Empirical Analysis of Economic Development Level of Chinese Provinces Based on PCA

Zile Xu[1], Jiaying Liao[2]

[1]Faculty of Mathematics and Statistics, Chongqing University, Chongqing, 401331, China
[2]School of Finance, Capital University of Economics and Business, Beijing, 110070, China

Abstract: In the past decades, China's overall economy has maintained growth, the economy of China's provinces in maintaining growth at the same time there are different degrees of change, and the development level of different provinces there are similarities and differences. Therefore, this paper takes the gross regional product (GDP) of 31 provinces including Beijing, Shanghai, Guangdong and Jiangsu from 2018 to 2020 as the dependent variable, uses principal component analysis (PCA) to comprehensively compare the development degree of 31 provinces, and evaluates and analyzes the results, so as to compare and find the similarities and differences of economic development of different provinces. In order to provide theoretical reference and guidance for national and regional development and policy formulation and implementation, the author compares the economic development gap among different provinces in China and reflects on the policy direction and path selection of China's acceleration of economic development in the context of the outbreak of COVID-19.

Keywords: Principal component analysis; China; Regional economy; Level of development

1. Introduction

Thanks to advantaged geographical advantages and tentative economic policies, China has entered an unprecedented period of rapid economic development. Under the circumstance of flexible use of comparative advantages, the overall economic strength is increasing day by day. China's GDP has grown from 149.541 billion US dollars in 1978 to 14.72 trillion US dollars in 2020 after 42 years of "pioneering" reform and opening-up. In the process of economic development, each province of our country has its own positioning and task, and belongs to different economic types. In December 2021, the Central Economic Work Conference made arrangements for the 2022 economic policy system from seven policy directions: macro, micro, structural, scientific and technological innovation, and reform and opening up, pointing out the direction for China's high-quality development in the future.

2. Construction of evaluation model of China's economic development level

(1) Collection of basic data of China's economic development level

In order to better analyze China's economic development, Choose to Beijing, Shanghai, Guangdong, Jiangsu and other 31 provinces (2018-2020) of GDP (X1), the population of permanent residents in the urbanization rate (X2), provincial resident population (X3), urban residents per capita disposable income (X4), rural residents per capita disposable income (X5), the local finance general budget revenue (X6), Total retail sales of consumer goods (X7), total household deposits of provinces (X8), growth rate of fixed asset investment in real estate development (X9) and growth rate of fixed asset investment in infrastructure (X10) were studied.

(2) Research methods and model construction

At present, there are principal component analysis, analytic hierarchy process, cluster analysis and Delphi method [1]. However, the evaluation of economic development level is a project, so it is particularly important to select suitable evaluation indicators, because some indicators have little correlation and cannot obtain effective data, while some indicators have great correlation and can obtain valuable reference information. When multiple variables were studied with statistical method, the different variables is complex, and principal component analysis (pca) is on the basis of the original variables, delete not important variables, establish new variables as little as possible, and were not associated with these two variables, and these new variables in reflect the information as much as possible

to keep the original information. Therefore, in this paper, principal component analysis (PCA) was used to investigate the correlation between multiple variables of provincial and regional economy in China, which was beneficial to reveal the regularity of variables of economic development level among 31 provinces and cities, and to find out the deficiencies in the development of each province in the comparison, in order to make relevant policy suggestions for government departments.

Principal Component Analysis (PCA) is a multivariate statistical method as well as dimension reduction Analysis. That is, the index data with certain correlation is grouped into one category, and the index without correlation is separated, and the multi-index and non-repeating information is converted into several principal components of multivariate statistical method [2]. After the transformation of variables through orthogonal transformation, the variables that may be related in the original are no longer related. This group of variables after transformation is called principal components, which aims to reflect as much information about the original variables as possible. In this coordinate system, variance represents the information content of this component. It is generally believed that the first principal component is the one with the largest variance value of the index. On the premise of little information loss, the index data of several principal components after linear transformation are analyzed to see whether there is correlation, so as to simplify the problems to be studied and improve analysis efficiency. If the first variable does not represent all variables, continue to select the second, third, and fourth variables...The PTH principal component, until the number of selected variables can represent all the information of the original indicator, and if it can become the principal component, it is required that after the total variance of the data is interpreted, the original variable must maintain more than 85% of the information (more than 80% of the social science category).To sum up, the analysis model can be expressed as

$$
\left\{
\begin{array}{l}
F_1 = u_{11}X_1 + u_{12}x_2 + \cdots + u_{1j}X_j \\
F_2 = u_{21}X_1 + u_{22}x_2 + \cdots + u_{2j}X_j \\
\vdots \\
F_m = u_{p1}X_1 + u_{p2}x_2 + \cdots + u_{pj}X_j
\end{array}
\right\}
\tag{1}
$$

3. Empirical analysis of economic development level of Chinese provinces

In terms of GDP, Guangdong and Jiangsu take the lead, with Guangdong topping 10 trillion yuan in 2019, reaching 10.77 trillion yuan. Jiangsu ranks second, with its GDP reaching 9.96 trillion yuan, accounting for more than one fifth of China's total. It has become the two major regions leading the high-quality development of China's market economy [3]. Meanwhile, at the bottom of the list, Hainan province only has more than 500 billion yuan, a big gap with the top provinces. In terms of per capita disposable income, Shanghai and Beijing are the top two cities, each with a per capita disposable income of around 70,000 yuan, while provinces such as Yunnan province, which ranks lower, have a per capita disposable income of over 20,000 yuan. The common point is that each province has a relatively rapid economic development of the region, such as Guangdong Province's Pearl River Delta region; The difference is that each province's economic indicators vary widely. China has the problem of north-south differentiation in regional economic development [4].

(1) Applicability analysis of the original data

Before factor analysis, relevant data were input using SPSS 24.0 for dimensionality reduction to form KMO and Bartlett analysis results, where KMO value was 0.776 (> 0.600) and Bartlett sphericity test sig. (significance level) was 0. It can be seen that there are large correlation coefficients among the 10 index variables, as shown in Table 1, indicating that the original data meets the requirements of adaptability analysis and can be analyzed by factor analysis.

Table 1: KMO and Bartlett test

KMO sampling suitability quantity.		0.776
Bartlett sphericity test	The approximate chi-square	1059.202
	Degrees of freedom	45
	significant	0.000

(2) Standardization of original data

Because our country partial data of 31 provinces of difference is too big, and the data with different units and orders of magnitude, so, after the data standardization, data comparable units and various orders of magnitude, can use the principal component analysis (pca) analysis comparison, greatly improves the accuracy of data analysis.

(3) Extraction of principal components

Table 2: Common factor variance

	The initial	extract
Gross Regional Product (100 million yuan)	1.000	0.933
Permanent population urbanization rate	1.000	0.869
Permanent Population of each province (ten thousand)	1.000	0.912
Per capita disposable income of permanent urban residents (YUAN)	1.000	0.923
Per capita disposable income of permanent rural residents	1.000	0.915
General budgetary revenue of local finance (YUAN)	1.000	0.923
Total retail sales of consumer goods	1.000	0.927
Total household deposits of Provinces (YUAN)	1.000	0.490
Real estate Development fixed Asset investment growth rate (%)	1.000	0.933
Growth rate of investment in infrastructure and Fixed Assets (%)	1.000	0.344
Note: Principal component analysis was used for extraction		

Table 3: Total variance interpretation

composition	Initial eigenvalue			Extract the sum of squares of loads		
	A total of	Percentage of variance	Cumulative %	A total of	Percentage of variance	Cumulative %
1	5.008	50.085	50.085	5.008	50.085	50.085
2	2.104	21.042	71.127	2.104	21.042	71.127
3	1.056	10.562	81.689	1.056	10.562	81.689
4	0.829	8.291	89.980			
5	0.567	5.672	95.652			
6	0.187	1.869	97.521			
7	0.113	1.134	98.655			
8	0.077	0.768	99.423			
9	0.041	0.412	99.835			
10	0.017	0.165	100.000			
Extraction method: principal component analysis.						

As shown in Table 2, after factor analysis, except for the large loss of growth rate of investment in infrastructure fixed assets, the extraction rate of other variables is high, and the effect of factor extraction is relatively ideal. According to the shown in table 3, the main component of variance is greater than 1 index data, if extract principal component 1 or less, then can't apply the principal component analysis, only extract principal component number > 1, can use the principal component analysis, therefore, this data can be extracted to table 3 3 principal components, and the former three principal component accumulation accounted for 81.689% of the total variance, The extraction conditions of principal components were satisfied.

(4) Calculate principal component score, comprehensive score and ranking of each city

Table 4: Ranking of composite scores

provinces	F1	F2	F3	F	ranking
Guangdong	5.71	2.14	0.02	3.31	1
Jiangsu	4.65	0.47	0.39	2.47	2
Shandong	2.78	1.68	0.04	1.74	3
Zhejiang	3.61	0.71	0.66	1.73	4
Henan	1.14	2.15	0.71	0.95	5
Shanghai	3.13	3.67	1.15	0.92	6
Sichuan	0.86	1.71	0.18	0.77	7
Beijing	2.52	3.57	0.69	0.58	8
Hebei	0.54	1.12	0.99	0.4	9
Anhui	0.21	1.04	0.21	0.3	10
Fujian	0.75	0.76	0.22	0.24	11
Hubei	0.43	0.4	0.63	0.23	12
Hunan	0.07	0.31	0.18	0.12	13
Liaoning	0.08	0.57	0.37	0.2	14
Jiangxi	0.55	0.41	0.29	0.22	15
Yunnan	1.17	1.08	0.38	0.32	16
Shaanxi	0.81	0.33	0.12	0.35	17
Chongqing	0.52	0.21	0.54	0.36	18
Guangxi	1.21	0.91	0.06	0.41	19
Shanxi	1.03	0.49	0.02	0.41	20
Tianjin	0.11	2.3	0.21	0.45	21
Guizhou	1.83	0.62	0.29	0.75	22
Heilongjiang	1.46	0.1	0.57	0.77	23
Ji Lin	1.67	0.16	0.2	0.85	24
Inner Mongolia	1.24	1.19	0.09	0.86	25
Xinjiang	1.99	0.09	0.16	1.03	26
Gansu	2.3	0.45	0.03	1.05	27
Tibet	3.39	0.91	3.36	1.15	28
Hainan	2.14	1.16	1.1	1.43	29
Qinghai	2.65	0.68	1.44	1.52	30
Ningxia	2.47	1.28	1.15	1.63	31

Based on the analysis of 10 index data of 34 cities in Guangdong and Jiangsu provinces in 2018,

principal component scores F1 and F2 of each city are obtained by principal component analysis method. Then, according to the ratio of variance contribution rate of each principal component in Table 3, multiply each principal component score. The comprehensive score of each region can be obtained (i.e. F=0.6131F1+0.2576F2+0.1293F3), and the calculation results are ranked in descending order, as shown in Table 4.

4. Countermeasures and suggestions for improving the level of economic development in China

(1) Continue to do a good job in epidemic prevention and related financial support

There is no doubt that the outbreak of COVID-19 has had a significant impact on our economy and the world economy. Ans as an exogenous shock variable, COVID-19 has had a significant impact on China's and the world's economic growth. The influence to economy mainly has macroscopical, medium view, microcosmic threefold mechanism [5]. At the macro level, it mainly refers to the overall impact of the epidemic on the social economy of each region, and refers to the impact of the epidemic on various macro indicators. The middle level is mainly the impact of the epidemic on specific industries and upstream and downstream industries. At the micro level, it is the direct and indirect economic impact of the epidemic on individuals. Contain the outbreak is the important precondition for economic development, only good as a whole the epidemic prevention and control work and the economic and social development, promote the medical and health system to further improve, intensify the vaccine development and application, and do a good job related financial support, the implementation of the support policy on the real economy, can get rid of the disease for a series of adverse impact of the economic operation in our country, Only in this way can we continue to optimize China's economic structure, promote China's economy to return to the normal track, and further release the potential and vitality of economic development. Therefore, on the basis of correctly judging the trend of the epidemic in the future, we should improve epidemic prevention measures, provide corresponding financial support, innovate and improve financial support methods, further optimize the policy coordination mechanism of various regions under abnormal economic conditions, and achieve a long-term balance between stable growth and risk prevention.

(2) improve the support force and driving force of the metropolitan area

Metropolitan area is a region formed by the core of megacities, megacities or large cities with strong radiation capacity, and the basic range of one-hour commuting circle or the radius of the economic radiation range of core cities. At present, there are about 30 cities in China that meet the criteria of being a city with a population of more than 3 million [6]. In our country has formed about 30 metropolitan areas under the background, during the "14th five-year plan", because of strengthening the planning of the existing metropolitan areas, to ensure that metropolitan areas can maintain high-quality economic development under the drive of the core cities. From the perspective of population, a city with more than three million permanent residents can be called a city, and the radiation radius of a city is about 100 kilometers. In order to maintain the high-quality economic development, during the "fourteenth Five-year Plan", we should make reasonable planning for the metropolitan area, and plan the future development direction and emphasis, so as to ensure that the metropolitan area can maintain the high-quality economic development driven by the core cities. Under the radiation effect of metropolitan circle, the small and medium-sized cities located in the radiation circle will also get many development opportunities, and will form their own economic development system.

Each province wants to give full play to its comparative advantage to improve its development speed and catch up with other provinces. However, in reality, due to China's vast territory, different regions tend to form a pattern of differentiated development, and unified economic progress is difficult to achieve. So a province alone tend to be more difficult, the advantage of development in the province, in this case, can take advantage of the city circle adjacent coupling characteristics, implement two city circle economic and communication resources, expand the scale of cities, urban structure more reasonable, to strengthen the exchanges and cooperation with neighboring provinces and cities, and use of the resources of neighboring provinces and city circle, We will address the problem of unbalanced development. During the 14th Five-Year Plan period, we will optimize the structure of urban agglomerations, complement each other's advantages, give full play to the radiation power of urban agglomerations, drive the development of surrounding areas through industrial transfer, resource utilization, docking and cooperation, and enhance the level of integration.

(3) Do a good job in forest resources protection measures, and achieve win-win with economic development

As the main channel of ecological civilization construction, forest resources lay the foundation for green development in China. At the same time, the protection of forest resources is linked to the development of economy. In the process of modern economic development, the consumption structure has undergone great changes. While the quality of people's life is getting higher and higher, our demand for forest resources is also getting higher and higher. Forest resources are important economic materials, so we must strike a balance between reasonable protection and effective utilization [7]. In China's economic development, forest resources bring us a large number of green products, and the direction of economic development. However, the excessive exploitation of resources aggravated the destruction of the environment, and even appeared ecological environment crisis in some areas. In the background of increasing demand for forest resources, the supply of forest resources in China is getting less and less. In this situation, we must do a good job of forest protection measures, and increase the rate of forest cultivation, to alleviate the imbalance. Actively do the relevant protection measures, take the road of ecological, can help remote areas to upgrade the industry, promote the development of forest tourism. The so-called "green mountains are mountains of gold and silver", while cultivating forestry resources, but also to create conditions for the development of circular economy.

References

[1] Zheng Yang, Xuan Xue. Evaluation of urban competitiveness in Liaoning Province based on principal Component Analysis [J]. Journal of Shenyang University of Technology (Social Science Edition), 2017(3): 251-255.

[2] Nui Hua. Evaluation of Urban economic development level in Sichuan Province based on Principal Component Analysis [J]. Journal of Aba Normal University, 2018(4): 63-68.

[3] Chen Hongtao's Empirical Analysis of Regional Economic Development Level in Guangdong Province based on PCA

[4] Liu Yingjie Chinese new pattern of regional development of four questions to think http://www.rreca.com/xueke/ShowArticle.asp?ArticleID=3774

[5] Qin Yu: A survey of the challenges and impacts of COVID-19 on China's economy

[6] Liu Xiaojin into the difference of regional development of the four strategic direction http://www.rreca.com/xueke/ShowArticle.asp?ArticleID=3766

[7] Li Liangliang on How to Achieve a win-win situation between Forest Resource Protection and Economic Development. Protection Engineering, No.23, 2017