

# 基于多状态马尔可夫链的股票价格预测

付 军, 李俊刚\*, 庞静雯

北方工业大学理学院统计系, 北京

收稿日期: 2023年3月10日; 录用日期: 2023年4月1日; 发布日期: 2023年4月14日

## 摘 要

随着人们投资理财的意识的逐渐建立, 股票已成为常见的投资对象之一。但是在股票带来高收益的同时伴随着高风险。如果能提供较准确的股票参考预测, 对投资者以及企业自身都将得到可观的收益。由于影响股票市场的因素众多, 因此股票波动具有一定的随机性。本文围绕股票只受当前状态的影响、与过去状态无关的特点, 将运用广泛的马尔可夫链作为预测模型。以在线旅游板块中的西域旅游股票作为案例, 并对其进行运算处理, 建立四状态和多阶十六状态马尔可夫链模型, 预测其发展趋势, 得出马尔可夫链在股票的实际应用中具有一定研究价值的结论。

## 关键词

股票, 马尔可夫链, 预测模型, 在线旅游

# Prediction of Stock Price Based on Multi-State Markov Chain

Jun Fu, Jungang Li\*, Jingwen Pang

Department of Statistics, College of Science, North China University of Technology, Beijing

Received: Mar. 10<sup>th</sup>, 2023; accepted: Apr. 1<sup>st</sup>, 2023; published: Apr. 14<sup>th</sup>, 2023

## Abstract

With the gradual establishment of people's awareness of investment and financial management, stocks have become one of the common investment objects. But with high returns come high risks. If we can provide more accurate stock reference forecasts, we will get considerable benefits for investors and companies themselves. Because of the many factors that affect the stock market, stock fluctuations are somewhat stochastic in nature. In this paper, around the feature that stocks

\*通讯作者。

are only affected by the current state and have nothing to do with the past state, the widely used Markov chain is used as a prediction model. The Western Region Tourism stock in the online tourism sector is used as a case study and processed arithmetically to establish a four-state and multi-order sixteen-state Markov chain model to predict its development trend, and conclude that Markov chain has some research value in the practical application of stocks.

## Keywords

Stock, Markov Chain, Prediction Model, Online Travel

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

2022年11月,文化和旅游部发布的《关于推动在线旅游市场高质量发展的意见(征求意见稿)》(以下简称《意见》)。其中提到“积极发挥在线旅游行业在旅游业中的枢纽和引领作用,推动旅游行业的创新发展和智慧化水平”,“提升行业管理的数字化水平和治理能力,推动中国在线旅游行业发展处于国际领先地位”,这无疑给在线旅游行业注入了新的活力。

在线旅游,指消费者通过网络的方式购买和预订旅游产品或者旅游服务,并通过互联网或者线下完成支付。最早在1999年,中国便产生了在线旅游行业,但是在当时互联网普及较低,发展受到了阻碍。后来随着技术的进步,在经过长达20年的发展与更迭后,在线旅游与国民经济的关系越来越密切了,形成了相互依赖、相互促进的关系。

因此研究和预测该行业股票走势,能帮助我们了解在线旅游行业的发展趋势,可以给广大投资者提供投资决策参考,本文选用在线旅游板块中龙头股之一的西域旅游股票作为研究对象。西域旅游开发股份有限公司是新疆地区第一家旅游开发行业的上市公司,以2022年11月2日的股价估算市值超40亿元。

## 2. 文献综述

目前,国内外对于股票价格的预测有很多研究,但是股票市场复杂多变,且交易往往具备一定随机性。由于受到市场内外部宏观环境、投资者微观因素等诸多不确定的影响,无疑增大了预测的难度,因此在研究方法上一直不断进行改进、优化和创新,并致力于形成更加稳定有效的预测体系。

针对股票价格的预测,国内外学者提出了方差分析、回归分析、模糊算法、遗传算法、马尔可夫链等预测模型。其中马尔可夫链的作为深度学习的基石,且应用比较广泛。例如,张宇山等人[1]将股价状态划分成六个区间,利用马尔可夫链构建转移概率矩阵,预测宝钢股份未来20个交易日的走势;彭志行等人[2]将23个精伦电子股票收盘价分成四个区间,通过构建转移概率矩阵预测未来两个交易日的收盘价。王佳伟等人[3]则利用神经网络和马尔可夫链结合的方式,预测特发信息股票的价格。

基于以上文献,本文通过马尔可夫模型来研究西域旅游(300859)的股价涨跌变化情况,探究马尔可夫链在预测股票价格中的实际应用。我们收集整理了西域旅游股票2022年7月29日至2022年11月2日的收盘价格,按照涨幅共划分成四个区间,构建四状态转移概率矩阵和十六状态转移概率矩阵,预测股票未来的走势情况。

### 3. 马尔可夫链模型介绍

马尔可夫链，是以俄国数学家安德烈·马尔可夫(A.A. Markov, 1856~1922)命名的，是随机过程领域中的重要研究方法之一。如果  $n$  个连续变动事物在变动过程中，其中任一次变动的结果都具有无后效性，那么，这  $n$  个连续变动事物的集合就叫做马尔可夫链，表明事物的状态由过去到现在、由现在到将来，一环接一环，像一根链条[4]。

马尔可夫预测法是基于马尔可夫链，是依据当前事件的情况来预测其在未来各个时期变化情况的一种预测方法。总结历史数据，形成状态转移概率矩阵，对预测对象各个状态的初始分布和各状态间的转移概率进行研究，在已知当前股价涨跌情况时，便可以预测下一时刻股票的涨跌情况。

#### 3.1. 马尔可夫模型分析法

用马尔可夫链模型对事件进行预测的方法被称为马尔可夫模型分析法，这是一种将研究的变量作为状态，变化过程作为“状态转移”，形成动态模型来研究随机事件变化趋势的方法。

股票市场的股票价格、收盘价格、涨幅情况都是一种依赖于时间  $t$  的随机变化的过程，并且股票未来的状态只与当下时刻状态有关，与过去无关。股票状态变化的过程是具有时间历程的不变性的，并且其一步转移概率与时间起点没有关系，只与时间差有关系，即转移概率矩阵稳定不变。同时股票价格、股票价格区间以及成交量只能产生可列个状态，且用同一标准划分的各状态应相互独立，并包含全部可能出现的状况[5]。所以，可以用马尔可夫模型分析法对股票市场的价格波动、未来发展趋势进行预测，探究马尔可夫链在预测股票价格中的实际应用，帮助投资者了解市场基本情况。

假设马尔可夫过程  $\{X_n, n \in T\}$  的参数集是  $T$  离散的时间集合，即  $T = \{0, 1, 2, \dots\}$ 。与  $X_n$  对应的全体可能取值组成了状态空间  $I$ ，即  $I = \{i_1, i_2, i_3, \dots\}$ 。

马尔可夫链[6]则是指，在随机过程  $\{X_n, n \in T\}$  中，对于任意整数  $n \in T$ ，其条件概率满足马尔可夫链的后无效性，即：

$$P\{X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} = P\{X_{n+1} = i_{n+1} | X_n = i_n\} \quad (1)$$

换言之，时刻  $X_{n+1}$  的状态条件概率仅仅依赖于时刻  $X_n$  的状态，与其他时刻的状态无关。

#### 3.2. 转移概率与转移概率矩阵

状态转移概率是指事物当前正处于某一状态，其下一时刻转移到其他状态的可能性。我们可以借助条件概率的定义，求解转移概率  $p$ 。例如事物从状态  $E_u$  到状态  $E_v$  的状态转移概率  $P(E_u \rightarrow E_v)$  为：

$$P(E_u \rightarrow E_v) = P(E_v | E_u) = p_{uv} \quad (2)$$

如果事物拥有  $n$  个可能的状态，即  $E_1, E_2, E_3, \dots, E_n$ 。那么，我们把  $p_{uv}$  记为事物从状态  $E_u$  转移到  $E_v$  状态的转移概率，若事物此刻正处于状态  $E_u$ ，其下一刻可能会转移到  $E_i$  ( $i = 1, 2, 3, \dots, n$ ) 的状态，共  $n$  种可能。总结所有的状态转移概率后，可得到矩阵  $P$ ：

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \quad (3)$$

如果矩阵  $P$  满足下列条件：

$$\begin{cases} 0 \leq \forall p_{ij} \leq 1, (i, j = 1, 2, 3, \dots, n) & (4) \\ \sum_{j=1}^n p_{ij} = 1 (i = 1, 2, 3, \dots, n) & (5) \end{cases}$$

式(4)要求矩阵  $P$  满足非负性, 即矩阵中所有元素均不得为负, 且均不大于 1。

式(5)要求矩阵  $P$  满足归一性, 即矩阵所有行和始终为 1。

若矩阵  $P$  满足以上两点, 我们称为  $P$  为状态转移概率矩阵。

### 3.3. 马尔可夫预测法

事物如果要转移到某种特定状态时, 在实际过程中可能需要进行多步才能达到。本文用  $\pi_v(k)$  表示事件在初始状态( $k=0$ )已知的情况下, 经过  $k$  次状态转移后, 到达状态  $E_v$  的概率。其具备以下性质:

$$\sum_{v=1}^n \pi_v(k) = 1 \quad (6)$$

我们可以把上述过程分解成两步: 事物先经过  $k-1$  次的状态转移至  $E_i (i=1, 2, 3, \dots, n)$  的状态, 再经过 1 次状态转移成  $E_v$  的状态。根据马尔可夫过程的特性两步之间相互独立, 由 Bayes 条件概率公式其概率表示为:

$$\pi_v(k) = \sum_{i=1}^n \pi_i(k-1) \times p_{iv} \quad (7)$$

如果我们考虑经过  $k$  次状态转移后, 所有可能转移到状态的概率, 我们用行向量  $\pi(k)$  来表示:

$$\pi(k) = (\pi_1(k), \pi_2(k), \dots, \pi_n(k)) \quad (8)$$

当  $k=0$  时, 我们称为初始态概率向量

$$\pi(0) = (\pi_1(0), \pi_2(0), \dots, \pi_n(0)) \quad (9)$$

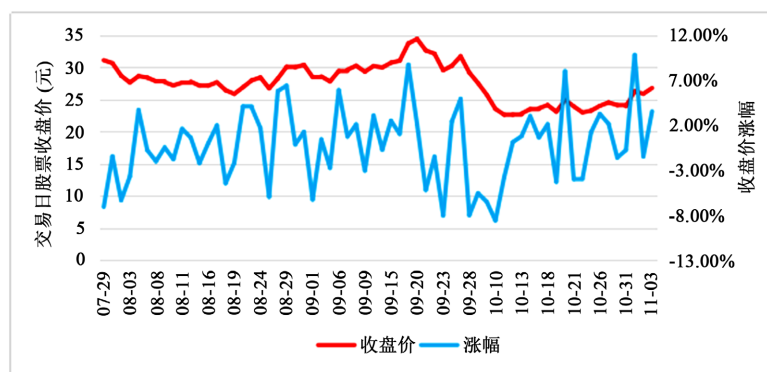
结合式(7), 我们可以推导出只用初始状态向量和状态转移矩阵, 表达  $\pi(k)$  的方法:

$$\begin{aligned} \pi(1) &= \pi(0) \times P \\ \pi(2) &= \pi(1) \times P = \pi(0) \times P^2 \\ \pi(3) &= \pi(2) \times P = \pi(1) \times P^2 = \pi(0) \times P^3 \\ \pi(4) &= \pi(3) \times P = \pi(2) \times P^2 = \pi(1) \times P^3 = \pi(0) \times P^4 \\ &\vdots \\ \pi(k) &= \pi(k-1) \times P = \pi(k-2) \times P^2 = \dots = \pi(0) \times P^k \end{aligned} \quad (10)$$

## 4. 在线旅游板块股票案例分析

选用在线旅游板块中的西域旅游(300859)股票作为本文研究对象, 收集该股自 2022 年 7 月 29 日至 2022 年 11 月 2 日的交易数据(详细数据见附表 1), 股票价格折线图如图 1 所示:

我们将西域旅游从 2022 年 7 月 29 日至 2022 年 11 月 2 日期间的每日收盘价与上一个交易日收盘价格进行比较, 把涨幅在  $(-\infty, -5\%]$  区间内的定义为大幅下降, 把涨幅在  $[5\%, +\infty)$  的股票定义为大幅上涨, 位于  $(0\%, 5\%)$  和  $(-5\%, 0\%)$  分别定义为上涨、下降。即总共分为大幅下降、下降、上涨、大幅上涨这四种情况, 将这四种情况分别作为马氏链模型中的四种状态, 利用马尔可夫链进行分析预测。



**Figure 1.** Stock price line chart of western region tourism from July 29, 2022 to November 2, 2022  
**图 1.** 西域旅游 2022 年 7 月 29 日~2022 年 11 月 2 日股价折线图

**Table 1.** Daily closing price status of western region tourism from July 29, 2022 to November 2, 2022  
**表 1.** 西域旅游 2022 年 7 月 29 日~2022 年 11 月 2 日每日收盘价状态

时间	收盘价	状态	时间	收盘价	状态	时间	收盘价	状态
2022-07-29, 五	31.08	大幅下降	2022-08-29, 一	30.04	大幅上涨	2022-09-28, 三	29.15	大幅下降
2022-08-01, 一	30.62	下降	2022-08-30, 二	29.99	下降	2022-09-29, 四	27.52	大幅下降
2022-08-02, 二	28.67	大幅下降	2022-08-31, 三	30.36	上涨	2022-09-30, 五	25.72	大幅下降
2022-08-03, 三	27.60	下降	2022-09-01, 四	28.45	大幅下降	2022-10-10, 一	23.50	大幅下降
2022-08-04, 四	28.61	上涨	2022-09-02, 五	28.56	上涨	2022-10-11, 二	22.60	下降
2022-08-05, 五	28.38	下降	2022-09-05, 一	27.77	下降	2022-10-12, 三	22.62	上涨
2022-08-08, 一	27.80	下降	2022-09-06, 二	29.40	大幅上涨	2022-10-13, 四	22.79	上涨
2022-08-09, 二	27.66	下降	2022-09-07, 三	29.61	上涨	2022-10-14, 五	23.47	上涨
2022-08-10, 三	27.16	下降	2022-09-08, 四	30.22	上涨	2022-10-17, 一	23.61	上涨
2022-08-11, 四	27.59	上涨	2022-09-09, 五	29.29	下降	2022-10-18, 二	24.10	上涨
2022-08-12, 五	27.75	上涨	2022-09-13, 二	30.18	上涨	2022-10-19, 三	23.06	下降
2022-08-15, 一	27.13	下降	2022-09-14, 三	29.95	下降	2022-10-20, 四	24.90	大幅上涨
2022-08-16, 二	27.12	下降	2022-09-15, 四	30.69	上涨	2022-10-21, 五	23.90	下降

Continued

2022-08-17, 三	27.65	上涨	2022-09-16, 五	31.00	上涨	2022-10-24, 一	22.95	下降
2022-08-18, 四	26.41	下降	2022-09-19, 一	33.70	大幅 上涨	2022-10-25, 二	23.23	上涨
2022-08-19, 五	25.83	下降	2022-09-20, 二	34.40	上涨	2022-10-26, 三	23.98	上涨
2022-08-22, 一	26.88	上涨	2022-09-21, 三	32.60	大幅 下降	2022-10-27, 四	24.49	上涨
2022-08-23, 二	27.94	上涨	2022-09-22, 四	32.10	下降	2022-10-28, 五	24.09	下降
2022-08-24, 三	28.40	上涨	2022-09-23, 五	29.51	大幅 下降	2022-10-31, 一	23.90	下降
2022-08-25, 四	26.70	大幅 下降	2022-09-26, 一	30.22	上涨	2022-11-01, 二	26.24	大幅 上涨
2022-08-26, 五	28.24	大幅 上涨	2022-09-27, 二	31.70	上涨	2022-11-02, 三	25.85	下降

#### 4.1. 划分状态确定状态空间

以每个收盘日作为离散的时间单位，收盘价变动情况分为四种状态：大幅下降、下降、上涨、大幅上涨。并把大幅下降记为 0(DD)、下降记为 1(ND)、上涨记为 2(NR)、大幅上涨记为 3(DR)。则状态空间为  $I_1 = \{DD, ND, NR, DR\}$ 。

状态概率是各种状态出现的可能性大小，用状态向量  $\mathbf{a}(n) = (p_0, p_1, p_2, p_3)$ ，其中  $n \in T$ ， $T$  是时间集， $T = \{0, 1, \dots\}$ ， $p_i$  为状态  $i$  出现的概率 ( $i \in I_1$ )。

表 1 中共 63 个交易日，其中大幅下降 10、下降 22、上涨 25、大幅上涨 6。

#### 4.2. 建立状态转移概率矩阵

设  $n_{uv}$  为  $X_1, X_2, X_3, \dots, X_n$  从状态  $u$  经过一步转移到状态  $v$  的频数，我们可以通过式(11)用来估计  $p_{uv}$ 。

$$\hat{p}_{uv} = \frac{n_{uv}}{\sum_{j=1}^n n_{ij}} \quad (11)$$

在表 1 中，大幅下降状态的总次数为 10 次，其中由大幅下降转移为大幅下降的次数是 3，故转移概率为： $p_{00} = 3/10 = 0.3$ ；由大幅下降转移为下降的次数是 4，故转移概率为： $p_{01} = 4/10 = 0.4$ ；由大幅下降转移为上涨的次数是 2，故转移概率为： $p_{03} = 2/10 = 0.2$ ；由大幅下降转移为大幅上涨的次数是 1，故转移概率为： $p_{04} = 1/10 = 0.1$ ；

下降状态的总次数为 22 次，其中由下降转移为大幅下降的次数是 2，故转移概率为： $p_{10} = 2/22 \approx 0.09$ ；由下降转移为下降的次数是 7，故转移概率为： $p_{11} = 7/22 \approx 0.32$ ；由下降转移为上涨的次数是 10，故转移概率为： $p_{12} = 10/22 \approx 0.45$ ；由下降转移为大幅上涨的次数是 3，故转移概率为： $p_{13} = 3/22 \approx 0.14$ ；

上涨状态的总次数为 25 次，其中由上涨转移为大幅下降的次数是 4，故转移概率为： $p_{20} = 4/25 = 0.16$ ；由上涨转移为下降的次数是 8，故转移概率为： $p_{21} = 8/25 = 0.32$ ；由上涨转移为上涨的次数是 12，故转移

概率为:  $p_{22} = 12/25 = 0.48$ ; 由上涨转移为大幅上涨的次数是 1, 故转移概率为:  $p_{23} = 1/25 = 0.04$ ;

大幅上涨状态的总次数为 6 次, 其中由大幅上涨转移为大幅下降的次数是 0, 故转移概率为:  $p_{10} = 0$ ; 由大幅上涨转移为下降的次数是 3, 故转移概率为:  $p_{11} = 3/6 = 0.5$ ; 由大幅上涨转移为上涨的次数是 2, 故转移概率为:  $p_{12} = 2/6 \approx 0.33$ ; 由大幅上涨转移为大幅上涨的次数是 1, 故转移概率为:  $p_{30} = 1/6 \approx 0.17$ , 将其各状态转移概率列成表 2。

**Table 2.** Table of state transition probability

**表 2.** 状态转移概率表

转向	收盘状态			
	大幅下降	下降	上涨	大幅上涨
大幅下降	0.30	0.40	0.20	0.10
下降	0.09	0.32	0.45	0.14
上涨	0.16	0.32	0.48	0.04
大幅上涨	0.00	0.50	0.33	0.17

表 2 中的所有元素组成了矩阵  $P$ , 且满足非负性和归一性, 因此  $P$  为所求的状态转移概率矩阵:

$$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} 0.30 & 0.40 & 0.20 & 0.10 \\ 0.09 & 0.32 & 0.45 & 0.14 \\ 0.16 & 0.32 & 0.48 & 0.04 \\ 0 & 0.50 & 0.33 & 0.17 \end{pmatrix}$$

### 4.3. 计算各收盘日状态概率

本文将不同时期的状态概率向量用  $\pi(n)$  表示, 根据式(10)计算不同时期的向量。由表 1 可知, 由于第 64 日收盘价处于下降状态, 而无后继资料[7], 所以可以认为初始状态向量  $\pi(0) = (0, 1, 0, 0)$ , 利用该向量和状态转移概率矩阵来预测以后各个收盘日价格状态概率。

第 64 日收盘价状态概率向量:

$$\pi(1) = \pi(0) \times P(0, 1, 0, 0) \begin{pmatrix} 0.30 & 0.40 & 0.20 & 0.10 \\ 0.09 & 0.32 & 0.45 & 0.14 \\ 0.16 & 0.32 & 0.48 & 0.04 \\ 0 & 0.50 & 0.33 & 0.17 \end{pmatrix} = (0.09, 0.32, 0.45, 0.14)$$

第 65 日收盘价状态概率向量:

$$\pi(2) = \pi(1) \times P \approx (0.1278, 0.3524, 0.4242, 0.0956)$$

第 66 日收盘价状态概率向量:

$$\pi(3) = \pi(2) \times P \approx (0.1379, 0.3474, 0.4193, 0.0953)$$

第 67 日收盘价状态概率向量:

$$\pi(4) = \pi(3) \times P \approx (0.1397, 0.3482, 0.4167, 0.0954)$$

第 68 日收盘价状态概率向量:

$$\pi(5) = \pi(4) \times P \approx (0.1399, 0.3483, 0.4161, 0.0956)$$



将以上结果变化为列表 3:

**Table 3.** The predicted probability of each closing day  
**表 3.** 各收盘日状态转移预测概率

各日状态	收盘日							
	概率	63	64	65	66	67	68	...
大幅下降	0	0.09	0.1278	0.1379	0.1397	0.1399	...	0.14
下降	1	0.32	0.3524	0.3474	0.3482	0.3483	...	0.35
上涨	0	0.45	0.4242	0.4193	0.4167	0.4161	...	0.42
大幅上涨	0	0.14	0.0956	0.0953	0.0954	0.0956	...	0.10

#### 4.4. 构造十六状态概率转移矩阵

从表 3 中计算值可以看出, 该股的收盘价状态的变化趋势: 随着交易日的增加, 即  $n$  足够大时, 只要状态转移概率矩阵不变(即稳定条件), 则状态概率会趋向和初始状态无关的值, 并稳定下来, 适合长期预测。即该股票在长期有 42%左右的可能性处于上涨状态, 35%的可能性处于下降状态, 14%的可能性处于大幅下降状态, 10%的可能性处于大幅上涨状态。预测的结果与实际情况基本一致。因此, 该股的前景还算不错。

马尔可夫链系统长期预测条件:

$$\begin{cases} w \times P = w \\ \sum_{i=1}^k p_i = 1, w = (p_1, p_2, \dots, p_k) \end{cases}$$

其中  $P$  为概率转移矩阵:

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}$$

我们将概率转移矩阵  $P$  代入式中, 即:

$$(p_1, p_2, p_3, p_4) \times \begin{pmatrix} 0.30 & 0.40 & 0.20 & 0.10 \\ 0.09 & 0.32 & 0.45 & 0.14 \\ 0.16 & 0.32 & 0.48 & 0.04 \\ 0 & 0.50 & 0.33 & 0.17 \end{pmatrix} = (p_1, p_2, p_3, p_4)$$

(其中,  $p_1 + p_2 + p_3 + p_4 = 1$ )。

求得:  $p_1 \approx 0.14, p_2 \approx 0.35, p_3 \approx 0.42, p_4 \approx 0.10$  可以看出, 由稳定状态下计算出的收盘价格状态概率值与递推公式推导的结论一致。我们利用数据来计算股票的期权价值, 并分析人们对于股票涨势的看法。

设昨日、今日连续两天大幅下降称为状态 1a (DDDD), 昨日下降、今日大幅下降称为状态 2a



(NDDD), 昨日上涨、今日大幅下降称为状态 $3a$  (NRDD), 昨日大幅上涨、今日大幅下降称为状态 $4a$  (DRDD);

设昨日大幅下降、今日下降称为状态 $1b$  (DDND), 设昨日下降、今日下降称为状态 $2b$  (NRND), 设昨日上涨、今日下降称为状态 $3b$  (NRND), 设昨日大幅上涨、今日下降称为状态 $4b$  (DRND);

设昨日大幅下降、今日上涨称为状态 $1c$  (DDNR), 设昨日下降、今日上涨称为状态 $2c$  (NDNR), 设昨日上涨、今日上涨称为状态 $3c$  (NRNR), 设昨日大幅上涨、今日上涨称为状态 $4c$  (DRNR);

设昨日大幅下降、今日大幅上涨称为状态 $1d$  (DDDR), 设昨日下降、今日大幅上涨称为状态 $2d$  (NDDR), 设昨日上涨、今日大幅上涨称为状态 $3d$  (NRDR), 设昨日大幅上涨、今日大幅上涨称为状态 $4d$  (DRDR)。

据表 1, 股票价格模型可看作十六状态的马尔可夫链。例如:

$$p_{1a2a} = P\{ND_{\text{今}} DD_{\text{明}} | DD_{\text{昨}} DD_{\text{今}}\} = \frac{1}{3} \approx 0.33$$

$$p_{1a2a} = P\{ND_{\text{今}} DD_{\text{明}} | DD_{\text{昨}} DD_{\text{今}}\} = \frac{1}{3} \approx 0.33$$

以此类推, 其转移概率矩阵  $P$  如下:

$$P = \begin{pmatrix} p_{1a1a} & p_{1a2a} & \cdots & p_{1a4d} \\ p_{2a1a} & p_{2a2a} & \cdots & p_{2a4d} \\ \vdots & \vdots & \ddots & \vdots \\ p_{4d1a} & p_{4d2a} & \cdots & p_{4d4d} \end{pmatrix} = \begin{pmatrix} 0.67 & 0 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0 & 0.25 & 0 & 0 & 0 & 0.25 & 0 & 0 & 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.29 & 0 & 0 & 0 & 0.57 & 0 & 0 & 0 & 0.14 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.25 & 0 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.11 & 0 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0.56 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.17 & 0 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0.42 & 0 & 0 & 0 & 0.08 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.67 & 0 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

其中, 第四行的概率总和为 0, 因为在表 1 中不存在昨日大幅上涨、今日大幅下降的情况。

#### 4.5. 多阶十六状态概率转移矩阵

二阶转移概率矩阵  $P^2$  为:

$$P^2 = P \times P = \begin{pmatrix} P_{1a1a} & P_{1a2a} & \cdots & P_{1a4d} \\ P_{2a1a} & P_{2a2a} & \cdots & P_{2a4d} \\ \vdots & \vdots & \ddots & \vdots \\ P_{4d1a} & P_{4d2a} & \cdots & P_{4d4d} \end{pmatrix} \times \begin{pmatrix} P_{1a1a} & P_{1a2a} & \cdots & P_{1a4d} \\ P_{2a1a} & P_{2a2a} & \cdots & P_{2a4d} \\ \vdots & \vdots & \ddots & \vdots \\ P_{4d1a} & P_{4d2a} & \cdots & P_{4d4d} \end{pmatrix}$$

$$= \begin{pmatrix} 0.449 & 0.165 & 0.000 & 0.000 & 0.221 & 0.000 & 0.000 & 0.000 & 0.000 & 0.165 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.250 & 0.000 & 0.000 & 0.000 & 0.000 & 0.250 & 0.000 & 0.000 & 0.250 & 0.250 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.168 & 0.125 & 0.000 & 0.000 & 0.083 & 0.000 & 0.125 & 0.000 & 0.000 & 0.125 & 0.125 & 0.000 & 0.000 & 0.000 & 0.000 & 0.250 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.055 & 0.000 & 0.250 & 0.000 & 0.165 & 0.000 & 0.250 & 0.000 & 0.280 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.063 & 0.000 & 0.000 & 0.084 & 0.188 & 0.000 & 0.000 & 0.165 & 0.319 & 0.046 & 0.000 & 0.041 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.028 & 0.000 & 0.000 & 0.145 & 0.083 & 0.168 & 0.000 & 0.285 & 0.140 & 0.083 & 0.000 & 0.070 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.055 & 0.000 & 0.000 & 0.145 & 0.165 & 0.000 & 0.000 & 0.285 & 0.280 & 0.000 & 0.000 & 0.070 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.085 & 0.000 & 0.000 & 0.250 & 0.165 & 0.000 & 0.000 & 0.125 & 0.210 & 0.000 & 0.000 & 0.125 & 0.000 & 0.000 \\ 0.028 & 0.000 & 0.095 & 0.000 & 0.028 & 0.165 & 0.185 & 0.000 & 0.028 & 0.083 & 0.235 & 0.000 & 0.028 & 0.083 & 0.045 & 0.000 \\ 0.043 & 0.000 & 0.071 & 0.000 & 0.043 & 0.165 & 0.139 & 0.000 & 0.043 & 0.083 & 0.176 & 0.080 & 0.043 & 0.083 & 0.034 & 0.000 \\ 0.125 & 0.000 & 0.085 & 0.000 & 0.125 & 0.000 & 0.165 & 0.000 & 0.125 & 0.000 & 0.210 & 0.000 & 0.125 & 0.000 & 0.040 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.165 & 0.000 & 0.000 & 0.335 & 0.000 & 0.000 & 0.000 & 0.335 & 0.165 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.050 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.050 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.050 & 0.000 & 0.000 & 0.000 & 0.050 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \end{pmatrix}$$

昨日、今日股票价格大幅下降，后天股票价格大幅下降的概率  $p_{1a1a}^2 = 0.449$ ；

昨日、今日股票价格大幅上涨，后天股票价格下降的概率  $p_{4d2a}^2 = 0.500$ ；

昨日、今日股票价格上涨，后天股票价格上涨的概率  $p_{3c3c}^2 = 0.176$ ；

昨日、今日股票价格下降，后天股票价格下降的概率  $p_{2b2b}^2 = 0.084$ 。

这样，我们以两天的股票状况对股票未来进行预测，增加了预测系统的稳定程度，在金融方面具有一定的实际意义。

转移概率矩阵  $P^3$  为：

$$P^3 = P^2 \times P = \begin{pmatrix} P_{1a1a} & P_{1a2a} & \cdots & P_{1a4d} \\ P_{2a1a} & P_{2a2a} & \cdots & P_{2a4d} \\ \vdots & \vdots & \ddots & \vdots \\ P_{4d1a} & P_{4d2a} & \cdots & P_{4d4d} \end{pmatrix}^2 \times \begin{pmatrix} P_{1a1a} & P_{1a2a} & \cdots & P_{1a4d} \\ P_{2a1a} & P_{2a2a} & \cdots & P_{2a4d} \\ \vdots & \vdots & \ddots & \vdots \\ P_{4d1a} & P_{4d2a} & \cdots & P_{4d4d} \end{pmatrix}$$

$$= \begin{pmatrix} 0.301 & 0.111 & 0.018 & 0.000 & 0.231 & 0.000 & 0.054 & 0.000 & 0.083 & 0.111 & 0.092 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.070 & 0.000 & 0.125 & 0.125 & 0.165 & 0.000 & 0.125 & 0.063 & 0.245 & 0.000 & 0.000 & 0.063 & 0.020 & 0.000 \\ 0.112 & 0.041 & 0.035 & 0.000 & 0.118 & 0.063 & 0.083 & 0.250 & 0.063 & 0.073 & 0.123 & 0.000 & 0.000 & 0.031 & 0.010 & 0.250 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.014 & 0.125 & 0.048 & 0.000 & 0.014 & 0.083 & 0.217 & 0.000 & 0.014 & 0.166 & 0.243 & 0.000 & 0.014 & 0.041 & 0.022 & 0.000 \\ 0.016 & 0.000 & 0.096 & 0.000 & 0.016 & 0.165 & 0.160 & 0.027 & 0.016 & 0.142 & 0.250 & 0.013 & 0.016 & 0.059 & 0.026 & 0.000 \\ 0.007 & 0.000 & 0.096 & 0.000 & 0.007 & 0.167 & 0.140 & 0.047 & 0.007 & 0.187 & 0.260 & 0.023 & 0.007 & 0.041 & 0.011 & 0.000 \\ 0.014 & 0.000 & 0.079 & 0.000 & 0.014 & 0.125 & 0.186 & 0.047 & 0.014 & 0.124 & 0.277 & 0.023 & 0.014 & 0.062 & 0.0022 & 0.000 \\ 0.021 & 0.000 & 0.049 & 0.000 & 0.021 & 0.155 & 0.111 & 0.084 & 0.021 & 0.184 & 0.158 & 0.081 & 0.021 & 0.076 & 0.017 & 0.000 \\ 0.042 & 0.014 & 0.049 & 0.000 & 0.033 & 0.140 & 0.119 & 0.055 & 0.024 & 0.154 & 0.259 & 0.072 & 0.024 & 0.069 & 0.019 & 0.028 \\ 0.046 & 0.021 & 0.079 & 0.000 & 0.032 & 0.117 & 0.107 & 0.055 & 0.018 & 0.150 & 0.182 & 0.061 & 0.018 & 0.058 & 0.014 & 0.043 \\ 0.105 & 0.063 & 0.036 & 0.000 & 0.063 & 0.083 & 0.132 & 0.000 & 0.021 & 0.104 & 0.151 & 0.040 & 0.021 & 0.041 & 0.017 & 0.125 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.500 & 0.000 & 0.000 & 0.000 & 0.500 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.041 & 0.000 & 0.065 & 0.000 & 0.041 & 0.097 & 0.165 & 0.000 & 0.041 & 0.191 & 0.257 & 0.000 & 0.041 & 0.047 & 0.013 & 0.000 \\ 0.125 & 0.000 & 0.085 & 0.000 & 0.125 & 0.000 & 0.165 & 0.000 & 0.125 & 0.000 & 0.210 & 0.000 & 0.125 & 0.000 & 0.040 & 0.000 \\ 0.000 & 0.000 & 0.055 & 0.000 & 0.000 & 0.145 & 0.000 & 0.000 & 0.000 & 0.285 & 0.280 & 0.000 & 0.000 & 0.070 & 0.000 & 0.000 \end{pmatrix}$$

## 5. 预测与检验

我们利用概率矩阵预测第 64 日(11 月 3 日)的股票情况, 已知第 62、63 日收盘价分别处于大幅上涨、下降的状态, 从十六状态转移矩阵  $P$  来看, 第 64 日处于下降和上涨的状态各 0.5, 没有明显的倾向; 我们从第 64 日收盘价状态概率向量  $\pi(1)$  来看, 收盘价转移到上涨状态的概率最大, 达到 0.45; 因此预测第 64 日股票收盘价涨幅在 (0%,5%) 区间范围内; 实际第 64 日收盘价涨幅为 3.37%, 位于区间范围内, 预测结果准确;

预测第 65 日(11 月 4 日)的股票情况, 从十六状态转移矩阵  $P^2$  和从第 65 日收盘价状态概率向量  $\pi(2)$  来看, 第 65 日位于上涨状态的概率最大, 即预测收盘价涨幅在 (0%,5%) 区间内; 实际第 65 日收盘价涨幅为 12.57%, 不位于区间范围内, 但预测和实际都在上涨;

预测第 66 日(11 月 7 日)的股票情况, 第 66 日收盘价状态概率向量  $\pi(3)$  中下降状态的概率较前一天有所增加, 再结合十六状态转移矩阵  $P^3$  进行分析, 第 66 日位于下降状态的概率会较大一些, 即预测涨幅在 (-5%,0%) 区间内; 实际第 66 日收盘价涨幅为 -1.40%, 位于区间范围内, 预测准确, 这也体现了随着状态的增多系统的稳定程度在增加。

## 6. 结论

本文利用马尔可夫模型对西域旅游股票少量的历史数据对股票收盘价格进行预测和分析, 结果表明预测与实际结果比较相符。但是仍有预测结果与实际股票波动不吻合的状况, 归根结底在于股票还受到市场环境、经济政策、汇率变化、政治等因素的影响。根据东方财富网的资料显示, 2022 年 11 月 4 日创业板指数上涨幅度超过 3%, 本文认为 11 月 4 日涨幅不在预测区间内的主要原因: 模型没有考虑到创业板股票当天市场交易环境因素的影响, 使得预测涨幅略微保守。

因此, 我们要综合考虑各种因素, 并随着股票收盘价的更新, 及时调整状态转移概率矩阵, 从而提高预测结果的可信度, 也为投资者提供参考依据, 为公司股价提供参考走势。

## 基金项目

本论文工作由北京市属高校基本科研业务费(No. 110052971921/103)和北京市教委基本科研业务费(No. KM202010009013), 大学生科学研究与创业项目(X202210009106)资助。

## 参考文献

- [1] 张宇山, 廖芹. 马尔可夫链在股市分析中的若干应用[J]. 华南理工大学学报(自然科学版), 2003, 31(7): 74-77.
- [2] 彭志行, 夏乐天. 马尔可夫链及其在股市分析中的应用[J]. 应用数学, 2004(S2): 159-163.
- [3] 王佳伟, 王志. 基于神经网络的马尔科夫链和遗传算法对股票价格预测研究[J]. 宁波工程学院学报, 2012, 24(3): 29-37.
- [4] 刘次华. 随机过程[M]. 第四版. 武汉: 华中科技大学出版社, 2008: 58.
- [5] 陈柯伊. 基于时间序列马尔科夫链的股市预测[J]. 数学学习与研究, 2014(23): 112.
- [6] 张波, 张景肖. 应用随机过程[M]. 北京: 清华大学出版社, 2004.
- [7] 王洲, 范爱华. 基于 Markov 模型对股票的价格预测[J]. 统计学与应用, 2022, 11(4): 778-784.  
<https://doi.org/10.12677/sa.2022.114081>

## 附录

Table A1. Stock data of western region tourism from July 29, 2022 to November 3, 2022

表 A1. 西域旅游 2022 年 7 月 29 日~2022 年 11 月 3 日股价数据

时间	开盘	最高	最低	收盘	涨幅	金额	成交次数
2022-07-29, 五	32.50	32.76	30.70	31.08	-7.11%	318,999,250	23,633
2022-08-01, 一	30.96	31.45	30.20	30.62	-1.48%	177,763,620	15,434
2022-08-02, 二	30.05	30.30	28.41	28.67	-6.37%	192,256,680	16,743
2022-08-03, 三	28.68	29.25	27.50	27.60	-3.73%	167,520,790	14,978
2022-08-04, 四	27.88	28.61	27.50	28.61	3.66%	168,529,020	15,163
2022-08-05, 五	28.15	28.40	27.60	28.38	-0.80%	126,976,004	10,868
2022-08-08, 一	27.20	27.99	27.10	27.80	-2.04%	131,779,771	10,223
2022-08-09, 二	27.51	28.29	27.31	27.66	-0.50%	113,439,568	9552
2022-08-10, 三	27.45	27.74	27.02	27.16	-1.81%	122,462,047	9983
2022-08-11, 四	27.20	27.88	27.03	27.59	1.58%	134,108,995	11,013
2022-08-12, 五	27.32	28.08	27.27	27.75	0.58%	157,900,090	12,819
2022-08-15, 一	27.30	27.50	26.33	27.13	-2.23%	126,926,553	9946
2022-08-16, 二	27.14	27.61	27.00	27.12	-0.04%	96,875,270	7923
2022-08-17, 三	27.13	27.88	26.78	27.65	1.95%	200,554,260	12,651
2022-08-18, 四	27.47	27.64	26.15	26.41	-4.48%	199,595,890	13,839
2022-08-19, 五	26.49	26.55	25.78	25.83	-2.20%	113,571,136	8451
2022-08-22, 一	25.84	27.33	25.55	26.88	4.07%	183,846,700	12,322
2022-08-23, 二	26.90	28.26	26.68	27.94	3.94%	233,093,430	16,360
2022-08-24, 三	27.60	29.46	27.55	28.40	1.65%	262,030,560	15,350
2022-08-25, 四	28.10	28.95	26.49	26.70	-5.99%	244,321,720	15,111
2022-08-26, 五	27.46	29.20	27.00	28.24	5.77%	273,859,550	15,621
2022-08-29, 一	27.78	30.48	27.46	30.04	6.37%	318,126,900	18,562
2022-08-30, 二	29.50	30.84	29.24	29.99	-0.17%	258,905,400	15,310
2022-08-31, 三	29.69	30.98	29.65	30.36	1.23%	301,580,330	15,455
2022-09-01, 四	30.12	30.78	28.37	28.45	-6.29%	268,464,030	15,558
2022-09-02, 五	27.74	28.56	27.12	28.56	0.39%	197,938,840	12,694
2022-09-05, 一	28.06	28.40	27.44	27.77	-2.77%	145,055,750	11,142
2022-09-06, 二	27.85	29.83	27.61	29.40	5.87%	257,677,210	15,675
2022-09-07, 三	29.00	30.24	28.90	29.61	0.71%	219,665,800	13,424
2022-09-08, 四	29.80	31.28	29.80	30.22	2.06%	329,878,770	19,753
2022-09-09, 五	29.13	29.33	28.36	29.29	-3.08%	240,000,940	14,847
2022-09-13, 二	29.44	30.18	28.90	30.18	3.04%	231,177,510	13,522

## Continued

2022-09-14, 三	29.56	30.66	29.42	29.95	-0.76%	195,083,390	11,531
2022-09-15, 四	30.26	30.87	29.60	30.69	2.47%	277,660,030	14,631
2022-09-16, 五	30.21	32.77	30.18	31.00	1.01%	417,616,460	21,392
2022-09-19, 一	30.10	34.27	30.10	33.70	8.71%	487,761,400	26,164
2022-09-20, 二	33.00	34.75	32.20	34.40	2.08%	456,949,700	25,571
2022-09-21, 三	33.50	34.22	32.50	32.60	-5.23%	334,911,940	19,454
2022-09-22, 四	32.54	32.98	31.55	32.10	-1.53%	273,429,060	16,441
2022-09-23, 五	31.95	31.95	29.51	29.51	-8.07%	307,600,030	18,567
2022-09-26, 一	29.59	31.35	29.59	30.22	2.41%	288,246,200	16,466
2022-09-27, 二	30.49	32.62	28.23	31.70	4.90%	411,458,140	25,478
2022-09-28, 三	30.74	31.20	29.05	29.15	-8.04%	281,848,520	18,375
2022-09-29, 四	29.49	29.78	27.50	27.52	-5.59%	227,211,810	17,255
2022-09-30, 五	27.67	28.14	25.68	25.72	-6.54%	187,634,580	14,777
2022-10-10, 一	25.31	25.65	23.50	23.50	-8.63%	162,961,230	14,618
2022-10-11, 二	23.70	23.72	22.41	22.60	-3.83%	114,752,593	13,289
2022-10-12, 三	22.70	22.77	21.45	22.62	0.09%	119,035,048	12,192
2022-10-13, 四	22.50	23.48	22.00	22.79	0.75%	141,318,030	13,421
2022-10-14, 五	23.15	23.59	22.99	23.47	2.98%	145,924,190	14,172
2022-10-17, 一	23.35	23.90	22.90	23.61	0.60%	129,947,509	12,681
2022-10-18, 二	24.26	25.73	23.92	24.10	2.08%	280,940,730	20,175
2022-10-19, 三	23.40	24.27	23.06	23.06	-4.32%	168,682,600	11,074
2022-10-20, 四	22.67	26.00	22.60	24.90	7.98%	282,987,400	18,285
2022-10-21, 五	24.18	24.50	23.70	23.90	-4.02%	170,894,460	12,593
2022-10-24, 一	23.52	24.00	22.94	22.95	-3.97%	119,037,294	9331
2022-10-25, 二	22.70	23.96	22.70	23.23	1.22%	140,515,780	10,009
2022-10-26, 三	23.50	24.82	23.50	23.98	3.23%	180,632,260	11,992
2022-10-27, 四	24.00	24.98	23.81	24.49	2.13%	179,495,860	12,455
2022-10-28, 五	24.40	25.15	23.77	24.09	-1.63%	177,340,950	14,352
2022-10-31, 一	23.03	23.94	23.02	23.90	-0.79%	142,179,520	8627
2022-11-01, 二	23.68	26.26	23.52	26.24	9.79%	277,228,500	16,196
2022-11-02, 三	25.55	27.39	25.55	25.85	-1.49%	322,643,820	17,319
2022-11-03, 四	25.60	27.40	25.40	26.72	3.37%	236,751,270	16,519