Analysis on The Status Quo of Social and Economic Development and Regional Economic Difference of Guangdong Province Based on Principal Component Analysis

Xiufang Yuan

School of Mathematics and Statistics, Sichuan University of Science & Engineering, China.

yxiufang1990@163.com

Abstract

This paper takes 21 prefecture-level cities in Guangdong Province as the research object, selects the relevant economic index to carry on the principal component analysis to the Guangdong Province economic development level, analyzes the Guangdong Province regional economic level difference.

Keywords

Guangdong Province, Principal Component Analysis, Regional economy.

1. Introduction

The economic strength of a region is determined by a variety of factors. The present situation of the social and economic development of the region can be reflected by the distribution of the region's GDP and industrial structure. The total economic volume of Guangdong Province is constantly increasing, by 2019, Guangdong's economy accounted for 1 / 8 of China's total. Based on the Statistical Yearbook of Guangdong, this paper selects the data of the relevant indicators of the economic development of Guangdong Province, analyzes the current situation of the economic development of Guangdong Province, and uses the principal component analysis 1-10 method to analyze the differences of the economic development level of 21 cities in Guangdong Province, and run a comparison.

2. Theory and Calculation Steps Principal Component Analysis

In order to deeply understand the difference of economic development level of 21 cities in Guangdong Province, this chapter uses principal component analysis, which is a multi-index method, to analyze and sort the basic indicators of economic development. This paper selects 21 ECONOMIC INDICATORS OF PREFECTURE LEVEL CITIES: First Industry added value, secondary sector of the economy added value, tertiary sector of the economy added value, fixed asset investment, per capita disposable income, per capita consumption expenditure, local fiscal revenue, and local fiscal expenditure, consumer Price Index. Using SPSS statistical analysis software, the results of economic development in Guangdong Province are obtained.

Principal Component Analysis (PCA) is an analytical method to simplify the complicated relations among the interrelated variables. The main idea of PCA is to reduce the dimension. Due to the different degree of correlation among the index variables, the main information can be extracted from the complex and varied indexes by linear combination, and expressed in a few factors, in other words, when the contribution rate of a few factors can explain most of the information in the original data, principal component analysis is successful.

(1) to standardize the data processing, the meaning of which is to eliminate the dimension of the data so that the data can be compared among different categories; (2) to calculate the correlation coefficient, which refers to the correlation among the variables of the indicators, it is meaningful to carry out principal component analysis only if the indexes are correlated to a certain extent; (3) calculating characteristics; (4) determining the principal components of this study, and in the process of carrying out principal component analysis, the cumulative contribution rate of each principal

component can be obtained, in general, the cumulative contribution rate should be more than 80% when choosing principal component; (5) synthesizing principal component; (6) calculating comprehensive score.

3. Results of principal component analysis

3.1 Validation of data

Table 3-1 KMO and Bartlett test results

A Kaiser-Meyer-Olkin metric of sufficient intensity		.707
Globular test of Bartlett's	Approximate chi-square	331.128
	df	36
	Sig.	.000

The KMO and Bartlit spherical tests determine whether the raw data is suitable for principal component analysis. The KMO statistic is 0.707 And 0 & Lt; 0.707 & Lt; 1, as can be seen from the results of Bartlit's spherical test. In the actual analysis, when KMO statistic is greater than 0.7, it is considered that the principal component analysis can get good results. The larger the KMO value, the better the results of principal component analysis, and the higher the overlap of the information carried by the variables.

3.2 Identify the principal component

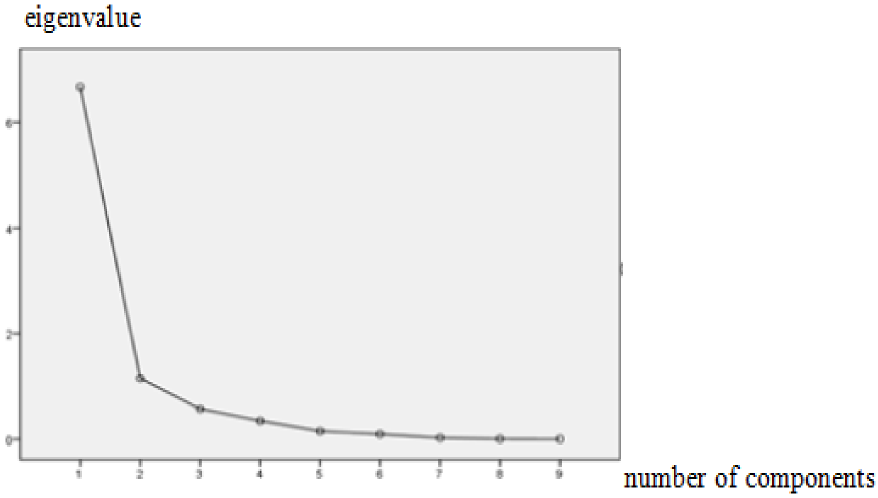

Fig. 3-1 Lithograph

Two common factors can be selected by extracting common factors according to the criterion that the characteristic root is greater than 1. The function of the lithograph is to judge whether the extraction standard is reasonable or not, and to indicate the importance of the overall information contribution of each factor. The horizontal axis is the factor ordinal and the vertical axis is the characteristic root size. The size of characteristic roots is arranged in order in the lithograph. The higher the value is, the more information the factor carries. According to the above gravel map obtained by SPSS software, it can be seen that the steep slope between the first factor and the third factor is larger, which corresponds to a larger characteristic root, and the effect is obvious, indicating that the first factor has most information of the original variable. The line behind is flat, indicating that the corresponding feature root is small, and the effect is weak, which means that the factor behind has little information about the original variable. There are only two points on the slope, and the other scattered points make up the flat Mesa, so only the first two factors are considered.

3.3 Explanation of total variance

Table 3-2 Total variance as explained

Composition	Initial eigenvalue			Extract Square sum load		
	Total	Variogram%	Accumulates%	Total	Variogram%	Accumulates%
1	6.669	74.097	74.097	6.669	74.097	74.097
2	1.155	12.830	86.927	1.155	12.830	86.927
3	.569	6.320	93.247			
4	.342	3.802	97.050			
5	.147	1.636	98.686			
6	.091	1.014	99.700			
7	.023	.255	99.955			
8	.003	.035	99.990			
9	.001	.010	100.00			
Extraction method: Principal Component Analysis						

According to the total variance interpretation table data, only two principal components have eigenvalues greater than 0.9, so two principal components are selected and recorded as F_1 and F_2 . According to the selection principle that the contribution rate is more than 80% , the cumulative contribution rate of eigenvalues of the first two principal components is 86.927% , which indicates that the nine variables can be reflected by the first two principal components.

3.4 Analysis of the results of the component matrix after rotation

According to the rotating composition Matrix, the first principal component F_1 supports tertiary sector of the economy X_3, fixed asset investment X_4, secondary sector of the economy X_2, local fiscal revenue X_7, local fiscal expenditure X_8, and per capita consumption expenditure X_6, the disposable income X_5 and consumer price indices X_9 are highly correlated with X_3, X_4, X_2, X_8, X_7. The second principal component F_2 is highly correlated only to the value added of the primary industry.

Table 3-3 rotation component Mattria

	Ingredients	
	1	2
x3	.954	.129
x4	.943	
x2	.939	.267
x8	.932	.179
x7	.914	.248
x6	.766	.522
x5	.758	.506
x1		-.951
x9	.563	.583
Extraction method: Principal Component. Rotation method: An orthogonal rotation method with Kaiser Standardization.		
a. Rotation converges after three iterations.		

3.5 Calculate composite scores

According to the percentage of variance corresponding to the first principal component and the second principal component multiplied by the scores of each factor, the comprehensive scores of 21 cities in Guangdong Province are calculated as follows:

$$y = 74.097 \times F_1 + 12.830 \times F_2$$

$$F_1 = 74.097 \times FAC1\_1$$

Table 3-4 factor load table

Region	FAC1_1	FAC2_2
Guangzhou	2.4962	-0.40742
Shenzhen	2.95499	0.88887
Zhuhai	-0.22522	1.64164
Shantou	-0.4007	0.11805
Foshan	0.87964	0.34951
Shaoguan	-0.61977	0.3559
Heyuan	-0.74957	0.40583
Meizhou	-0.54197	-0.25017
Huizhou	0.01977	0.03073
Shanwei	-0.72697	0.29304
Dongguan	0.14936	1.73982
Zhongshan	-0.28009	0.83464
Jiangmen	-0.19578	-0.25241
Yangjiang	-0.55499	-0.36227
Zhanjiang	0.27587	-2.28815
Maoming	0.1751	-2.05186
Zhaoqing	-0.1304	-1.22658
Qingyuan	-0.48916	0.01243
Chaozhou	-0.86274	0.8266
Jieyang	-0.44493	-0.36933
Yunfu	-0.72865	-0.28888

Table 3-5 ranking table of comprehensive scores by city

Region	F1	F2	Y
Guangzhou	218.955894	11.4042021	230.3600961
Shenzhen	184.9609314	-5.2271986	179.7337328
Zhuhai	65.17868508	4.4842133	69.66289838
Shantou	11.06712792	22.3218906	33.38901852
Foshan	-16.68812634	21.0622412	4.37411486
Shaoguan	1.46489769	0.3942659	1.85916359
Heyuan	20.44113939	-29.3569645	-8.91582511
Meizhou	-20.75382873	10.7084312	-10.04539753
Huizhou	12.9743847	-26.3253638	-13.3509791
Shanwei	-14.50671066	-3.2384203	-17.74513096
Dongguan	-9.6622488	-15.7370214	-25.3992702
Zhongshan	-29.6906679	1.5145815	-28.1760864
Jiangmen	-36.24528852	0.1594769	-36.08581162
Yangjiang	-32.96797821	-4.7385039	-37.70648211
Zhanjiang	-45.92309769	4.566197	-41.35690069
Maoming	-40.15835109	-3.2096811	-43.36803219
Zhaoqing	-41.12309403	-4.6479241	-45.77101813
Qingyuan	-53.86629609	3.7597032	-50.10659289
Chaozhou	-55.54088829	5.2067989	-50.33408939
Jieyang	-63.92644578	10.605278	-53.32116778
Yunfu	-53.99077905	-3.7063304	-57.69710945

Using the factor load Table 3-4, according to the comprehensive score calculation formula, 21 cities in Guangdong Province score and ranking table as shown in Table 3-5.

According to the data of the comprehensive score table, it can be seen that the city with the highest comprehensive score in Guangdong Province is Shenzhen City, with the overall score as high as 230.36, followed by Guangzhou City, and the two largest first-tier cities in Guangdong province both have comprehensive scores above 150, than any other city in the country. The top five cities in the combined score are all Pearl River Delta Economic Zone. The five cities with the lowest scores were Western Guangdong Yangjiang, eastern Guangdong Chaozhou and Shanwei, northern Guangdong Heyuan, Meizhou and Yunfu. The lowest level of economic development in north Guangdong, Yunfu, Guangdong. The future economic and social development of Guangdong Province still depends to a large extent on the Pearl River Delta Economic Zone, which has a high overall score. There are two major first-tier cities, and the total economic volume of Pearl River Delta accounts for an overwhelming proportion of the total economic volume of the province, with the exception of the Pearl River Delta, the economic performance of other regions is weak. Obviously, the economic development of Guangdong Province has high regional differences and uneven development level.

4. Conclusion

According to the characteristics of economic development in Guangdong Province and the analysis of regional economic disparities by Principal Component Analysis (PCA) , the GDP of Guangdong Province has been rising year by year, and the Pearl River Delta (PRD) accounts for most of the GDP, showing very significant regional differences. From the perspective of economic and industrial structure, the industry of Guangdong Province has changed from the original "2,3,1" type of industrial structure to the "3,2,1" type, and the population industrial structure type is also the "3,2,1" type, it shows that the industrial structure of Guangdong Province is relatively reasonable. In terms of regional differences, the economic level of Pearl River Delta is high, of which Shenzhen, Guangdong, and Guangzhou's economic comprehensive strength is strong, and far higher than other prefecture level cities.

Acknowledgements

This paper was financially supported by fundation: Key Laboratory of Bridge Nondestructive Testing and Engineering Calculation (2019QYY03).

References

[1] Physics -Chemical Physics; Reports from Sorbonne Universite Provide New Insights into Chemical Physics (Using principal component analysis for neural network high-dimensional potential energy surface) [J]. Network Weekly News, 2020.

[2] Signal Processing; Findings from Andhra University Provides New Data about Signal Processing (An Optimized Blind Watermarking Scheme Based On Principal Component Analysis In Redundant Discrete Wavelet Domain) [J]. Journal of Engineering, 2020.

[3] Hazell, Rinner. The impact of spatial scale: exploring urban butterfly abundance and richness patterns using multi-criteria decision analysis and principal component analysis[J]. International Journal of Geographical Information Science, 2020, 34(8).

[4] Cotta, H.H.A., Reisen, V.A., Bondon, P. et al. Identification of Redundant Air Quality Monitoring Stations using Robust Principal Component Analysis. Environ Model Assess 25, 521–530 (2020).

[5] Khaoula Ghoulem, Tarek Kormi, Nizar Bel Hadj Ali. Damage detection in nonlinear civil structures using kernel principal component analysis. 2020, 23(11):2414-2430.

[6] Zinan Zhao, Shijie Li, Shuaikang Li. Optimal batching plan of deoxidation alloying based on principal component analysis and linear programming. 2020, 3(2)

[7] Environmental Atmospheric Research; Tsinghua University Reports Findings in Environmental Atmospheric Research (Locating Unknown Number of Multi-point Hazardous Gas Leaks Using Principal Component Analysis and a Modified Genetic Algorithm)[J]. Journal of Mathematics, 2020.

[8] Remote Sensing; Researchers at Aeronautics Institute of Technology Target Remote Sensing (Change Detection in UWB SAR Images Based on Robust Principal Component Analysis)[J]. Journal of Technology, 2020.

[9] Hongtao Li, Ersegun Deniz Gedikli, Raed Lubbad. Exploring time-delay-based numerical differentiation using principal component analysis[J]. Physica A: Statistical Mechanics and its Applications, 2020, 556.

[10] Machine Learning; Study Data from School of Information Technology and Engineering Update Understanding of Machine Learning (An Empirical Model In Intrusion Detection Systems Using Principal Component Analysis and Deep Learning Models)[J]. Journal of Robotics & Machine Learning, 2020.