Application of Factor Analysis in the Evaluation of Economic Development Levels of Cities in Shandong Province

Manli Zhang, Guangli Xu , Yanyan Zhu, Jiaojiao Zhao
TaiShan University, Tai'an 271000, China.

Abstract: Shandong province is the most populous province in China and at the same time a major economic province. Many aspects of the economic development of Shandong province are very similar to the economic development of China and are highly representative, so it is necessary to study the level of economic development in Shandong province. Nowadays, many fields of research are very comprehensive, i.e. we can observe a lot of variables, but in the actual analysis of research it is difficult to analyse so many variables, so it is usually necessary to use fewer variables to reveal the problem through transformation, which is the problem of dimensionality reduction, the method of dimensionality reduction in multivariate statistics has the method of principal component and factor analysis, principal component is mainly to reduce the dimensionality while factor analysis not only This paper focuses on the use of factor analysis to analyse the variables in Shandong Province. In this paper, we use factor analysis to evaluate the level of economic development in Shandong Province, with data from the Shandong Provincial Bureau of Statistics website in 2021.

Keywords: Factor Analysis; Shandong Province; Evaluation of Economic Development Level

1. Introduction

Shandong Province is a province with a large population and a large economy in China. Although the economy of Shandong Province has been developed vigorously with the reform and opening up, there is an imbalance in the economic development among the cities. In this paper, we study 16 cities in Shandong Province and select 7 indicators, namely gross domestic product, real estate investment, disposable income, consumption expenditure, service enterprises, the number of domestic patent applications received by each city, and consumption expenditure of rural residents, with specific data from the website of Shandong Provincial Bureau of Statistics.

2. The Factor Analysis Method

The idea of factor analysis was developed by Charles Spearman [1] in 1904 in relation to the study of student performance. The idea of factor analysis is to find the similarities between many variables and to express these similarities as new variables, called factors. It is a statistical analysis method that converts complex variables into a few composite factors. It is different from principal component analysis: principal component analysis uses the original variables to represent new variables that are not related to each other, whereas factor analysis looks for common factors that can explain the similarity of the original variables and uses the factor variables to represent the original variables. The number of factors in factor analysis is usually determined in advance, and factor analysis can be rotated to make the results easier to interpret.

The core components of factor analysis are finding the common factors and loading matrices, factor rotation and calculating factor scores. The most common methods used to solve for the common factors and loading matrices are the principal component method, the great likelihood method and the principal axis factor method, which are used in this thesis.

The face gives the model analysis of factor analysis [2]: first we have a premise that there are some samples, each sample observes certain indicators, and there is a strong correlation between the indicators, usually the magnitude between the variables is different, so

we first the magnitude for processing [1], that is, the observation data for standardization, first set the original variable with, this is the variable studied in the actual research problem, where it has a vector mean of 0, we set $F=(F_1,F_2\cdots F_P)'$ are unobservable variables, which are factor variables to be determined, and the factors are unrelated to each other, i.e. the components are independent of each other, ε with F Mutually independent,Then the matrix form of the model is as follows:

$$X=AF+\varepsilon \quad \text{that is} \quad \begin{cases} X_1=a_{11}F_1+a_{12}F_2+\cdots a_{1m}F_m+\varepsilon_1 \\ X_2=a_{21}F_1+a_{22}F_2+\cdots a_{2m}F_m+\varepsilon_2 \\ \quad\vdots \\ X_p=a_{p1}F_1+a_{p2}F_2+\cdots a_{pm}F_m+\varepsilon_p \end{cases}$$

The common factors are independent of each other and are not measurable. The meaning of the common factors must be generalised to the actual problem. In practical problems we can express the factors in terms of variables by regression and weighted least squares to obtain factor scores. Often the initial loading matrix obtained is not unique and in practice the common factors are often rotated to make it easier to analyse their actual meaning.

3. Example application of the factor analysis method

We extracted seven variables for 16 cities in Shandong Province from the data downloaded from the website of the Shandong Provincial Bureau of Statistics for the description of the economic situation in Shandong Province. These seven variables are gross domestic product (billion yuan), real estate investment (billion yuan), disposable income (yuan/person), service enterprises (number), the number of domestic three kinds of patent applications received by each city, and consumption expenditure of rural residents (yuan/person). Factor analysis was conducted for the above seven variables on the economic situation of Shandong Province.

Firstly, we used the principal component method to conduct the factor analysis, and we obtained the loadings of the rotated factors as shown in Figure 1: from Figure 1, we can see that Factor 1 has a strong positive correlation with variables such as gross domestic product, real estate investment, service enterprises, and the number of domestic patent applications received by each municipality, so we named it the "welfare factor". Factor 2 has a strong positive correlation with disposable income, consumption expenditure and rural consumption expenditure, and is therefore named the "consumption factor".

Rotational factor load

	Factor 1	Factor 2
Gross Domestic Product (100 million yuan)	0.9040304	0.3652993
Real estate investment (100 million)	0.9590179	0.2630890
Disposable income (yuan/person)	0.4221624	0.8499133
Consumption expenditure (yuan/person)	0.4248894	0.8929596
Service enterprises (number)	0.9491288	0.2876503
Number of domestic patent applications accepted in each city	0.9482142	0.2917367
Rural residents' consumption expenditure (yuan/person)	0.1292484	0.8955065

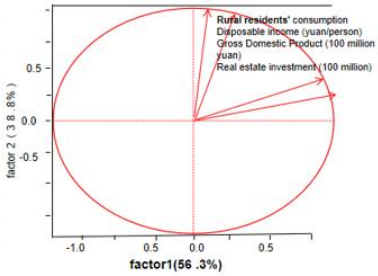

Figure 1: Factor loadings

Next we look at the factor loadings obtained using the principal axis factor method as shown in Figure 2: from the figure we can see that the results are the same as those for the principal components so we take the same two factors as above.

Rotational factor load

	Factor 1	Factor 2
Real estate investment (100 million)	0.9590179	0.2630890
Service enterprises (number)	0.9491288	0.2876503
Number of domestic patent applications accepted in each city	0.9482142	0.2917367
Gross Domestic Product (100 million yuan)	0.9040304	0.3652993
Rural residents' consumption expenditure (yuan/person)	0.1292484	0.8955065
Consumption expenditure (yuan/person)	0.4248894	0.8929596
Disposable income (yuan/person)	0.4221624	0.8499133

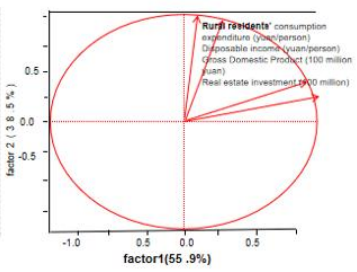

Figure 2: Factor loadings for the principal axis factor approach

We can further derive the factor scores to analyse the relationship between the cities and the two factors, and we have ranked the two factors in descending order as shown in Table 1; from Table 1 we can see that the top two rankings according to factor 1, the 'welfare factor', are Qingdao and Jinan, which are the two largest cities in Shandong Province and are This is in line with the reality. The top two cities according to factor 2, the "consumption factor", are Dongying and Zibo, which is not unrealistic either. Readers can look at the specific rankings and give specific analyses according to the cities they study.

Table 1 Factor scores

Region	In descending order of factor 1 score	Region	In descending order of factor 2 score
Qingdao	2.551748081	Dongying	1.629984208
Jinan	2.12426855	Zibo	1.319592022
Linyi	0.45752937	Yantai	1.255728336
Weifang	0.246798368	Weihai	1.193485861
Jining	0.09739135	Qingdao	0.917199528
Yantai	0.011404397	Jinan	0.206210965
Heze	-0.243356651	Weifang	0.199344239
liaocheng	-0.243958946	Tai'an	-0.097177541
Rizhao	-0.268748932	Binzhou	-0.106449776
Dezhou	-0.426026218	Dezhou	-0.503610789
Zaozhuang	-0.486858812	Jining	-0.672071795
Tai'an	-0.582791758	Zaozhuang	-0.720777143
Binzhou	-0.635914849	Heze	-0.929734269
Zibo	-0.691255713	Rizhao	-1.149224084
Weihai	-0.813060319	liaocheng	-1.187057263
Dongying	-1.097167918	Linyi	-1.355442499

4. Summary

In this paper, we have introduced the application of factor analysis in the evaluation of the economic development level of cities in Shandong Province, mainly using the principal factor method and the principal component method. The factor scores show the differences between cities and are backed up by actual data, so factor analysis is useful for evaluation purposes.

References

[1] He XQ, Ma XJ. Multivariate statistical analysis - based on R language [M]. Beijing: People's University of China Press, 2021.

[2] Wu XZ. Multivariate Statistical Analysis [M]. Beijing: People's University of China Press, 2019.

Author: Manli Zhang (1985.09---) Mujer, Nacionalidad Han, Tai'an City, Shandong Province, Estudiantes de posgrado,Research direction: Applied statistics

Fund: TaiShan University Teaching Reform and Research Project (Project number: 4101210116)