, it becomes  $X_i$  k As long as k t, then phase point  $X_i$  k Only one of the components in is unknown, the rest are known.

$$\frac{\left\|X_{\eta(n)} - X_{n}\right\|_{m+1} - \left\|X_{\eta(n)} - X_{n}\right\|_{m}}{\left\|X_{\eta(n)} - X_{n}\right\|_{m}}$$
(4)

At this point,  $X_{\eta(n)}$  by  $X_n$  False nearest neighbor.

$$C_{m}(r) = \frac{1}{k} \sum_{i,j} H\{r - \|X_{i} - X_{j}\|\}$$
(5)

Correlation dimension:

$$D_c = -\lim \left[ \ln \frac{C_m(r)}{r \to 0} / \ln r \right]$$
(6)

The specific steps of the algorithm are as follows:

- 1) Using time series  $x_1, x_2, x_n = 1, x_n$  Firstly, a smaller embedding dimension m is given
  - 2) The correlation function is obtained
- 3) For an appropriate range of R, the linear relationship<sup>[8]</sup> between the dimension D and the

cumulative integral distribution function is satisfied, and then the estimated correlation dimension of M is

4) Increase the embedding dimension m and repeat the calculation to get the corresponding dimension estimation until D (m) converges.

### Model establishment and solution of problem two

According to the theory of chaos dynamics, the distance between the initial two phase points in the evolution process of chaotic system is assumed  $L_0$  After k-step iteration, the maximum distance between trajectories is set as follows:

$$L_k = L_0 e^{\lambda_1 k} \tag{7}$$

Suppose  $\frac{L_k}{L}$  When the critical value C is exceeded, the two initial points of the chaotic system are considered to diverge after iteration.

$$k = \frac{1}{\lambda_1} \log c \tag{8}$$

When the distance between the two initial phases is n times, we can get the maximum distance between them.

$$k_0 \sim \frac{1}{\lambda_1} \tag{9}$$

For lyqpunov index  $\lambda_1$ . The reciprocal of T = 1/ $\lambda_1$ . Usually represents the upper limit of predictable time.

Step1:Initialize the time series data and determine the appropriate sample space n; Step2:Calculate and select time series delay  $\tau$ .

Step3:In the proposed range, the initial value of interval is selected as the space dimension m

Step4:Based on the initialization of time series set n, the Lyapunov exponent is calculated

 $\lambda$  And the results were broken.

Step5:The embedding dimension m and time delay of reconstruction are calculated by using the determined time series  $\tau$  Reconstruction of phase space

Step6:The last known phase space point is calculated  $X(N-(m-1)\tau)$ . Nearest point is X(Min-dist)

Step7:Calculating phase space points  $X(Min\_dist)$  . And phase space point  $X(N-(m-1)\tau)$ 

Linear distance of<sup>[9]</sup>

Step8: Calculating phase space points

$$X(Min_dist+1)$$
 and phase space point

 $X(N-(m-1)\tau)$  distance of.

Step9:Finding phase space point X(Min \_ dist+1) In the reconstructed phase space, the Lyapunov exponent is used to predict the value when x is n + 1.

Step10:In the fourth step, we can get more accurate results by embedding in the M space

Finally, the predicted value is compared with the actual situation.

### Model establishment and solution of problem three

In question 3, we introduce the wavelet transform theory to analyze the prediction results in question 2, and then establish the stock turnover judgment model and improve the prediction method. Because of the temporality of the fluctuation, the fluctuation speed is different in different time. Therefore, it has the characteristics of high frequency and ground frequency, which can be captured by wavelet transform. Using the algorithm of Xiaobian, the daily yield series is decomposed into four levels, which are divided into floor part  $D_1$  And the high frequency part  $D_2$ ,  $D_3$ ,  $D_4$  On the basis of this decomposition, wavelet reconstruction is used to recover the data.

 $X = \{X(1), X(2), , X(k)\}$  It is the most original time series. After four levels of wavelet analysis, we can get the following results:

$$X = \sum_{i} D_i + A_4 \tag{10}$$

 ${d_i(t)}(i = 1, , 4; t = 1, , K)$  They among  $D_i$ are the decomposed signals.

$$X(t) = \sum_{i=1}^{5} d_i(t) + a_4(t)$$
(11)

Now known $\{t \ T\}$ All x (T) values of time are obtained by k-step prediction.

$$X(T+k) = \sum_{i=1}^{4} d_i(T+k) + \alpha_4(T+k)$$
(12)

We decompose the time series of correlation and get the prediction results.

#### 5 Sensitivity Analysis

For the principal component analysis, decision-making and other models, the factors analyzed have not been verified. Therefore, we carry out sensitivity analysis on the main influencing factors. The main influencing factors can be obtained by sensitivity analysis, and observe the influence on their comprehensive ability value when their values change. The change rate was calculated and compared. The time delays corresponding to the three stocks are normalized as follows:After normalization, the interval of the three factors is set in between, so that the three stocks can be numerically substituted into the equation, corresponding to the different comprehensive capacity values of the companies of the three stocks, as shown in the following figure:

With the time delay of the chaos theory of the three stocks, the corresponding comprehensive ability value is also changing constantly. As can be seen from the above figure, when the range changes, the comprehensive capacity value changes the most, and the increase of the stock market value has the smallest change. Then the factors explored in this model are reasonable<sup>[10]</sup>.

### 6 Conclusion

Aiming at the first problem, the nonlinear motion of stock market and its application to chaotic model are clarified through the behavior of investors. By analyzing the relevant data, we get the correlation coefficient of chaos theory and confirm that the three stocks are chaotic. And the time delay is obtained, and the change trend of these three stocks is calculated every day, every week and every month.

For the second question, we establish the rank sum ratio comprehensive evaluation model. In the first question, we determine the chaotic characteristics of the stock, and then we use the largest Lyapunov index to predict, where we get the time delay  $\tau$  And embedding dimension m.The error range of prediction is only 0-4%~3% between.

In view of the third question, we analyze and establish the stock trading volatility model through the wavelet change theory. Through the analysis, we aim at the profit level and stability of the same stock in different periods, and take the maximum return of the whole cycle as the most objective, and use the evaluation model of the second question to judge, Finally, the correctness of the model is obtained by sensitivity analysis.

### 7 Strengths and Weakness

### 7.1 Strengths

- 1) The rank sum ratio is a new statistic and a carrier of composite information, with large capacity, strong plasticity, strong pertinence, simple operation and high application value.
- 2) The comprehensive evaluation method of rank-tosum ratio is an innovation in quantitative methods and has extremely broad development prospects.
- 3) The results obtained by the problem three model are very intuitive and accurate for short-term prediction results.
- 4) In the third question, the problem is divided into multiple cycles and combined with reality to make the model have practical significance

### 7.2 Weakness

1) For the model of problem two, the main basis of sorting is to use the rank of the original data, and the final RSR value reflects the difference between the comprehensive rank and the order of the original data

irrelevant.

- 2) For the model of question three, the prediction based on the exponential rate does not consider the randomness of the system, and it is impossible to predict long-term data.
- 3) The less data used in the model makes the objectivity of the model not strong enough

### 7.3 Improvement

- 1) Improve the model of problem 1, adopt the weighted global least squares estimation criterion (WTLS) to estimate the parameters to be estimated, and finally, for the newly added observation information, derive the parameter estimation correction recursive algorithm to avoid complicated repeated calculations
- 2) For chaos theory, you can refer to a large amount of stock data, and then build a BP neural network to predict stocks on this basis

### References

- 1. Yuan Bo. Application research of Chaos Theory in stock Market Trend Prediction [D]. Northeast Forestry University.
- Pu Yongjian, Yang Xiutai. Chaos model of Economic fluctuation in China [J]. Journal of Chongqing University (Social Science edition), 1995, 001(002):24-29.
- Liu Daowen, Fan Zhizhi. Modeling and Prediction of Stock Price Index based on Support Vector Machine [J]. Statistics and Decision-making, 2013, 000(002):76-78.
- 4. Wei Junda. Research and Application of Chaotic Time Series based on Neural Network [D].
- Xun Sun. Research on short-term Stock Market Predictability and Prediction Methods based on Chaos Theory [D]. Jiangsu University, 2009.
- 6. Dong Yihao. Research on chaotic modeling and prediction based on Chinese stock index series [D]
- Zeng Wu-xu, Qian Wen-bin, WANG Ying-long, et al. A Stock prediction method based on Python and BP Neural Network [J]. Computer Age, 2018, 000(006):72-75,80.
- 8. Han Wenlei. Nonlinear Analysis and Prediction of China's Stock Market [D]. Northwestern Polytechnical University, 2006.
- Nonlinear Analysis and Forecast of Chinese Stock Market Han Wenlei
- Li Hongquan. Nonlinear Dynamic characteristics of capital market and Risk Management research [D]. Hunan University, 200