Prediction of stock price movement based on Markov chain prediction model

Qiong Qin[1], *Fang Zhang*[2]
1.Dalian Naval Academy, Dalian 116018, China
2.Dalian University of Science and Technology, Dalian 116036, China

Abstract: Based on the closing price of Ping An Bank (2023/6/7-2023/7/28) 38 trading days, this paper uses Markov chain prediction model to forecast the stock price trend of Ping An Bank in early August. By dividing the forecast objects into three states (up, flat and down), each state is independent of each other, using the Markov prediction method to forecast, using the initial state probability distribution and state transfer probability matrix to predict Ping An Bank in the 39th trading day with a probability of 0.4118 to show an upward state, with a probability of 0.5294 to show a decline state; On the 40th trading day, Ping An Bank showed a rising state with a probability of 0.451, and a falling state with a probability of 0.475.

Key words: Markov chain, stock price trend prediction, probability transfer matrix

The prediction of the stock market has been widely concerned. For investors, they all hope to seek the maximization of benefits in the stock market. However, stock is affected by the complex factors of market behavior, and has the characteristics of high noise and non-stationary nonlinear, which increase the challenge of stock forecasting. Based on the previous stock data, investors can estimate the stock price trend in the future time, which is more convenient for investors to make reasonable investment decisions according to the forecast data. In this paper, the Markov chain model is used to predict the stock price trend. A Markov chain model is a way of representing real-world random systems and processes that are modeled as a series of states and that shift between states with specific probabilities over time. Transitions between states are also conditional, depending on the state they were in before the transition occurred.

1. The basic theory of Markov chain

[Definition 1] A random process $X_T = \left\{X_t, t \in T\right\}$ is provided, where time $T = \left\{0,1,2,\cdots\right\}$,state space is $I = \left\{1,2,\cdots\right\}$,if for any positive integer k , any $t_i \in T$, $t_i < t_{i+1}$ $(i = 0,1,2,\cdots,k+1)$, and any positive integer $i_0, i_1, \cdots, i_{k+1}$,there is

$$P\left\{X_{t_{k+1}} = i_{k+1} \middle| X_{t_0} = i_0, X_{t_1} = i_1, \cdots, X_{t_k} = i_k\right\} = P\left\{X_{t_{k+1}} = i_{k+1} \middle| X_{t_k} = i_k\right\}$$

then X_T is called a discrete time Markov chain, referred to as Markov chain or Markov chain.

The change of state is a state transition, and the state transitions one step per unit time, so the conditional probability

$$P_{ij}^{(k)}(n) = P\left\{X_{n+k} = j \middle| X_n = i\right\}$$

It is called the transition probability that is in the state n after the step k transition of the system from the state i at the time, which is called the step j transition probability.

The transition probability $P_{ij}^{(k)}(n)$ is related to the state i and j , and the Markov chain independent of the moment n has a smooth transition probability.

A matrix of k step transition probabilities of Markov chains

$$P^{(k)} = \begin{pmatrix} p_{11}^{(k)} & p_{12}^{(k)} & p_{13}^{(k)} & \cdots \\ p_{21}^{(k)} & p_{22}^{(k)} & p_{23}^{(k)} & \cdots \\ \vdots & \vdots & \vdots & \\ p_{i1}^{(k)} & p_{i2}^{(k)} & p_{i3}^{(k)} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

It is called k step transition probability matrix.

The transition matrix has the following properties:

(1) $P_{ij}^{(k)} \geq 0, i, j \in I$ (2) $\sum_{j \in I} P_{ij}^{(k)} = 1, i \in I, k = 1,2,\cdots$

[Theorem] (C-K equation) For any positive numbers k, l and $i, j \in I$ have

$$P_{ij}^{(k+l)} = \sum_{r \in I} p_{ir}^{(k)} p_{rj}^{(l)}$$

The relation between n step transition probability matrix $P^{(n)}$ and step transition probability matrix P is $P^{(n)} = P^n$.

[Definition 2] Markov chain $X_T = \left\{X_n, n = 0,1,2,\cdots\right\}$, the initial time takes the probability of each state

$$P\left\{X_0 = i\right\} = p_i , i \in I$$

It is called the initial probability distribution of X_T, abbreviated as the initial distribution. Take the probability of each state at time n

$P\{X_n = i\} = p_i^{(n)}, i \in I$

It is called the absolute probability distribution at the moment n, or simply the absolute distribution.

For a finite number of state Spaces $I = \{1, 2, \cdots, N\}$ of Markov chains, there is

$$\left(p_1^{(n)}, p_2^{(n)}, \cdots, p_N^{(n)}\right) = \left(p_1, p_2, \cdots, p_N\right)\begin{pmatrix} p_{11}^{(n)} & p_{11}^{(n)} & \cdots & p_{1N}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & \cdots & p_{2N}^{(n)} \\ \vdots & \vdots & & \vdots \\ p_{N1}^{(n)} & p_{N2}^{(n)} & \cdots & p_{NN}^{(n)} \end{pmatrix}$$

The absolute probability distribution of the Markov chain at time n can be determined by the initial distribution and the one-step transition probability matrix P.

2. Markov prediction method

Markov method is the estimation of the transition probability matrix of the system state. In order to find out each transition probability, the idea of frequency approximation probability is generally used to calculate. Set in the M observation, the predicted object X is in the state i of the total number of times n_i, obviously,

$$M = \sum_{i=1}^{N} n_i$$

using frequency instead of probability can be obtained

$$p_i = P\{X = i\} = \frac{n_i}{M}, i = 1, 2, \cdots, N$$

Where $0 \le p_i \le 1$ and $\sum_{i=1}^{N} p_i = 1$, p_i can be used as a state probability estimate of the object X being in a state i.

If the predicted object X is in a state i, the next transition to a state j has occurred a total of times n_{ij}, obviously

$$n_i = \sum_{j=1}^{N} n_{ij}, i = 1, 2, \cdots, N$$

Where $0 \le p_{ij} \le 1$ and $\sum_{j=1}^{N} p_{ij} = 1$, p_{ij} can be used as a step transfer probability estimate of the phenomenon.

The basic steps of Markov chain prediction:

Step 1: Divide the state space and determine the state space $I = \{1, 2, \cdots, N\}$. The state of the prediction object can be directly divided according to the obvious state boundary value of the prediction object, or it can be artificially judged according to the actual situation. The purpose of prediction and the comprehensiveness of the state should be taken into account when dividing.

Step 2: According to the state interval divided in step 1, determine the state corresponding to the index value of each period in the data series.

Step 3: Statistical calculation according to the results obtained in step 2, the one-step transition probability matrix of Markov chain can be obtained, which determines the probability law of the state transition process of the index value. P

Step 4: Determine the initial distribution $\left(p_1, p_2, \cdots, p_N\right)$

and obtain $P(n) = \left(p_1, p_2, \cdots, p_N\right) \cdot P^n$

Step 5: Further discuss ergodicity, determine stationary distribution, and calculate the distribution law for long-term stability

3. The stock price trend forecast

In this paper, the closing price of Ping An Bank stock in 38 trading days from June 7, 2023 to July 28, 2023 is selected. The data are shown in Table 1. In the stock market, the daily closing price of individual stocks compared with the previous day can be divided into three situations: rising, flat and falling. According to the data in Table 1, the comparison between the daily closing price of Ping An stock and the previous day can be divided into three states, that is, the rise is recorded as 1, the flat is recorded as 2, and the fall is recorded as 3

Time	6/7	6/8	6/9	6/12	6/13	6/14	6/15	6/16	6/19	6/20
Closing price	11.94	12.12	11.88	11.79	11.77	11.47	11.61	11.63	11.45	11.36
Status	1	1	3	3	3	3	3	1	3	1
Time	6/21	6/22	6/23	6/26	6/27	6/28	6/29	6/30	7/3	7/4
Closing price	11.35	11.18	11.28	11.3	11.18	11.23	11.49	11.4	11.34	11.24
Status	3	3	1	1	3	1	1	3	3	3

Time	7/5	7/6	7/7	7/10	7/11	7/12	7/13	7/14	7/17	7/18
Closing price	11.2	11.2	11.21	11.24	11.43	11.46	11.31	11.25	11.31	11.32
status	3	2	1	1	1	1	3	3	1	1
Time	7/19	7/20	7/21	7/24	7/25	7/26	7/27	7/28		
Closing price	11.34	11.26	11.72	11.67	11.75	12.26	12.32	12.16		
Status	1	3	1	3	1	1	1	3		

Figure 1 Ping An stock's 38 trading days closing price (2023/6/7-7/28)

Markov chain prediction method is used to forecast, Markov prediction model is established, and the initial state probability distribution and state transfer probability matrix are used to predict the rise and fall of Ping An stock in early August. According to the status bar in Table 1, it can be calculated that the closing price compared with the previous day rose for 19 days, was flat for 1 day, and fell for 18 days. Since the state of the 38th trading day is "down", because there is no transfer on the last day, the number of falls is 17 times, the total number of rises is 19 times, the number of times from the rising state to the rising state (the transfer of state 1 to 1) is 9 times, the frequency is, the frequency replaces the probability; $\frac{9}{19}$ The number of times from the rising state to the flat state (the transfer of states 1 to 2) is 0; The number of times from the rising state to the falling state (transitions from states 1 to 3) is 9, the frequency is, and so on to get a one-step transition probability matrix is $\frac{9}{17}$

$$P=\begin{pmatrix} \frac{9}{18} & 0 & \frac{9}{18} \\ 1 & 0 & 0 \\ \frac{7}{17} & \frac{1}{17} & \frac{9}{17} \end{pmatrix}$$

According to the state data, since the 38th trading day is in the "down" state, the initial state distribution of Ping An Bank price is

$$\left(p_1, p_2, p_3\right)=(0,0,1)$$

The state probability distribution of predicting Ping An Bank price on the 39th trading day is

$$\left(p_1^{(1)}, p_2^{(1)}, p_3^{(1)}\right)=(0,0,1)\begin{pmatrix} \frac{9}{19} & 0 & \frac{9}{19} \\ 1 & 0 & 0 \\ \frac{7}{17} & \frac{1}{17} & \frac{9}{17} \end{pmatrix}=\left(0.4118, 0.0588, 0.5294\right)$$

According to the calculation results, Ping An Bank is in the state of rising with a probability of 0.4118, unchanged with a probability of 0.0588, and falling with a probability of 0.5294 on the 39th trading day. Through the comparison of the three probabilities, it is obvious that the probability of falling is greater.

The state probability distribution of Ping An Bank's price on the 40th trading day is further predicted as

$$\left(p_1^{(2)}, p_2^{(2)}, p_3^{(2)}\right)=\left(\frac{7}{17},\frac{1}{17},\frac{9}{17}\right)\begin{pmatrix} \frac{9}{19} & 0 & \frac{9}{19} \\ 1 & 0 & 0 \\ \frac{7}{17} & \frac{1}{17} & \frac{9}{17} \end{pmatrix}^2=\left(0.451, 0.028, 0.475\right)$$

According to the calculation results, Ping An Bank is in the state of rising with a probability of 0.451, unchanged with a probability of 0.028, and falling with a probability of 0.475 on the 40th trading day. Through the comparison of the three probabilities, it is obvious that the probability of falling is greater.

4. Conclusion

Using Markov chain theory, the paper forecasts the stock price trend of Ping An Bank and improves the performance of the trading strategy. However, the application of Markov chain model to stock trend prediction needs further improvement. In the next step, the closing price data can be divided into different price state ranges, which not only predicts the three states of "up", "flat" and "down" of the stock, but also further predicts the price range of the closing price of the stock, so as to facilitate investors to properly adjust their investment strategies. In order to obtain greater returns.

References:

[1] Xiaoyang Qu. Stock Price Prediction based on Markov Chain [J]. Times Finance, 2017(8):1.

[2] Jiayao Hu,Jie Luo. Prediction of Trend distribution of fund Heavy stock index based on Markov chain [J]. Chinese Science and Technology Journal Database (full-text Edition) Economic Management, 2022(1):3.