



REVIEW ARTICLE

CHINA'S ECONOMY BASED ON GREY FORECAST MODEL AND K-MEANS CLUSTERING ALGORITHM

Chan Wen, Ziqi Li, Xiaomin Zhou*

School of Economics and Management, Hunan University of Science and Engineering, Yongzhou 425100, China

*Corresponding Author E-mail: 18229445798@163.com

This is an open access article distributed under the Creative Commons Attribution License CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ARTICLE DETAILS

Article History:

Received 30 August 2023
Accepted 14 November 2023
Available online 13 December 2023

ABSTRACT

In recent years, there are more and more studies on financial asset allocation and portfolio. Managing the idle assets and maximizing the capital income have become an important research issue. In the common asset classes, equity assets, which refer to the assets related to stocks and corporate equity, have lower investment thresholds and larger investment returns than other investment assets. However, this investment method has high risks, making the asset allocation necessary. There is always one industry in the actual investment portfolio that will outperform others. In this dynamic investment, we need to build a model to predict the optimal investment portfolio. The core of the dynamic adjustment of the investment portfolio is to judge the future monetary and credit cycle. By building the grey forecast model and the K-means clustering model, we divided China's economy in 2001.01-2021.12 into four stages: tight money and tight credit, easy money and easy credit, tight money and easy credit, easy money and tight credit based on the theory and index data of the monetary and credit cycle, and predicted China's economy in 2022.10-2023.12. Finally, we put forward investment suggestions to help investors to optimize their investment portfolio and maximize their investment benefits.

KEYWORDS

K-means clustering algorithm, grey forecast model, economy, investment

1. INTRODUCTION

With the rapid development of the world economy, market traders are looking for the best investment way to make use of idle asset allocation, so as to maximize their investment benefits. Due to the instability of the market industry, it is necessary to build a model to predict the optimal investment portfolio. The core of the dynamic adjustment of the investment portfolio is to judge the future monetary and credit cycle. The risk-return characteristics of different industries are often affected by the monetary and credit cycle. Thus, only when we accurately predict the future monetary and credit cycle can we effectively determine the optimal industry and investment portfolio. Referring to the Basic Classification of China's New System of National Accounting of Xu (1999), we divided China's economy in 2001.01-2021.12 into four stages: tight money and tight credit, easy money and easy credit, tight money and easy credit, easy money and tight credit based on the theory and index data of the monetary and credit cycle. Referring to the Grey Forecast and Decision-Making of Deng (2002) and the Research on Cluster Analysis Algorithm and Its Effectiveness Evaluation of Li (2021), we built the grey forecast model and the K-means clustering model, so as to analyze China's economy in 2001.01-2021.12 and predict China's economy in 2022.10-2023.12. In the Asset Allocation Theory: Research Review and

Outlook of Zhou et al., (2022), it mentioned that the neoclassical finance establishes the theory and model of asset pricing and allocation through the hypothesis of rational man and optimization method. However, the financial market does show randomness, and the non-measurable theory in quantum theory is also applicable to the financial market. The unexplainable abnormal market price constantly urges scholars to expand the pricing factor and seek to build an asset pricing model for explanation. The asset allocation model based on asset pricing model has been developed to include macro factors, and the quantitative investment strategy based on financial anomaly has also become an important development direction in the field of asset allocation. In the Portfolio Selection: The Role of Extreme Loss Constraints of Jiang and Zhang (2022), it mentioned that it is of great theoretical value and practical significance to consider the possibility of extreme losses in the future in the investment decision-making process. Build and solve the asset portfolio selection model based on the extreme loss constraint, and conduct numerical and empirical analysis of the data, which expands the research on the asset portfolio selection based on the downside risk control. In the Revisiting Carl Marx's Theory of Money, Credit and Economic Cycle of Li and Wei (2021), it focused on the Role of money and credit in the modern capitalist market economy system. If investors want to obtain higher investment benefits, they should allocate their

Quick Response Code



Access this article online

Website
www.aiem.com.my

DOI:
10.7508/aiem.02.2024.84.90

investment reasonably, realize the diversification of investment methods, and optimize their investment portfolio. How to maximize their investment benefits while fully controlling the risks? Therefore, we used the grey forecast model and the K-means clustering algorithm to put forward investment suggestions to help investors to optimize their investment portfolio and maximize their investment benefits under the premise of economy forecast and classification.

2. METHODS

2.1 Grey Forecast Model

The basic idea of GM (1,1) is as follows. First, the accumulate technique is used to make the data have exponential rules, and then the first order differential equation is established and solved, and the inverse accumulating generation operation of the results is performed to obtain the predict, so as to predict the future. Step 1: Before building the grey forecast model, the stepwise ratio test of the known raw data must be performed to guarantee the feasibility of the modeling method. Let the initial non-negative data sequence be:

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

The model can be built only if all $\sigma(k)$ fall into the calculation range. The calculation and judgment formulas of the stepwise ratio are respectively:

$$\sigma(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, \sigma(k) \in \left(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}} \right) \quad (2)$$

Through the accumulation operation, the first-order accumulation sequence of $x^{(0)}$ is obtained, which can weaken the perturbation of $x^{(0)}$:

$$x_k^{(1)} = \sum_{i=1}^k x_i^{(0)}, k = 1, 2, \dots, n \quad (3)$$

$Z^{(1)}$ is the generated mean sequence of consecutive neighbors of $X^{(1)}$

$$Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)\} \quad (4)$$

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)) \quad (5)$$

Thus, the corresponding differential equation of GM (1,1) is obtained:

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (6)$$

Step 2: Build the data matrix B and the data vector Y, respectively

$$B = \begin{bmatrix} -z(2) \cdots 1 \\ -z(3) \cdots 1 \\ \vdots \\ -z(n) \cdots 1 \\ \dots \end{bmatrix} \quad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \quad (7)$$

Then the least squares estimation parameter column of the grey differential equation satisfies.

$$u = [a \quad b]^T = (B^T B)^{-1} B^T Y \quad (8)$$

Where: a is the development coefficient that controls the system development trend; b is the grey action quantity that reflects the relationship of changes in the data.

Step 3: Build the model and solve the generated values and the reduced values. The forecast model is obtained according to the formula

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

$$k = 1, 2, \dots, n \quad (10)$$

After the inverse accumulating generation operation, the reduced predict is obtained.

2.2 K-means Clustering Algorithm

K-means clustering algorithm is an unsupervised learning algorithm

and a partition-based algorithm. Generally, Euclidean distance is used as an indicator to measure the similarity between the data points. The similarity is inversely proportional to the distance between the data points, and the greater the similarity, the smaller the distance between the data points. The K-means clustering algorithm starts with k cluster centroids as specified. According to the similarity between the data points and the cluster centroid, the location of the cluster centroid is constantly updated, and the sum of squared error (SSE) of the class cluster is constantly reduced. When the SSE no longer changes or the objective function converges, the clustering ends and the final result is obtained.

The basic idea of the algorithm is divided into 4 steps. Step 1: Randomly initialize K cluster centroid (data points) $C_i (1 \leq i \leq k)$. Step 2: Compute the Euclidean distance from each data point to the cluster centroid, and assign each data point to the closest cluster centroid. Step 3: Compute the mean of all points assigned to each cluster centroid, and re-assign the K cluster centroids with the newly computed mean. Step 4: Repeat steps 2-3, until the cluster centroid no longer changes or the number of clusters reaches the upper limit of iteration.

The formula of Euclidean distance from each data point to the cluster centroid is:

$$d(x, C_i) = \sqrt{\sum_{j=1}^m (x_j - C_{ij})^2} \quad (11)$$

Where: x is the data point; C_i is the i th cluster centroid; m is the dimension of the data point; x_j and C_{ij} are the j th attribute value of x and C_i , respectively.

The formula of SSE for the entire data set is:

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} |d(x, C_i)|^2 \quad (12)$$

Where: SSE indicates the quality of the clustering results; k is the number of clusters.

3. EMPIRICAL ANALYSIS

3.1 Analysis of the Results of Running the Grey Forecast Model

Due to the lack of data on social financing scale, we built a grey forecast model using the R language to predict the data of social financing scale in 2001 and completed the data. Then, we used the grey forecast model to predict the RMB deposit-reserve ratio (small and medium-sized deposit financial institutions) and the social financing scale (current month) in 2022.10-2023.12 respectively, as shown in Table 1.

On this basis, we built a grey forecast model using SPSSPRO to predict the RMB deposit-reserve ratio (small and medium-sized deposit financial institutions) and the social financing scale in 2022.10-2023.12 respectively, as shown in Table 2, Figure 1 and Figure 2:

3.2 Analysis of the results of running the K-means clustering algorithm

We ran the K-means clustering algorithm to analyze the social financing scale (current month) and the RMB deposit-reserve ratio (small and medium-sized deposit financial institutions) in 2001.01-2021.12 using MATLAB. We divided the data into four stages: tight money and tight credit, easy money and easy credit, tight money and easy credit, easy money and tight credit, as shown in Figure 3.

According to the index data of the monetary and credit cycle, we divided China's economy in 2001.01-2021.12 into four stages using the K-means clustering algorithm. From Figure 3, the economy of each period can be clearly and intuitively seen.

4. PREDICTION AND CLASSIFICATION OF CHINA'S ECONOMY IN 2022.10-2023.12

Cluster analysis was performed again based on the prediction results of economic indicators in 2022.10-2023.12. During this period, China's economy has a stable cycle law, and the monetary policy is from easy

Table 1: Predict of Social Financing Scale in 2001.1-2001.12

Year	Social Financing Scale (Current Month)
2001-1	53.41545
2001-2	57.74631
2001-3	223.7272
2001-4	242.8048
2001-5	27.74214
2001-6	125.2771
2001-7	193.1846
2001-8	224.3779
2001-9	509.7627
2001-10	587.6624
2001-11	587.4152
2001-12	941.7388

Table 2: Predict of RMB Deposit-Reserve Ratio and Social Financing Scale in 2022.10-2023.12

Year	RMB Deposit-Reserve Ratio: Small and Medium-Sized Deposit Financial Institutions (Month)	Social Financing Scale (Current Month)
2022-10	8.25	30668.084
2022-11	7.687	30860.681
2022-12	7.593	31053.293
2023-1	7.5	31245.92
2023-2	7.407	31438.562
2023-3	7.315	31631.218
2023-4	7.222	31823.889
2023-5	7.13	32016.576
2023-6	7.037	32209.277
2023-7	6.946	32401.993
2023-8	6.854	32594.723
2023-9	6.762	32787.469
2023-10	6.671	32980.23
2023-11	6.58	33173.005
2023-12	6.489	33365.795

to tight to east, while the credit policy is from tight to easy. Each period lasts for a similar time, and the economy is stable. From Figure 3, the economy of each month during this period is obtained, as shown in Table 4.

5. CONCLUSION AND SUGGESTION

To sum up, we divided China's economy in 2001.01-2021.12 into four stages: tight money and tight credit, easy money and easy credit, tight

money and easy credit, easy money and tight credit. Each economy has a different impact on different industries. Some industries may boom, while some industries may slump. The easy monetary policy reduces the economic pressure of enterprises, and lowers their financing threshold, which facilitates their development. On the contrary, it hinders their development. The easy credit policy makes the cash flow of enterprises increase and improves their willingness to expand, which improves their profits. By predicting China's future economy, we put forward investment suggestions to help investors to optimize their investment



Figure 1: Predict of Social Financing Scale (Current Month) in 2022.10-2023.12
Actual Fitted Predict
Backward predict



Figure 2: Predict of RMB Deposit-Reserve Ratio (Small and Medium-Sized Deposit Financial Institutions) in 2022.10-2023.12
Actual Fitted Predict
Backward predict

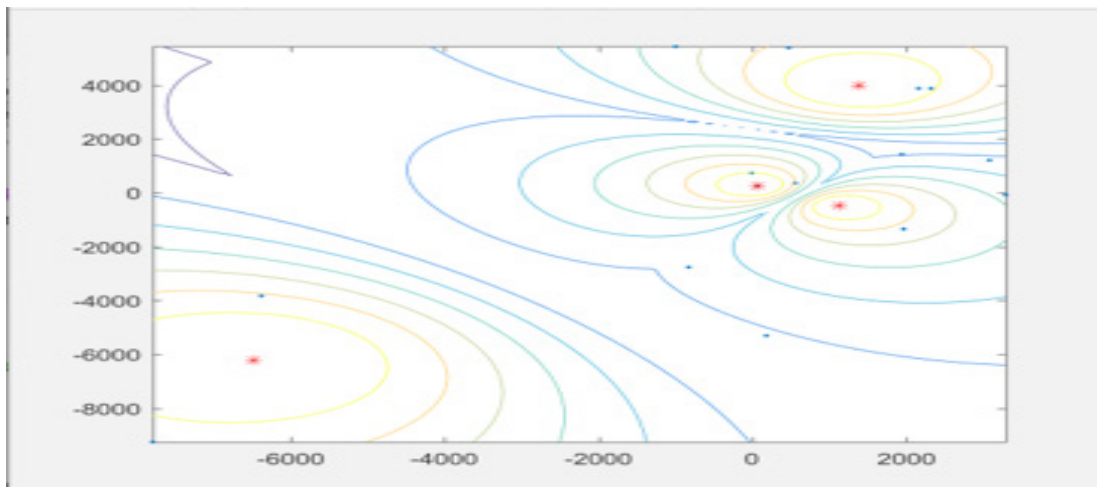


Figure 3 Diagram of running the K-means clustering algorithm

Table 3: China's Economy in 2001.01-2021.12

No.	Start-stop Time	Stage of Money and Credit Cycle	Duration	No.	Start-stop Time	Stage of Money and Credit Cycle	Duration
1	2001.01-2006.02	1	62	51	2014.12-2015.01	2	2
2	2006.03	4	1	52	2015.02-2015.05	4	4
3	2006.04-2006.12	1	9	53	2015.06	2	1
4	2007.01	4	1	54	2015.07-2015.09	4	3
5	2007.02-2007.05	1	4	55	2015.10	1	1
6	2007.06	4	1	56	2015.11	4	1
7	2007.07	1	1	57	2015.12	2	1
8	2007.08	4	1	58	2016.01	3	1
9	2007.09-2007.12	1	4	59	2016.02	4	1
10	2008.01	4	1	60	2016.03	2	1
11	2008.02-2008.03	1	2	61	2016.04-2016.05	4	2
12	2008.04	4	1	62	2016.06	2	1
13	2008.05-2008.11	1	7	63	2016.07	1	1
14	2008.12-2009.02	4	3	64	2016.08	4	1
15	2009.03	2	1	65	2016.09	2	1
16	2009.04	1	1	66	2016.10	4	1
17	2009.05	4	1	67	2016.11-2016.12	2	2
18	2009.06	2	1	68	2017.01	3	1
19	2009.07-2009.09	4	3	69	2017.02	4	1
20	2009.10	1	1	70	2017.03-2017.12	2	10
21	2009.11-2009.12	4	2	71	2018.01	3	1
22	2010.01	2	1	72	2018.02	4	1
23	2010.2-2010.12	4	11	73	2018.03-2018.04	2	2
24	2011.01	2	1	74	2018.05	4	1
25	2011.02	1	1	75	2018.06-2018.09	2	4

Table 3 (continued): China's Economy in 2001.01-2021.12

No.	Start-stop Time	Stage of Money and Credit Cycle	Duration	No.	Start-stop Time	Stage of Money and Credit Cycle	Duration
26	2011.03	2	1	76	2018.10	4	1
27	2011.04-2011.06	4	3	77	2018.11-2018.12	2	2
28	2011.07	1	1	78	2019.01	3	1
29	2011.08	4	1	79	2019.02	4	1
30	2011.09	1	1	80	2019.03	3	1
31	2011.10-2012.02	4	5	81	2019.04-2019.06	2	3
32	2012.03	2	1	82	2019.07	4	1
33	2012.04-2012.05	4	2	83	2019.08-2019.09	2	2
34	2012.06	2	1	84	2019.10	4	1
35	2012.07-2012.08	4	2	85	2019.11-2019.12	2	2
36	2012.09	2	1	86	2020.01	3	1
37	2012.10-2012.11	4	2	87	2020.02	4	1
38	2012.12-2013.01	2	2	88	2020.03-2020.06	3	4
39	2013.02	4	1	89	2020.07	2	1
40	2013.03-2013.04	2	2	90	2020.08-2020.09	3	2
41	2013.05-2013.07	4	3	91	2020.10	4	1
42	2013.08	2	1	92	2020.11-2020.12	2	2
43	2013.09-2013.12	4	4	93	2021.01	3	1
44	2014.01	2	1	94	2021.02	2	1
45	2014.02	4	1	95	2021.03	3	1
46	2014.03	2	1	96	2021.04-2021.05	2	2
47	2014.04-2014.05	4	2	97	2021.06	3	1
48	2014.06	2	1	98	2021.07	4	1
49	2014.07	1	1	99	2021.08-2021.09	3	2
50	2014.08-2014.11	4	4	100	2021.10-2021.12	2	3

Table 4: China's Economy in 2022.10-2023.12

No.	Start-stop Time	Stage of Money and Credit Cycle	Duration
1	2022.10-2023.01	4	4
2	2023.02-2023.05	1	4
3	2023.06-2023.09	3	4
4	2023.10-2023.12	2	3

portfolio and maximize their investment benefits.

74.

ACKNOWLEDGEMENTS

This paper is the research result of "A Study on the Optimal Site Selection of Firefly Recreation Areas Based on Grey Relational Analysis and Analytic Hierarchy Process - Assisting the Development of Tourism Economy" (No.S202210551019).

REFERENCES

- Deng, J. 2002. Grey forecast and decision-making. Wuhan: Huazhong University of Science & Technology Press
- Jiang, C., Zhang, J. 2022. Portfolio selection: The role of extreme loss constraints. *Contemporary Finance & Economics*, 2022(09), Pp. 64-74.
- Li, D. 2021. Research on cluster analysis algorithm and its effectiveness evaluation. *Computer Knowledge and Technology*, 17(30), Pp. 26-27.
- Li, X., Wei, S. 2021. Revisiting carl marx's theory of money, credit and economic cycle. *Journal of Shandong University (Philosophy and Social Sciences)*, 2021(05), Pp. 157-173.
- Xu, X. 1999. Basic classification of China's new system of national accounting. *Journal of Shanxi University of Finance and Economics*, 1999(01), Pp. 111-114.
- Zhou, W., Gao, J., An, G. 2022. Asset allocation theory: Research review and outlook. *Exploration of Financial Theory*, 2022(05), Pp. 72-80.

