ISSN 2522-6398 Vol. 4, Issue 4: 99-102, DOI: 10.25236/FER.2021.040420

A Study of Higher Education Factors Based on Principal Component Analysis

Bingxue Li[1], Jinqi Wang[2], Jiarui Zhang[3]

[1]College of Marine Electrical Engineering, Dalian Maritime University, Dalian, Liaoning 116026, China
[2]College of Public Administration and Humanities and Arts, Dalian Maritime University, Dalian, Liaoning 116026, China
[3]College of Science, Dalian Maritime University, Dalian, Liaoning 116026, China

Abstract: *In recent years, the higher education system occupies a higher and higher position in all countries. Our goal is to measure and evaluate the health of the higher education system at the national level, to determine a healthy and sustainable state for the higher education system of each country, and to propose and analyze a set of policies. And verify the feasibility and importance of the proposed policies in improving the health status of higher education systems in various countries. For the factors that affect the health status of the ten higher education systems in China, we have established a principal component analysis model to screen out four main factors, which can be roughly divided into two categories-the amount of money invested in education and the number and quality of teachers. In view of these two kinds of influencing factors, we put forward a realizable and reasonable vision for China's higher education system. Next, we use our model to measure China's current system and the proposed system respectively. By comparison, we can see that the health level of China's higher education system has risen from the third level to the second level. It is clear that our proposal is very effective.*

Keywords: *Principal component analysis, Higher education system, Policy, Health Levels*

1. Introduction

The higher education system is an important part of a country, and each country's higher education system has its strengths and weaknesses. After adjustments need to be made during the current pandemic, countries have the opportunity to reflect on what is effective and which can be done better. However, change is often difficult.

In order to measure and evaluate the health of higher education systems at home level, and to determine a healthy and sustainable state for higher education systems in some countries, we need to develop models to analyze and solve some problems.

2. Principal component analysis model

2.1 Analysis of ideas

Aiming at the selected country-China, we use principal component analysis to model and analyze the higher education data in China, eliminate the indicators with less impact, reduce the application of these indicators, and determine the indicators with greater impact. Combined with the characteristics of China's development, we can properly consider strengthening the indicators with greater impact, further achieve its scope of influence and make a breakthrough in the education system. Finally, a realizable and reasonable vision is put forward according to the development of China.

ISSN 2522-6398 Vol. 4, Issue 4: 99-102, DOI: 10.25236/FER.2021.040420

2.2 Symbol description

Symbol	Description
y_i	The ith evaluation object
b_{ij}	J-th index of the i-th evaluation
$\overline{b_{ij}}$	Standardized index value
v_j	Sample mean of the j th index
g_j	Sample standard deviation of the j index
$\overline{\overline{y_j}}$	Standardized index variable
r_{ij}	The correlation coefficient between the I index and the j index
λ_j	Eigenvalue
c_j	Information contribution rate
b_q	Cumulative contribution rate
H	Comprehensive score

2.3 Modeling and solution of Principal component Analysis

There are 10 index variables for principal component analysis, which are $y_1, y_2, \ldots, y_{10}$, there are 10 evaluation objects, and the value of the j index of the i evaluation object is b_{ij}. Convert each indicator value b_{ij} into a standardized indicator value $\overline{b_{ij}}$ indicator value, such as:

$$\overline{b_{ij}}=\frac{b_{ij}-v_j}{g_j},\ i=1,2,\ldots,10;\ j=1,2,\ldots,10$$

1) In the formula: $v_j=\frac{1}{10}\sum_{i=1}^{10} b_{ij}$; $g_j=\sqrt{\frac{1}{9}\sum_{i=1}^{10}(b_{ij}-v_j)^2}$, j=1, 2, …10, that is v_j, g_j is the sample mean and sample standard deviation of the j th index.

Correspondingly, called:

$$\overline{\overline{Y_j}}=\frac{y_j-v_j}{g_j},\ j=1, 2, \ldots, 10$$

is a standardized index variable.

2) The correlation coefficient matrix R is calculated. Correlation coefficient matrix $R=(r_{ij})_{m\times m}$:

$$R_{ij}=\frac{\sum_{k=1}^{n}\overline{b_{ki}\cdot b_{kj}}}{9},\ i, j=1, 2, \ldots 10$$

In the formula: $r_{ii}=1$, $r_{ij}=r_{ji}, r_{ij}$ is the correlation coefficient between the i index and the j index.

3) Calculate eigenvalues and Eigenvectors. Calculate the eigenvalue $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{10} \geq 0$ of the correlation coefficient matrix R, and the corresponding eigenvector $v_1, v_2, \ldots, v_{10}$,, in which $v_j = [v_{1j}, v_{2j}, \ldots, v_{10j}]^T$, which is composed of 10 new index variables:

$$z_1 = v_{11}\bar{y}_1 + v_{21}\bar{y}_2 + \cdots + v_{10\,1}\bar{y}_{10},$$
$$z_2 = v_{12}\bar{y}_1 + v_{22}\bar{y}_2 + \cdots + v_{10\,2}\bar{y}_{10},$$
$$\vdots$$
$$z_{10} = v_{1\,10}\bar{y}_1 + v_{2\,10}\bar{y}_2 + \cdots + v_{10\,10}\bar{y}_{10}$$

4) Select q ($q \leq 10$) principal components and calculate the comprehensive evaluation value.

① Calculate the eigenvalue λ_j (j=1,2,…,10) information contribution rate and cumulative contribution rate,called:

$$c_j = \frac{\lambda_j}{\sum_{k=1}^{10}\lambda_k},\ j=1,2,\ldots, 10$$

The information contribution rate of the principal component z_j, at the same time, there are

$$b_q = \frac{\sum_{k=1}^{q}\lambda_k}{\sum_{k=1}^{10}\lambda_k}$$

ISSN 2522-6398 Vol. 4, Issue 4: 99-102, DOI: 10.25236/FER.2021.040420

The main ingredient z_1, z_2 ,… ,z_q, the cumulative contribution rate of b_q,. When b_q is close to 1 (usually take $b_q$bq=0.85,0.90,0.95), then select the first q index variables z_1, z_2,… ,z_q, as Q principal components, replaces the original 10 index variables, so that q principal components can be analyzed comprehensively.

② Calculate the comprehensive score:

$$H=\sum_{j=1}^{q} c_j z_j$$

In the formula: cj is the information contribution rate of the j principal component, which can be evaluated according to the comprehensive score.

5) The principal component analysis of 10 evaluation indexes is carried out by using Matlab software. the first eigenvalues of the correlation coefficient matrix and their contribution rates are shown in the table.

Table 1: Correlation matrixte

Serial number	Characteristic root	Contribution rate	Cumulative contribution rate
1	8.9256	89.2558%	89.2558%
2	0.5833	5.8329%	95.0887%
3	0.2193	2.1930%	97.2817%
4	0.1121	1.1214%	98.4031%
5	0.0680	0.6796%	99.0827%
6	0.0389	0.3889%	99.4717%
7	0.0321	0.3209%	99.7925%
8	0.0201	0.2007%	99.9932%
9	0.0007	0.0067%	100.0000%
10	0.0000	0.0000%	100.0000%

It can be seen that the cumulative contribution rate of the first four characteristic roots is more than 90%, and the effect of principal component analysis is very good. Below, the first four principal components (The cumulative contribution rate reached 98.430%) are selected for comprehensive evaluation. The eigenvectors corresponding to the first four eigenvalues are shown in the table.

Table 2: Feature vector scale

	$\overline{x_1}$	$\overline{x_2}$	$\overline{x_3}$	$\overline{x_4}$	$\overline{x_5}$	$\overline{x_6}$	$\overline{x_7}$	$\overline{x_8}$	$\overline{x_9}$	$\overline{x_{10}}$
1	0.3181	-0.3511	0.0368	-0.0545	0.4135	-0.2141	-0.0490	0.7400	0.0477	-0.0463
2	0.3208	-0.3192	0.0100	-0.1392	0.2547	0.5356	0.3075	-0.2197	-0.4128	0.3259
3	0.2905	0.5925	-0.2463	0.0275	0.5183	-0.3190	-0.0592	-0.1983	-0.2975	-0.0278
4	0.3105	-0.4462	0.1971	-0.1165	0.1261	-0.2211	-0.3567	-0.5547	0.1719	-0.3505

2.4 A reasonable vision

The four main factors we have selected can be divided into two categories: the funds invested by the state in education and the quantity and quality of teachers.

In this regard, our reasonable vision for China's higher education system is as follows:

(1) It is hoped that China will spend more than 8 per cent of its GDP on education in the future.

(2) It is hoped that there will be more and more teaching staff in China in the future, and the treatment of teaching staff will be better and better.

(3) It is hoped that the quality of China's future teaching staff will be greatly improved and that they will be able to teach better students.

2.5 Policy

Based on the analysis of the conclusions of the above model, we have formulated the following policies to solve the problem of the health status of China's higher education system:

(1) Pay attention to the advantages and limitations of higher education research brought byregional differences, increase financial support and resource investment for higher education in poor areas, and provide necessary material support for higher education innovation in poor areas.

ISSN 2522-6398 Vol. 4, Issue 4: 99-102, DOI: 10.25236/FER.2021.040420

(2) Through the establishment of a community of academic exchanges and collaborative development, we can promote the balanced development of higher education in different regions and provinces.

(3) Fully tap the educational potential of vulnerable colleges and universities, comprehensively consider the relationship between educational input and output, pay attention to the input of educational human resources, innovate the talent training system, and strengthen the cultivation of core educational groups and emerging scientific research forces. to provide a continuous supply of human resources for the production of higher education, and strive to occupy a place in the production of higher education in the country by integrating academic resources and cultivating innovative teams.

(4) Enrich the research theme category of higher education, coordinate the distribution proportion of scientific research projects on various topics, create a more professional and original principle system of higher education, and on the basis of drawing lessons from the development experience of foreign higher education, strengthen the study of the basic theories of higher education and establish a discourse system of higher education research with unique Chinese characteristics.

3. Conclusion

(1) For students: more funds, better facilities and environment, higher learning experience and fun, and more choices for students.

(2) For teachers: due to the increase in the strength of teachers and the number of teachers, the burden of each teacher becomes smaller, and the treatment of each teacher is improved.

(3) For schools: due to the increase of state investment in education, the facilities and environment of schools become better, which makes the teaching methods more rich and colorful.

(4) To the country: as the country's investment in education increases, the country's burden may become heavier, but the country's future will become stronger and stronger.

References

[1] Zhuo Jinwu. The application of MATLAB in mathematical modeling [M] Beijing University of Aeronautics and Astronautics Press, 2014-9, 137-159
[2] Jiang Qiyuan. Mathematical model [M] Higher Education Press, 2011, 249-260
[3]Si Shoukui .Mathematical modeling algorithm and Application [M] National Defense Industry Press, 2015, 231-239
[4] Ma Lichao. Analysis and development vision of scientific research projects of higher education in China [M] National Defense Industry Press, 2018, 106-117