

Original Research Article

Analysis on Evaluation of Economic Vitality of Some Cities in China Based on Principal Component Analysis

Xuming Li¹, Teng Pu^{2*}, Jing Huang¹, Xiaoyuan Li²

¹ Sichuan Agricultural University, Dujiangyan 611830, Sichuan Province, China

² Southwest University of Science and Technology, Mianyang 621000, Sichuan Province, China. E-mail: 1752560404@qq.com

Abstract: Through the research on the development indicators of major cities in China, this paper analyzes the development factors and achievements as the basis of data analysis, tries to rank some cities in China for economic vitality, and finds out the relevant factors that can promote the development of urban economy, and provides a reference for the government to formulate policies.

Keywords: Principal Component Analysis; Vitality Evaluation; Urban Policy Reference; Regional Economics; Unsupervised Learning

1. Background analysis

Nowadays, the economic development of cities is getting faster and faster, with different developing speed in different cities. How to evaluate the economic vitality of a city is an important issue. This paper evaluates the economic vitality of some cities in China, which can help the research of relevant institutions and provide reference for the government to formulate policies.

2. Construction of evaluation index system

In order to facilitate the quantitative evaluation of economic vitality, this paper selects the following ten indicators as the evaluation index system of economic vitality by looking up the literature and materials: Real GDP per capita, GDP growth rate, proportion of tertiary industry, investment in fixed assets, total amount of actual utilization of foreign capital, number of patent authorizations, number of students in ordinary colleges and universities, general public budget income per capita disposable income, and employment rate^[3,4].

3. Evaluation of economic viability through principal component analysis

In this paper, principal component analysis is adopted to establish the economic vitality evaluation model. PCA is the most widely used data dimensionality reduction algorithm, which can reduce the dimensionality of high-dimensional data and leave some of the most important features, subtract unimportant features, and improve the speed of data processing^[1,2].

First, the sample raw data matrix is established, which consists of n dimension vectors, where n takes 1,2,3...

$$X = (x_1, x_2, x_3, \dots, x_n)$$

In order to eliminate the difference in dimension and order of magnitude of each index data, the index data is standardized to obtain the standardized matrix. The standardizing methods are as follows:

$$A_{ij} = \frac{x_{ij} - \bar{x}}{\sum_i x_{ij}}$$

The larger the value of covariance Matrix R, a correlation matrix that reflects the relationships between standardized data, the more necessary it is to perform principal component analysis on the data. R_{ij} is the correlation coefficient between the Primitive Variables X_i and X_j . R is a real symmetric matrix, and it only needs to calculate its upper triangular elements or lower triangular elements, the formula of which is

$$R_{ij} = \frac{\sum_{k=1}^n (X_{kj} - X_i)(X_{kj} - X_j)}{\sqrt{\sum_{k=1}^n (X_{kj} - X_i)^2 (X_{kj} - X_j)^2}}$$

The EIGENVALUES, principal component contribution, and cumulative variance contribution are calculated according to covariance Matrix R, and the number of principal components is determined. Find the EIGENVALUES. R is a positive definite matrix whose eigenvalues are all positive numbers. The EIGENVALUES are arranged in order of magnitude. The magnitude of the EIGENVALUES reflects the magnitude of the influence of each principal component. The EIGENVALUES can be obtained by solving the following equations:

$$|R - \lambda I_p| = 0$$

Select the appropriate number of principal components. The eigenvalue should be more than 1 and the cumulative contribution rate of the eigenvalues is 80%-95%. The contribution rate is calculated as follows:

$$\frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^p \lambda_j}$$

The cumulative contribution rate of the first six eigenvalues is over 95%.

Serial number	Contribution rate	Cumulative contribution
1	0.395783179542776	0.395783179542776
2	0.249998889033788	0.645782068576563
3	0.120641170901503	0.766423239478066
4	0.0938136680846127	0.860236907562679
5	0.0577267035397899	0.917963611102469
6	0.0328727400561022	0.950836351158571

Table 1.

Establish the initial factor load matrix to explain the principal component. The eigenvectors corresponding to each eigenvalue are shown in **Table 2**.

1	2	3	4	5	6	
0.301	0.353	0.259	-0.039	0.598	0.201	X1
-0.072	0.046	0.807	0.160	-0.524	0.100	X2
0.385	-0.087	-0.206	-0.461	-0.376	0.075	X3
0.050	-0.516	0.321	0.209	0.429	-0.151	X4
0.400	-0.320	-0.025	0.143	-0.040	-0.122	X5
0.441	-0.106	0.059	0.052	-0.067	-0.525	X6
-0.015	-0.357	0.298	-0.754	0.140	0.094	X7
0.436	-0.168	-0.085	0.304	-0.129	0.240	X8
0.419	0.256	0.094	-0.004	0.018	0.489	X9
0.179	0.514	0.166	-0.179	0.010	-0.570	X10

Table 2.

From the eigenvectors in the table, it can be seen that:

The first principal component mainly reflects the impact of indicators such as real GDP per capita, the proportion of tertiary industry, the total amount of actual utilization of foreign capital, the number of patent authorizations, the general public budget income, and the per capita disposable income. These indicators are basically the most direct feedback data for economic development, and this component can be named as the principal component of income and investment.

The second principal component mainly reflects indicators including the impact of real GDP per capita, per capita disposable income, employment rate, which directly reflect the level of people's quality of life, and can be named as the principal component of people's quality of life.

The third principal component mainly reflects indicators including the impact of real GDP per capita, GDP growth rate, and fixed assets investment. These indicators reflect the impact of economic investment, and this component can be named as the principal component of investment volume.

The fourth principal component mainly reflects indicators of the impact of fixed assets investment and general public budget income, which reflect the use of public property, and can be named as the principal component of public assets.

The fifth principal component mainly reflects the impact of indicators such as real GDP per capita and GDP growth rate, which reflect the impact of visual changes of these indicators, and can be named as the principal GDP component.

The sixth principal component mainly reflects the influence of general public budget income, per capita disposable income and other indicators, which reflect the level of people's living income, but also reflect the quality of life of the people, and can be named as the principal component of people's living standard.

Taking each eigenvector as factor equation coefficients, the following factor equations are obtained:

$$Z_i = u_{i1}X_1 + u_{i2}X_2 + \dots + u_{ip}X_p$$

Calculate the scoring function F , and then get the comprehensive evaluation value of economic vitality of each city.

$$F = W_1Z_1 + W_2Z_2 + \dots + W_iZ_i$$

W_i is the corresponding contribution rate of each component, that is, the weight value.

The evaluation function is obtained by substituting data analysis as follows:

$$F = 0.40 * Z_1 + 0.25 * Z_2 + 0.12 * Z_3 + 0.09 * Z_4 + 0.06 * Z_5 + 0.03 * Z_6$$

By looking up the relevant index data of some cities, the paper evaluates and sorts through the PCA model, the result of which is shown in **Table 3**.

Urban	Score	Ranking
Shanghai	2.853224798	3
Shenzhen	5.99502061	1
Beijing	1.696093074	7
Guangzhou	0.613213428	9
Chongqing	-2.688876584	14
Chengdu	-2.799247606	16
Nanjing	2.65500397	4
Hangzhou	1.728419714	6
Suzhou	3.68293473	2
Tianjin	-1.121081492	13
Qingdao	0.330134127	10
Dongguan	-0.081196667	11
Zhengzhou	-2.701544041	15
Wuhan	-0.272396054	12
Xi'an	-4.057106823	18
Ningbo	2.563592128	5
Changsha	1.421694641	8
Shenyang	-6.165244978	19
Kunming	-3.652636971	17

Table 3.

4. Conclusion

From this, it can be concluded that people's traditional consciousness of the economy is relatively strong in the north; Shanghai and Shenzhen are of high rank, while the economic vitality performance of other cities is no worse than the north, such as Suzhou and Nanjing, which has a long history and cultural heritage, and whose economic heritage is not weaker than other cities^[5].

The main aspects that affect the economic development of a city are the life quality of people, the use of public property and the influence of foreign investment, and so on. It can be seen that for the economic development of a city, the most important thing is internal investment and foreign investment. In addition, the level of people's living standard is also closely related to the economic situation of a city. To develop the economy, it is necessary to start with people's living and asset investment.

References

1. Lu X, Guo W. The comprehensive evaluation index system of urban economic vitality (in Chinese). *Statistics and Decision Making* 2006; 11: 77-78.
2. Liu Z, Zhang Z. Index framework of state-owned economy vitality, control power and influence power and its

analysis: Taking Shanghai state-owned economy as an example (in Chinese). Volume of Economic Management of the Sixth Annual Conference of Shanghai Social Sciences; 2008.

3. Zhang P. Research on comprehensive evaluation based on principal component analysis (in Chinese). Nanjing: Nanjing University of Science and Technology; 2004.
4. Li J, Guo Y. Principal component analysis for multi-index evaluation (in Chinese). *Journal of Management Engineering* 2002; 16(1): 39-43. doi: 10.3969/j.issn.1004-6062.2002.01.011.
5. Lin Y, Liu P. China's economic development strategy and regional income gap (in Chinese). *Economic Research* 2003; (3):19-25+89.