

An investment theory model based on linear programming

Summary

Emerging industries are booming and mobilizing rapid economic growth, and this paper provides a discussion of the problems posed by the interaction of the industries, modeling how to promote the industries as well as increase employment to achieve economic sustainability.

To address question 1, the annual growth rates of the nine industries were first calculated. The correlation and causality of each group of industries is quantified by grouping the industries in pairs and applying Pearson's correlation coefficient and Granger causality test. Then, based on the correlation coefficients of each group, the impact between them and on economic development is evaluated.

To address question 2, the investment and GDP of each industry were used separately using linear regression, the goodness of fit was calculated, and the accuracy of the fit was tested using the coefficient of determination R^2 to derive a theoretical model of investment and GDP for each industry.

For Problem 3, the particle swarm algorithm is used to construct an objective function to maximize GDP, set up constraints to see that the industries that maximize profits are industry, wholesale and retail trade, and other industries, and solve the equation to arrive at an optimal investment ratio of 30.9% : 52.6% : 16.5%.

For problem 4, the elasticity values were calculated by replacing the original values with the elasticity values. Normalizing it and further calculating it to get the composite elasticity value and allocating the investment ratio of 39.21% : 30.41% : 30.39% according to the size of the elasticity value to the other sectors, finance and transport, storage and postal services will maximize the value of production.

For problem 5, a multi-objective planning model is constructed to maximize the employment rate and GDP, using hierarchical analysis to give weights to the variables to establish the objective function through the particle swarm algorithm to obtain the optimal solution for the industry, other industries and wholesale and retail trade investment ratio of 32.6% : 18.6% : 9.8%.

Keywords: Granger Causality Test; Particle Swarm Algorithm; Multi-Objective Optimization; Hierarchical Analysis;

Contents

1	Introduction	4
1.1	Problem Background	4
1.2	Restatement of the Problem	4
1.3	Our Work	4
2	Analysis of Problems	5
2.1	Analysis of problem 1	5
2.2	Analysis of problem 2	5
2.3	Analysis of problem 3	5
2.4	Analysis of problem 4	6
2.5	Analysis of problem 5	6
3	Model Assumptions	6
4	Notations	6
5	Model Preparation	6
5.1	Modeling based on Pearson's correlation coefficient for solving	6
5.1.1	Preparation Based on the Pearson Correlation Coefficient Model	7
5.1.2	Establishment Based on the Pearson Correlation Coefficient Model	7
5.1.3	Solution Based on The Pearson Correlation Coefficient Model . .	8
5.2	Establishment and Solution of Linear Programming Models Based on Least Squares Methods	11
5.2.1	Preparation of Linear Programming Models Based on Least Squares Methods	11
5.2.2	Establishment of a Linear Programming Model Based on The Least Squares Method	11
5.2.3	Solving Linear Programming Models Based on Least Squares Meth- ods	12
5.3	Particle Swarm Algorithm Based Linear Programming Model Building and Solving	14
5.3.1	Preparation of a linear programming model based on particle swarm algorithm	14
5.3.2	Modeling of linear programming based on particle swarm algo- rithm	14
5.3.3	Particle Swarm Algorithm Based Solving of Linear Programming Models	15
5.4	Modeling and solving for investment elasticity	16
5.4.1	Preparation of investment elasticity models	16
5.4.2	Modeling of investment elasticity	17
5.4.3	Solving the investment elasticity model	18
5.5	Multi-Objective Planning Modeling and Solving	19
5.5.1	Preparation of multi-objective planning models	19
5.5.2	Multi-objective planning modeling	20
5.5.3	Multi-objective planning model solution	21
6	Evaluation and generalization of the model	22
6.1	Advantages of the model	22
6.2	Disadvantages of the model	22

1 Introduction

1.1 Problem Background

Guided by Xi Jinping's thought of socialism with Chinese characteristics in the new era, it comprehensively implements the spirit of the Third Plenary Session of the 20th CPC Central Committee and promotes the development of the national economy. China's industrial structure has also undergone significant changes, and emerging industries have flourished. Compared with the traditional manufacturing industry, emerging industries have provided new impetus for China's economic growth, and have had a profound impact on economic development as well as the economic landscape. At present, the interrelationships between domestic industries have various impacts, and this paper examines this issue.

1.2 Restatement of the Problem

The interrelationships between domestic industries are intricate and complex, and this paper will comprehensively consider the complexity and industrial interconnect-edness of these industries with the aim of achieving balanced and sustainable economic growth, enhancing employment, and focusing on the following problems:

- Question 1: Collect data related to major domestic industries and analyze and discuss the interrelationships within China's major industries and their impact on economic development.
- Question 2: Design one or more theoretical models of investment, study the relationship between investment and the gross domestic product of various industries, and evaluate the models.
- Question 3: Analyze the percentage allocation of investments in all or only three types of industries to achieve the highest GDP using a total investment fund of 1 trillion units.
- Question 4: Analyze the proportion of investment allocated to all or just three types of industries that use 1 trillion units to effectively stimulate employment.
- Question 5: Optimize investment plans to ensure a steady increase in employment while increasing GDP.

1.3 Our Work

The problem requires us to fight fires by optimizing the locations of two type of drones. In order to avoid complicated description, intuitively reflect our work process, the flow chart is shown in Figure 1:

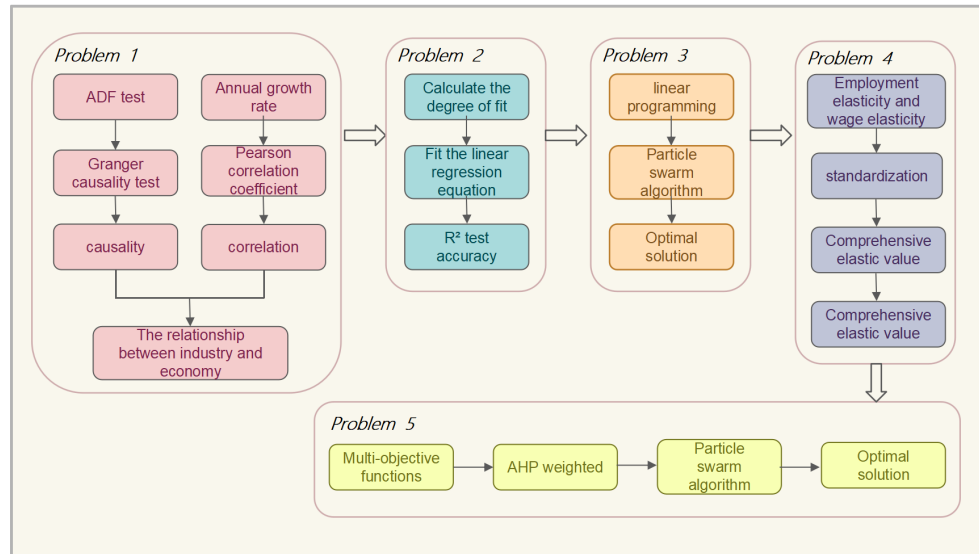


Figure 1: Ourwork

2 Analysis of Problems

2.1 Analysis of problem 1

The major industries in China include agriculture, forestry, animal husbandry and fishery, industry, construction, wholesale and retail trade, transportation, storage and postal service, accommodation and food service, finance, real estate, and other industries. Gross product and GDP of each industry in recent years were collected from the National Bureau of Statistics to analyze the year-on-year growth rate of the industry. By calculating the Pearson's correlation coefficient between each two industries, explore the interactions between industries and their role in the economy.

2.2 Analysis of problem 2

An investment theory model is a theoretical framework used to analyze and predict investment behavior and its impact on the economy. Therefore, we utilize a linear regression model to derive the fitted parameters for each industry for each year using known data, which forms the core of the investment theory model. Finally, the fitted parameters are evaluated to ensure the accuracy and reliability of the model.

2.3 Analysis of problem 3

To maximize the benefits, it is necessary to predict the changes in the gross product of each industry and analyze the changes in the total economic value of the industry. Therefore, the future gross product of each industry can be predicted through the fitting function derived from problem two. According to the proportion of each industry to divide the amount of investment, in the proportion of the size of the first three industries as the three industries limited to investment, and finally through the proportion of the division of 1 trillion dollars of investment.

2.4 Analysis of problem 4

To improve the employment rate as well as the quality of employment, two factors should be considered: the employed population and the wages of jobs. Firstly, the data is cleaned in order to ensure the integrity and consistency of the data. Secondly, we calculate the growth rate of employment population and wages in each industry, through which we calculate the employment elasticity and wage elasticity, take the average of the elasticity values, calculate the comprehensive elasticity value, and then get the investment allocation of each industry, and choose the three industries with the highest comprehensive elasticity for investment.

2.5 Analysis of problem 5

To improve the employment rate as well as the quality of employment, two factors should be considered: the employed population and the wages of jobs. Firstly, the data is cleaned in order to ensure the integrity and consistency of the data. Secondly, we calculate the growth rate of employment population and wages in each industry, through which we calculate the employment elasticity and wage elasticity, take the average of the elasticity values, calculate the comprehensive elasticity value, and then get the investment allocation of each industry, and choose the three industries with the highest comprehensive elasticity for investment.

3 Model Assumptions

This paper makes the following assumptions to address the above questions:

Assumption 1: Only the movement of the entire investment is considered, without specifically distinguishing between specific forms of cash flows such as own funds and borrowed funds.

Assumption 2: Errors arise only from science and technology, policy, labor, etc. and are not related to force majeure factors.

4 Notations

Some important mathematical notations used in this paper are listed in Table 1.

5 Model Preparation

5.1 Modeling based on Pearson's correlation coefficient for solving

The Granger causality test[1] is utilized to study the interactions between major industries in China, and the Pearson correlation number[2] is used to study the relationship between industries, and the two methods are combined to study the impact on economic development.

Table 1: Model assumptions.

Notation	Hidden Meaning
S_1	Gross Domestic Product (GDP)
S_2	Agriculture, forestry, animal husbandry and fisheries
S_3	Industries
S_4	Building industry
S_5	Wholesale and retail trade
S_6	Transportation, storage and postal services
S_7	Accommodation and catering
S_8	Financial sector
S_9	Real estate industry
S_{10}	Other industries
r	Pearson correlation coefficient
R^2	Coefficient of determination (CoD)
α_i	Return on investment (ROI)

5.1.1 Preparation Based on the Pearson Correlation Coefficient Model

Before the model is built, ADF test should be carried out on the data of GDP and Gross Domestic Product (GDP) of each industry in recent years to determine whether the time series is smooth or not.

5.1.2 Establishment Based on the Pearson Correlation Coefficient Model

The data shows that, to date, the gross output of various industries has been on an upward trend. From a statistical perspective, the Granger causality test can be applied to the gross outputs of various industries over the past 20 years ($S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9$ and S_{10}) and 10 indicators including S_1 , to examine whether causal relationships exist between the indicators. The specific steps for the Granger causality test are illustrated in the following figure:

The following regression model was developed based on the lag order:

$$C_t = \alpha + \beta_1 C_{t-1} + \beta_2 C_{t-2} + \gamma_1 D_{t-1} + \gamma_2 D_{t-2} \quad (1)$$

where C_t is the consumption expenditure of an industry in the first t period, C_{t-1} and C_{t-2} are the lagged terms of the consumption expenditure of the industry, D_{t-1} and D_{t-2} are the lagged terms of another industry, α is a constant term, and $\beta_1, \beta_2, \gamma_1$ and γ_2 are the parameters to be estimated.

In testing whether an industry is a Granger cause of consumer spending, the original hypothesis(H_0) is that the industry is not a Granger cause of consumer spending. The alternative hypothesis(H_1) is that the industry is a Granger cause of consumption expenditure. The same is true when testing consumption expenditure.

The original hypothesis is usually tested using the F test, where the expression for the statistic F is:

$$F = \frac{(RSS_r - RSS_u)/q}{RSS_u/(n - k)} \quad (2)$$

Where RSS_r is the residual sum of squares for the restricted model, RSS_u is the

residual sum of squares for the unrestricted model, q is the number of parameters restricted, n is the sample size, and k is the number of parameters for the unrestricted model.

Comparing the F statistic from the above calculations with the critical value of the F distribution gives an indication of whether one industry is the Granger cause of another.

At the same time, GDP is analyzed by treating it as a more visual growth rate. First, any two industries are divided into a group, using X to indicate one industry in a group and Y to indicate another industry in a group. The Pearson's correlation coefficient is calculated with the change of growth rate of each industry in recent years, so as to present the relationship between each group of industries.

Calculate the year-to-year growth rate for the raw data:

$$a = \left(\frac{b_j - b_{j-1}}{b_{j-1}} \right) \cdot 100\% \quad (3)$$

where denotes the annual growth rate of the total output of the industry, b_j ($j = 1, 2, \dots, n$) denotes the GDP of an industry in the year b , and b_{j-1} ($j = 1, 2, \dots, n$) denotes the GDP of another industry in the year $j - 1$.

Pearson correlation coefficient was calculated based on the growth rate:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}, i = 1, 2, \dots, n \quad (4)$$

where n denotes the total number of years to be calculated, i denotes the year, X_i and Y_i denote the growth rate of GDP in the first i year respectively. \bar{X} and \bar{Y} denote the average of the growth rates in the sample totals, respectively.

Calculate each group of data one by one to get the correlation coefficient between each two industries and analyze the correlation. The larger the value of $|r|$, the stronger the correlation; conversely, the smaller the value of $|r|$, the weaker the correlation.

5.1.3 Solution Based on The Pearson Correlation Coefficient Model

From the analysis of the ADF test results, it can be seen that the gross domestic product and GDP of each industry are smooth time series. Granger causality test was conducted using SPSSPRO on 10 data of China's industries and GDP in the last 20 years to get the comparison of significance level of 90 groups of industries as shown in the figure below.

Paired Sample	F-value	P-value	Significance	Paired Sample	F-value	P-value	Significance	Paired Sample	F-value	P-value	Significance	Paired Sample	F-value	P-value	Significance	Paired Sample	F-value	P-value	Significance
S2 → S1	0.119	0.889	Not significant	S3 → S2	1.059	0.375	Not significant	S7 → S8	2.605	0.112	Not significant	S7 → S9	13.484	0.001	***	S9 → S3	2.282	0.141	Not significant
S1 → S2	3.29	0.07	*	S2 → S3	4.178	0.04	**	S8 → S7	1.954	0.181	Not significant	S9 → S7	0.204	0.818	Not significant	S3 → S9	6.412	0.012	**
S3 → S1	0.349	0.712	Not significant	S4 → S2	7.61	0.006	***	S10 → S8	0.913	0.426	Not significant	S10 → S9	0.286	0.756	Not significant	S7 → S3	8.649	0.004	***
S1 → S3	2.105	0.161	Not significant	S2 → S4	0.164	0.85	Not significant	S8 → S10	7.508	0.007	***	S9 → S10	2.012	0.173	Not significant	S3 → S7	1.438	0.273	Not significant
S4 → S1	0.475	0.632	Not significant	S8 → S2	2.459	0.124	Not significant	S6 → S5	1.333	0.297	Not significant	S10 → S7	2.639	0.109	Not significant	S10 → S3	2.668	0.107	Not significant
S1 → S4	1.176	0.339	Not significant	S2 → S8	1.056	0.376	Not significant	S5 → S6	1.229	0.324	Not significant	S7 → S10	4.429	0.034	**	S3 → S10	0.711	0.509	Not significant
S8 → S1	2.001	0.175	Not significant	S5 → S2	2.699	0.105	Not significant	S9 → S5	0.285	0.756	Not significant	S7 → S2	3.473	0.062	*	S8 → S4	0.785	0.476	Not significant
S1 → S8	3.868	0.048	**	S2 → S5	2.062	0.167	Not significant	S5 → S9	11.146	0.002	***	S2 → S7	4.152	0.04	**	S4 → S8	4.674	0.03	**
S5 → S1	2.28	0.142	Not significant	S6 → S2	1.463	0.267	Not significant	S7 → S5	4.63	0.03	**	S10 → S2	1.849	0.196	Not significant	S5 → S4	2.623	0.11	Not significant
S1 → S5	0.481	0.629	Not significant	S2 → S6	1.76	0.211	Not significant	S5 → S7	2.246	0.145	Not significant	S2 → S10	2.503	0.12	Not significant	S4 → S5	0.461	0.641	Not significant
S6 → S1	3.24	0.072	*	S9 → S2	6.183	0.013	**	S10 → S5	2.073	0.165	Not significant	S4 → S3	2.444	0.126	Not significant	S6 → S4	1.757	0.211	Not significant
S1 → S6	2.012	0.173	Not significant	S2 → S9	1.089	0.365	Not significant	S5 → S10	3.334	0.068	*	S3 → S4	0.42	0.665	Not significant	S4 → S6	1.365	0.29	Not significant
S9 → S1	1.796	0.205	Not significant	S5 → S8	8.109	0.005	***	S9 → S6	0.691	0.518	Not significant	S8 → S3	2.473	0.123	Not significant	S9 → S4	0.937	0.417	Not significant
S1 → S9	7.318	0.007	***	S8 → S5	1.271	0.313	Not significant	S6 → S9	10.767	0.002	***	S3 → S8	5.682	0.017	**	S4 → S9	2.372	0.132	Not significant
S7 → S1	6.257	0.012	**	S6 → S8	3.398	0.065	*	S7 → S6	9.846	0.002	***	S5 → S3	3.553	0.059	*	S7 → S4	3.226	0.073	*
S1 → S7	1.138	0.35	Not significant	S8 → S6	2.495	0.121	Not significant	S6 → S7	2.228	0.147	Not significant	S3 → S5	1.412	0.279	Not significant	S4 → S7	0.54	0.595	Not significant
S10 → S1	0.829	0.458	Not significant	S9 → S8	0.144	0.867	Not significant	S10 → S6	4.28	0.037	**	S6 → S3	4.903	0.026	**	S10 → S4	0.456	0.644	Not significant
S1 → S10	1.506	0.258	Not significant	S8 → S9	1.351	0.293	Not significant	S6 → S10	1.495	0.26	Not significant	S3 → S6	1.207	0.33	Not significant	S4 → S10	1.702	0.22	Not significant

Figure 2: Comparison of significance levels.

In the above table, ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

A significant $p < 0.05$ indicates the rejection of the original hypothesis (that one set of time series is not the cause of the other set of time series), i.e., the left-hand side variables can cause changes in the right-hand side variables with Granger causality, and vice versa, there is no Granger causality. In addition, the smaller the value of p , the more significant the difference between the observed data and the original hypothesis.

In response to the results of the data in the above table, there is Granger causality between the following industries.

Paired Sample	F-value	p-value	Significance	Paired Sample	F-value	p-value	Significance
S7 → S5	3.226	0.073	*	S6 → S3	4.903	0.026	**
S6 → S1	3.24	0.072	*	S3 → S8	5.682	0.017	**
S1 → S2	3.29	0.07	*	S9 → S2	6.183	0.013	**
S5 → S10	3.334	0.068	*	S7 → S1	6.257	0.012	**
S6 → S8	3.398	0.065	*	S3 → S9	6.412	0.012	**
S7 → S2	3.473	0.062	*	S1 → S9	7.318	0.007	***
S5 → S3	3.553	0.059	*	S8 → S10	7.508	0.007	***
S1 → S8	3.868	0.048	**	S4 → S2	7.61	0.006	***
S2 → S3	4.178	0.04	**	S5 → S8	8.109	0.005	***
S2 → S7	4.152	0.04	**	S7 → S3	8.649	0.004	***
S10 → S6	4.28	0.037	**	S5 → S9	11.146	0.002	***
S7 → S10	4.429	0.034	**	S6 → S9	10.767	0.002	***
S7 → S5	4.63	0.03	**	S7 → S6	9.846	0.002	***
S4 → S8	4.674	0.03	**	S7 → S9	13.484	0.001	***

Figure 3: Granger causality.

Among them, S_7 can directly affect S_9 , S_6 and S_3 , finance can directly affect other industries, construction can directly affect agriculture, forestry and fisheries, wholesale

and retail trade can directly affect finance and real estate, and transportation, warehousing and postal services can directly affect real estate.

At the same time, the economic growth rate of each group of different industries is calculated using matlab to get the change in growth rate of each group of data during 2004-2023 years. The following figure shows the line graph of the change in the growth rate of each industry over a period of 20 years.

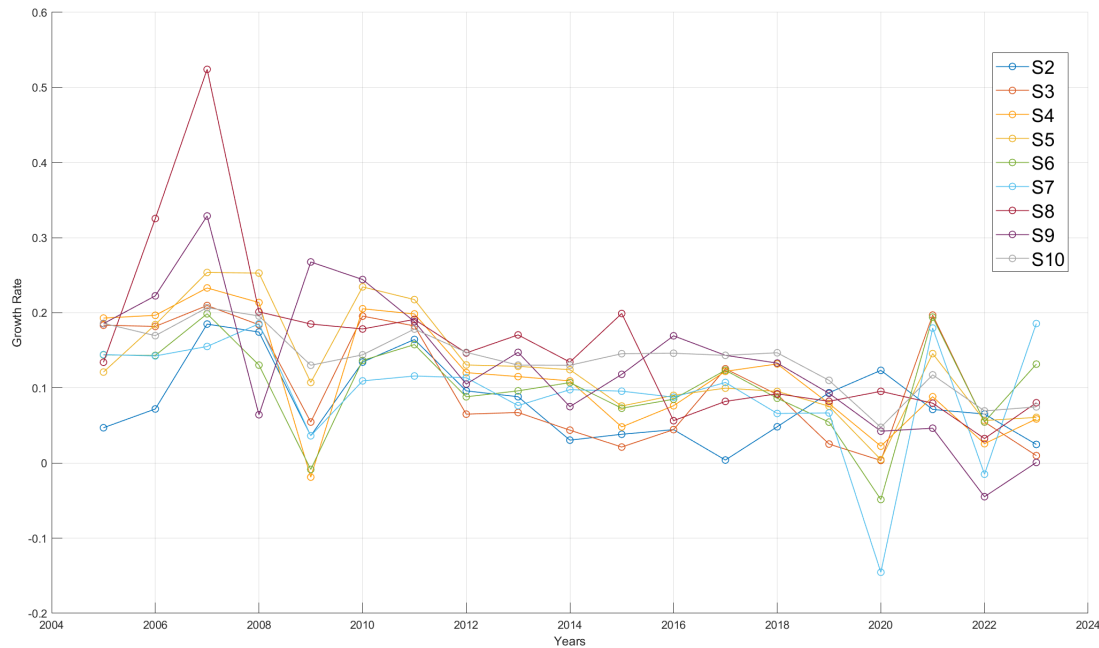


Figure 4: Growth rate over 20 years.

From the results of the above graph, it can be seen that in these 20 years, the vast majority of the industries had positive annual growth rates, i.e., the gross product of the industry showed an upward trend. However, in the year 2020, the impact of the New Crown Epidemic led to a downward trend in the gross product of most of the industries. Among them, the accommodation and food service industry showed a negative growth rate due to the tourism industry being more affected.

The data from the above graph was brought into the Pearson correlation coefficient formula to calculate the correlation between each of the two industries to obtain the following heat map.

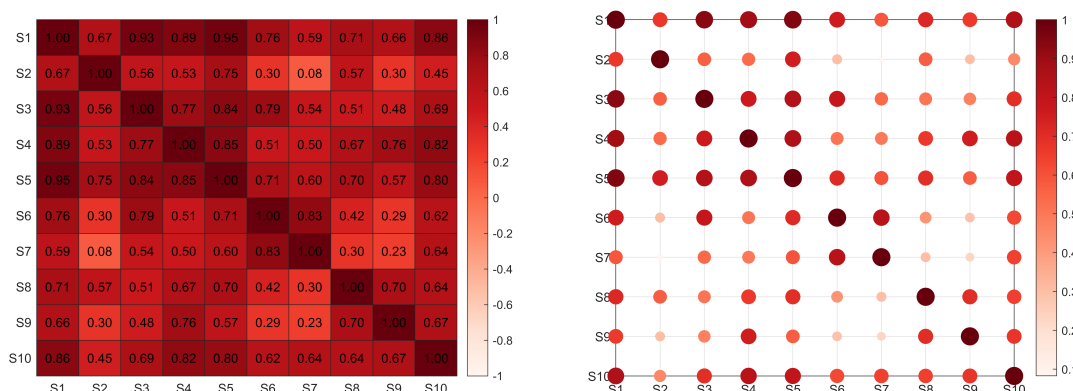


Figure 5: Pearson correlation coefficient heat distribution.

In Pearson correlation coefficient, (0.8, 1), denotes very strong correlation, (0.6, 0.8) denotes strong correlation, (0.4, 0.6) denotes medium strength correlation, (0.2, 0.4) denotes weak correlation, and (0, 0.2) denotes very weak correlation or no correlation. From the above figure, it can be seen that there is a very strong correlation between wholesale and retail trade and GDP, industry and transportation, storage and postal services. There is a very strong correlation between industry and GDP.

5.2 Establishment and Solution of Linear Programming Models Based on Least Squares Methods

The parameters of the linear model were estimated using the least squares method to study the relationship between the amount of investment and the gross product of each industry. The merit of the model fit was assessed by the coefficient of determination R^2 to compare the models.

5.2.1 Preparation of Linear Programming Models Based on Least Squares Methods

Typically, the expression for the form of the linear regression model is:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (5)$$

The fitting coefficients are calculated without taking into account the error term ε , i.e., random errors that cannot be explained by the model.

5.2.2 Establishment of a Linear Programming Model Based on The Least Squares Method

The fitting function assumptions for the linear regression for each industry are as follows:

$$GDP_{S_i} = \beta_i + \alpha_i X_{S_i}, (i = 2, 3, \dots, 10) \quad (6)$$

where GDP_{S_i} denotes the GDP of the i ($i = 2, 3, \dots, 10$) industry, $\alpha_i, \beta_i, \delta_i$ denotes the fit of the i industry, and X_{S_i} denotes the investment amount of the i industry.

R^2 denotes the accuracy of the fit, which is expressed as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

where n denote the number of industries, y_i denotes the gross production value of the i th industry, \hat{y}_i denotes the forecasted gross production value of the i th industry, and \bar{y} denotes the mean value of the gross production value of the industry. $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ denotes the variance not explained by the model, and $\sum_{i=1}^n (y_i - \bar{y})^2$ denotes the total variance of the total production value of the industry.

$R^2 = 1$ indicates that all of the variance is explained by the regression model, $0 < R^2 < 1$ indicates that a portion of the variance is explained, and $R^2 = 0$ indicates that the model does not predict better than the prediction using the mean of the industry's total output value.

5.2.3 Solving Linear Programming Models Based on Least Squares Methods

Let the number of industries $n = 9$, and build different linear regression models for different industries. Construct the design matrix X and estimate the model parameters using least squares, i.e., solve the formal equation:

$$(X^T X)\gamma = X^T y \quad (8)$$

where γ is the parameter vector, containing the intercept term and slope, and y is the vector of observations.

Parameter estimation $\hat{\gamma}$ can be done by:

$$\hat{\gamma} = (X^T X)^{-1} X^T y \quad (9)$$

The slope, intercept and R^2 in the linear regression equation for each industry were obtained by solving the above parameters using matlab as follows 2:

Table 2: Comparison of linear fitting degrees by industry.

estate	slope	intercept	R^2
S_2	1.67	23218.63	0.95
S_3	1.04	41009.20	0.97
S_4	-0.19	42586.76	-0.15
S_5	3.56	10751.51	0.87
S_6	0.59	4254.32	0.97
S_7	2.11	2029.97	0.75
S_8	50.78	6753.99	0.51
S_9	0.41	-784.52	0.98
S_{10}	1.35	8516.61	0.98

$R^2 = 1$ When the fit is perfect, (0.8, 1) means the fit is very good, (0.6, 0.8) means the fit good, (0.4, 0.6) means the fit is fair, (0.2, 0.4) means the fit is poor, and (0, 0.2) means the fit is very poor.

The plots of the fitted functions for each industry are shown below:

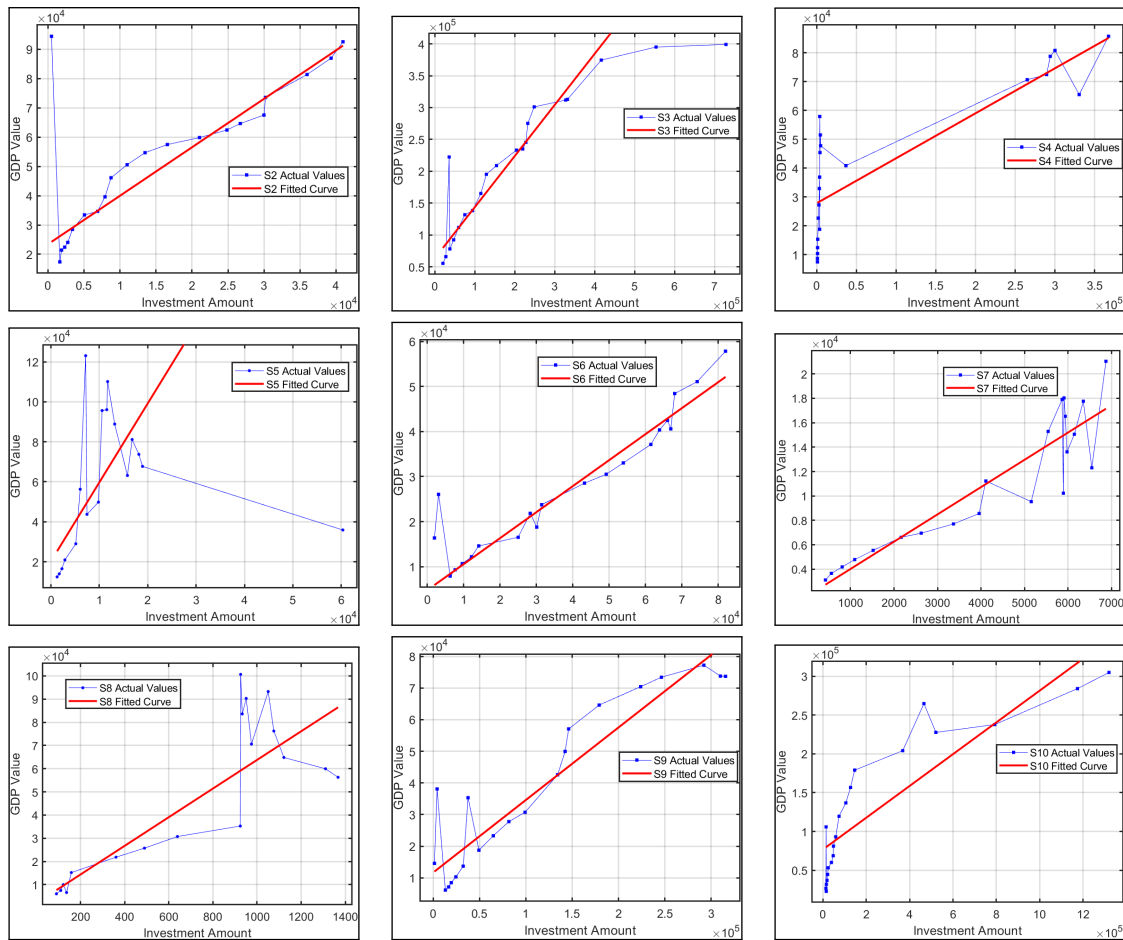


Figure 6: Linear fitting functions for nine major industries.

Combining the images and R^2 it is clear that S_2 , S_3 , S_5 , S_6 , S_9 and S_{10} industry fits very well, but S_4 industry has a negative value of R^2 , which indicates that this industry is not suitable for linear programming. The quadratic polynomial regression model, the exponential regression model, and the logarithmic regression model were built for this industry, and the resulting function plots are as follows Figure7.

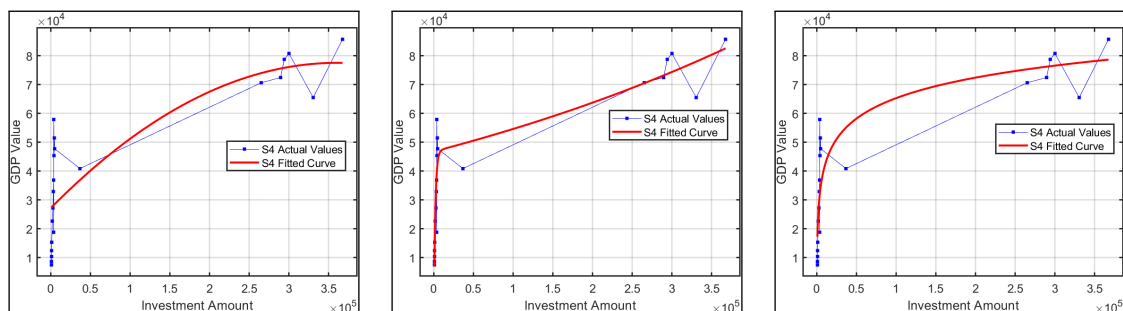


Figure 7: Image of a nonlinear fitting function.

The three regression models corresponding to R^2 are shown in the table below Table3:

Table 3: Comparison of R^2 under different models.

	quadratic polynomial regression	exponential regression	logistic regression
R^2	0.0084	-0.0043	0.0381

Since there is no significant correlation between the S_4 original data for the industry and there is no significant relationship among the four regression models, the linear regression model is taken as the investment model for the S_4 industry in order to simplify the model for this question.

5.3 Particle Swarm Algorithm Based Linear Programming Model Building and Solving

Based on the investment model derived from Problem 2, the optimal solution for investment allocation can be found by particle swarm algorithm.

5.3.1 Preparation of a linear programming model based on particle swarm algorithm

The steps of the particle swarm algorithm are shown below Figure 8.

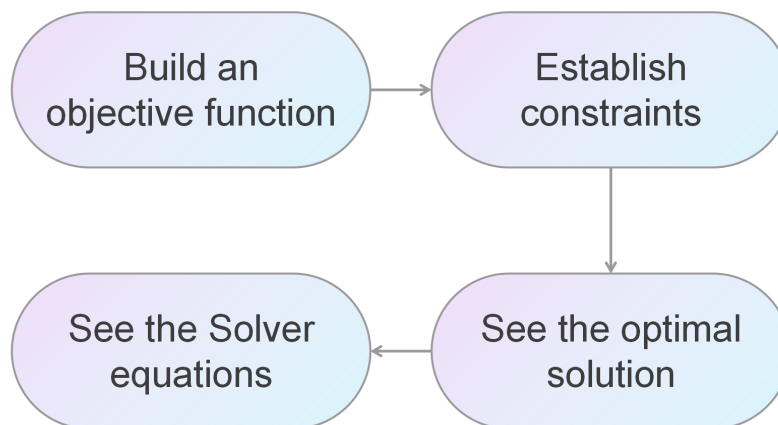


Figure 8: Basic steps of particle swarm algorithm.

5.3.2 Modeling of linear programming based on particle swarm algorithm

Total GDP is equal to the sum of the GDP of each industry, and profit is GDP minus investment costs:

$$profit = GDP - invest \quad (10)$$

To maximize profit with limited capital, the objective function is:

$$\begin{aligned} \max Z = & 16715 * I_2 + 10448 * I_3 - 18908 * I_4 + 35585 * I_5 + 5866.8 * I_6 \\ & + 2111.5 * I_7 + 507.79 * I_8 + 406.9 * I_9 + 1354.3 * I_{10} + 138337.38 \end{aligned} \quad (11)$$

The constraints are:

$$s.t. \begin{cases} I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9 + I_{10} \leq 1 \\ I_2 \in (0.073, 1) \\ I_3 \in (0.309, 1) \\ I_4 \in (0.054, 1) \\ I_5 \in (0.075, 1) \\ I_6 \in (0.04, 1) \\ I_7 \in (0.015, 1) \\ I_8 \in (0.04, 1) \\ I_9 \in (0.044, 1) \\ I_{10} \in (0.165, 1) \end{cases} \quad (12)$$

Particle swarm algorithm based on the above conditions leads to an optimal solution. That is, the industry that can realize the maximum profit is obtained and invested in it.

For investing in three industries to achieve the highest GDP, a linear fit can be made based on the return on investment for each industry to calculate the rate of return for each industry on the next investment. The ROI formula is as follows:

$$\alpha_i = \frac{G_i - I_i}{I_i} (i = 2, 3, \dots, n) \quad (13)$$

Based on the projections, the three industries with higher rates of return can be selected and invested in their respective proportions.

5.3.3 Particle Swarm Algorithm Based Solving of Linear Programming Models

The particle swarm algorithm is solved by matlab and the optimal solution is obtained as follows:

Table 4: Investment by industry.

parameters	Solved value (in trillion units)
Z	13770.82
I_2	0.079
I_3	0.312
I_4	0.054
I_5	0.272
I_6	0.042
I_7	0.016
I_8	0.04
I_9	0.044
I_{10}	0.165

This shows that GDP can be maximized by investing 0.312 trillion in the S_3 industry, 0.272 trillion in the S_5 industry, and 0.165 trillion in the S_{10} industry.

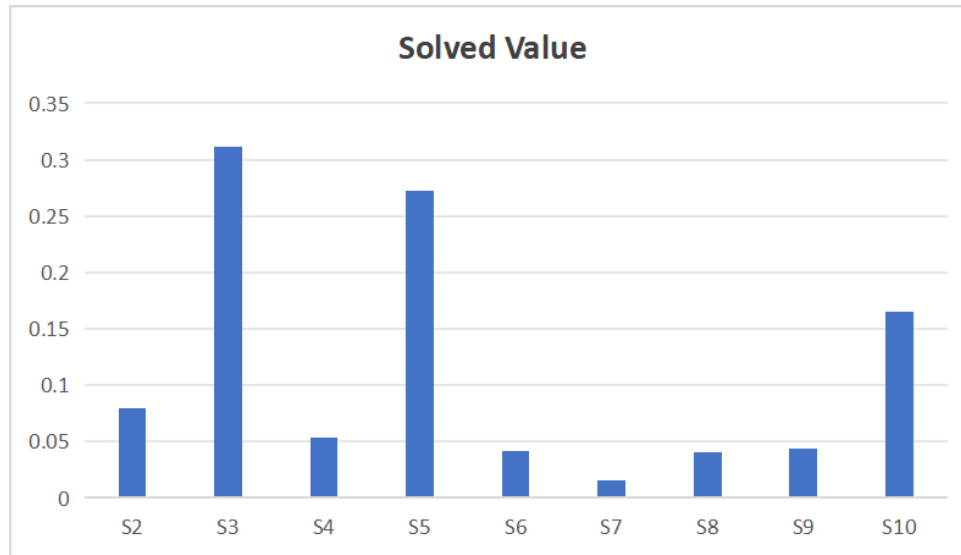


Figure 9: Proportion of investment allocation.

From the above figure, it can be seen that S_3 , S_5 and S_{10} industries have the largest solution values. The particle swarm optimization algorithm is performed again for these three industries to solve the best investment ratio as shown in the table below.

Table 5: Investment in the three industries.

parameters	Solved value (in trillion units)
Z	60279.737
X_3	0.309
X_5	0.526
X_{10}	0.165

5.4 Modeling and solving for investment elasticity

Models are built to calculate and standardize the elasticity values for different industries, to represent the relationship between investment and employment, and to plan investment allocation ratios.

5.4.1 Preparation of investment elasticity models

The cleaned data are first run through an elasticity formula that reflects the impact of investment on employment growth and wage growth [3]. And then the elasticity values are standardized to compare the elasticity values of different industries. The standardized elasticity values obtained in the previous step are further calculated to obtain the composite elasticity values. Finally, the calculation of investment allocation is carried out, in the case of not restricting the number of industries to be invested in, industries with higher comprehensive elasticity will receive more investment; if the investment is restricted to three industries, the three industries with the highest comprehensive elasticity value will be selected and the corresponding investment proportion will be allocated.

5.4.2 Modeling of investment elasticity

The raw data is processed to some extent to become investment growth rate, wage growth rate [5], and employment growth rate for the next step of analysis. The employment elasticity and wage elasticity are calculated for each industry using the following formulas:

$$X_i = \frac{\alpha_i}{\beta_i} \quad (14)$$

$$Y_i = \frac{\delta_i}{\beta_i} \quad (15)$$

X_i denotes the employment elasticity, which reflects the impact of investment on employment growth [3], α_i denotes the employment growth rate, and β_i denotes the investment growth rate. Y_i denotes wage elasticity, which reflects the impact of investment on wage growth, and δ_i denotes wage growth rate. $i = 2, 3, \dots, 10$ Indicates industry number.

In order to enable direct comparison of the elasticity values of different industries and to improve the stability of the numerical calculations, the average annual employment elasticity and the flexible wage of each industry are first calculated and finally the following standardized formula is utilized:

$$Z = \frac{(E - \mu)}{\sigma} \quad (16)$$

E denotes the value of the elasticity, μ denotes the average annual employment elasticity and flexible wage, and σ denotes the difference between the maximum and minimum values of the employment elasticity or flexible wage.

The combined elasticity value is the mean of the sum of the standardized employment elasticity and the standardized wage elasticity. The investment allocation ratio for each industry is calculated from the combined elasticity value using the following formula.

$$N_1 = \frac{T}{\sum_{i=1}^n T_i} M \quad (17)$$

The above formula shows the calculation of the sectoral investment allocation ratio N_1 in case there is no restriction on investment in the industry, where T is the value of combined elasticity, $\sum_{i=1}^n T_i$ is the value of combined elasticity, and M is the total investment amount.

$$N_2 = \frac{T'}{\sum_{i=1}^n T'_i} \quad (18)$$

The above formula shows the calculation of industry investment allocation ratio N_2 in case of restriction on investment in three industries, where T is the combined elasticity value of the selected industries, $\sum_{i=1}^n T_i$ is the combined elasticity value of the selected industries.

5.4.3 Solving the investment elasticity model

Using data-processed employment and wage growth rates, the elasticity measures the sensitivity of the amount of investment to the size of the employed population and wages, and calculating the employment elasticity as well as the wage elasticity yields the following Figure 10:

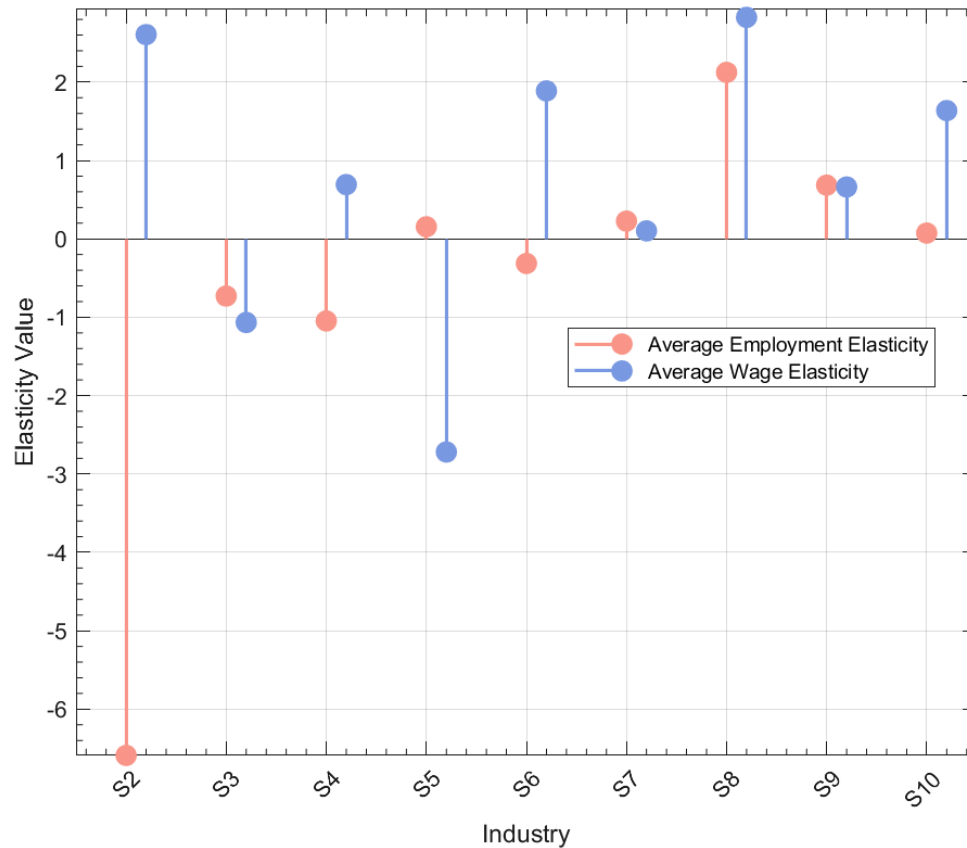


Figure 10: Employment elasticity and wage elasticity by industry.

The impact of investment on employment growth and wage growth can be visualized from the figure above. For industry S_1 , the employment fell while wages rose; employment and its wages fell in S_3 industries. Therefore, both should be considered when planning investment programs.

The combined elasticity is obtained through employment and wage elasticity, which is then calculated to obtain the investment allocation ratio, as shown below:

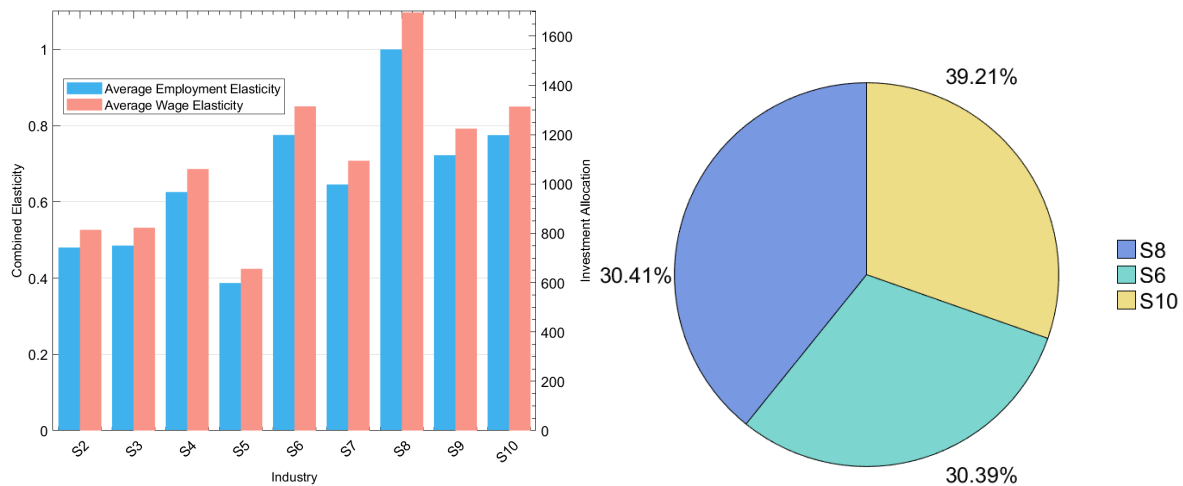


Figure 11: Investment allocation.

The left panel shows a graph of the composite elasticity indicator, with the blue bars showing the employment elasticity and the red bars showing the wage elasticity, the left vertical coordinate showing the result of the calculation based on these two variables, and the right vertical coordinate showing the amount of investment to be allocated. The chart on the right shows the proportion of investment that should be allocated to achieve the highest GDP if investment is limited to three sectors.

5.5 Multi-Objective Planning Modeling and Solving

A multi-objective planning model is used to weight the independent variables using hierarchical analysis, and the optimal solution is found by particle swarm algorithm.

5.5.1 Preparation of multi-objective planning models

In multi-objective optimization problems, there are often conflicting situations between different objective functions. Through the assignment, the relative importance of each objective can be clarified, so as to find a reasonable balance point between these multiple conflicting objectives.

The hierarchical analysis method (AHP) is utilized to assign weights to the objective functions in order to avoid conflicts between the objective functions as shown in the figure below:

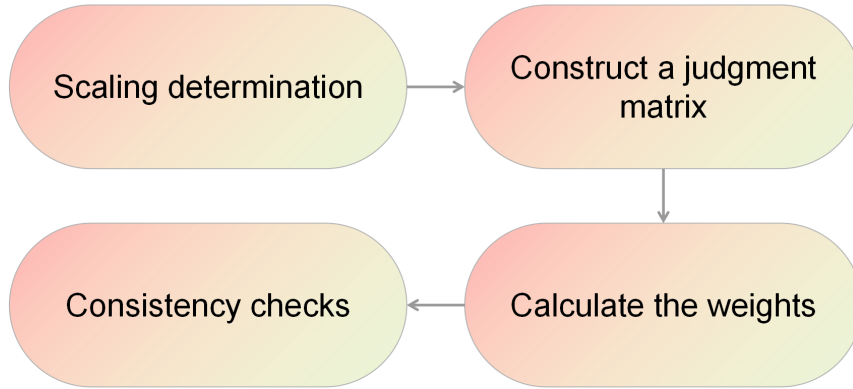


Figure 12: The basic steps of analytic hierarchy process.

The weight values can be calculated by sum and product method. Subsequently, these weight values are substituted into the objective function and finally the optimal investment allocation scheme is calculated with the help of Particle Swarm Optimization (PSO) algorithm.

5.5.2 Multi-objective planning modeling

A multi-objective planning model is a mathematical model used to optimize multiple objectives simultaneously. In order to increase employment while achieving the highest total GDP, a multi-objective planning model can be used to optimize employment as well as GDP at the same time; however, there is often a conflict between these objectives.

Define the objective function.

$$\max Y = nF + (1 - n)G \quad (19)$$

Clarify constraints:

$$s.t. \left\{ \begin{array}{l} F = -5.4 \times 10^{-7}I_2 - 5.9 \times 10^3I_3 + 9.9 \times 10^{-6}I_4 \\ + 1.7 \times 10^{-7}I_5 - 1.3 \times 10^{-8}I_6 + 1.3 \times 10^{-6} \cdot I_7 - 3 \times 10^{-5}I_8 + 1.6 \times 10^{-7}I_9 + 1.4 \times 10^{-6}I_{10} + 0.9 \\ G = 16715I_2 + 10448I_3 - 18908I_4 + 35585I_5 \\ + 5866.8I_6 + 2111.5I_7 + 507.8I_8 + 406.9I_9 + 1354.3I_{10} + 138337.3 \\ I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9 + I_{10} \leq 1 \\ I_2 \in (0.073, 1) \\ I_3 \in (0.309, 1) \\ I_4 \in (0.054, 1) \\ I_5 \in (0.075, 1) \\ I_6 \in (0.04, 1) \\ I_7 \in (0.015, 1) \\ I_8 \in (0.04, 1) \\ I_9 \in (0.044, 1) \\ I_{10} \in (0.165, 1) \end{array} \right. \quad (20)$$

Assigning weights to conflicting objectives through hierarchical analysis to determine the relative importance of each objective. The judgment matrix is processed using normalization in the sum-product method and the weight vector is calculated.

The computed weights are substituted into the objective function to transform the multi-objective optimization problem into a single-objective optimization problem.

Finally the particle swarm optimization algorithm is used to solve the single objective optimization problem to find the optimal solution.

5.5.3 Multi-objective planning model solution

The weight of F is obtained by hierarchical analysis as $\frac{1}{3}$ and the weight of G is $\frac{2}{3}$.

The particle swarm algorithm is solved and the optimal solution obtained is as follows:

Table 6: Investment in various industries.

Parameter	Solved Value
Objective Value (z)	31859.477
X_2	0.096
X_3	0.326
X_4	0.054
X_5	0.098
X_6	0.064
X_7	0.039
X_8	0.064
X_9	0.068
X_{10}	0.186

From the Table 6, it is clear that the solution values of S_3, S_5, S_{10} are the best solution values.

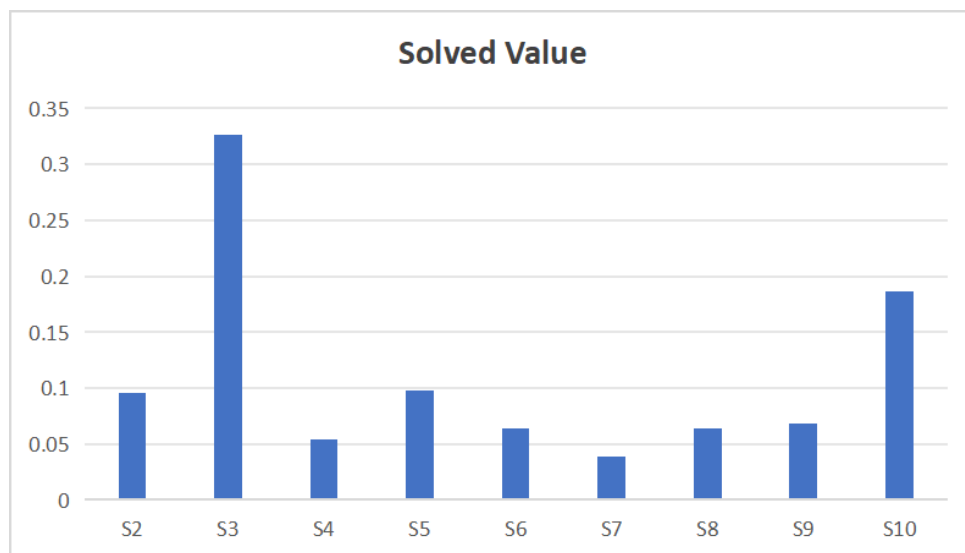


Figure 13: Proportion of investment allocation.

From the above figure, it can be seen that S_3 , S_5 and S_{10} industries have the largest solution value. Next, the particle swarm optimization algorithm is performed again for these three industries to solve for the optimal investment ratio.

6 Evaluation and generalization of the model

6.1 Advantages of the model

- Based on the model of Pearson's correlation coefficient, the interrelationship between industries is discussed from two aspects. The consideration is more comprehensive, which makes the results more convincing.
- The multi-objective planning model takes into account the conflict problem brought about by multiple optimization objectives, uses the hierarchical analysis method to weight the optimization objectives, and finally uses the particle swarm algorithm to derive the optimal solution, which uses a variety of analytical methods, is more rigorous in its consideration, and the results are more accurate.

6.2 Disadvantages of the model

- The simplification of the real economic environment by the investment resilience model may result in a model that does not fully capture the complexity of emergency activities and that relies heavily on the quality and reliability of the data.
- Least squares-based models are sensitive to outliers and cannot deal directly with categorical variables, with limited predictive power.

References

- [1] Diks, Cees, and Valentyn Panchenko. "A new statistic and practical guidelines for nonparametric Granger causality testing." *Journal of Economic Dynamics and Control* 30.9-10 (2006): 1647-1669.
 - [2] Cohen, Israel, et al. "Pearson correlation coefficient." *Noise reduction in speech processing* (2009): 1-4.
 - [3] Morén V, Wändal E. The employment elasticity of economic growth-A global study of trends and determinants for the years 2000-2017[J]. 2019.
 - [4] Burgi C, Hovhannisyan S, Mondragon-Velez C. GDP-Employment Elasticities across Develo** Economies[R]. The World Bank, 2024.
 - [5] Danziger L. The elasticity of labor demand and the minimum wage[J]. *Journal of Population Economics*, 2009, 22: 757-772.
-