ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0061

Quantitative Evaluation Model based on Linear Regression

Wenjing Liang, Jingli Wang, Haoyu Guo

School of Artificial Intelligence. North China University of Technology, Tangshan 063210, China

Abstract

Considering that countries have different levels of development, there will be different criteria for evaluating the degree of fair distribution among countries. In order to accurately adopt a policy of fair distribution for any country, we have studied a set of indicators and additional parameters. A linear regression model capable of measuring distributive equity in any country. First, a total of 16 representative indicators in two categories were selected. After data preprocessing, a factor analysis model was established to extract three comprehensive indicators: government support, scientific and technological development level, and national contribution, with Topsis score as the dependent variable. , the comprehensive index is the independent variable to perform regression fitting, and finally the quantitative equation model is obtained.

Keywords

Linear Regression; Mineral Resources; Comprehensive Evaluation; TOPSIS.

1. Introduction

1.1 Background

Exploring changes in the asteroid mining sector by developing and implementing analytical methods could have different implications for global stock markets.

1.2 Our Work

In view of the changes in the asteroid mining sector, taking into account the different development levels of various countries, linear regression processing is performed using the index parameters in question 1.2. First, a total of 16 representative indicators in two categories were selected, and after data preprocessing, a factor analysis model was established to extract three comprehensive indicators, government support, technology development level, and national contribution. The index is the independent variable to perform regression fitting, and finally the quantitative equation model is obtained.

2. Model Assumptions

It is assumed that the data collected in this paper can fully reflect the state of human society. The data collected in this paper can explain the state of human society, and then the establishment of the model can explain global equity and measure global equity.

3. Notation

Symbols	Definition	Units
A	Population	10000
B	GDP per capita	Dollar

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0061

C	Energy consumption per unit of GDP	USD/kg
F	Import of ores and metals	%
S	Forest cover rate	%

4. Modeling

4.1 Factor Analysis to Extract Principal Component Indicators

In order to avoid information duplication and confusion among the 15 indicators in questions 1 and 2, resulting in data duplication, etc., this paper uses factor analysis to extract comprehensive indicators.

4.2 KMO, Bartlett Sphericity Test

This paper uses KMO and Bartlett sphericity test to test the original data to verify whether the original data can use the factor analysis model. The results show that the original data can be used for analysis within the 90% confidence interval.Equation Chapter 2 Section 2

4.3 The Establishment of Factor Models

The relationship between the 16 indicators and the common factors is constructed as follows:

$$x_1 = u_1 + a_{11} f_1 + a_1 f_2 + \cdots + a_{1m} f_{1m} + \varepsilon_{16}$$
$$\vdots$$
$$x_{16} = u_{16} + a_{161} f_1 + a_{162} f_2 + \cdots + a_{16m} f_{16m} + \varepsilon_{16}$$

aij represents the covariance between the original variable xi and the common factor fi, and aij is used to construct the A matrix.

By calculating the row element square sum hi2 and the column element square sum gi2 of the A matrix respectively, the dependence of the original variable xi on the common factor and the contribution of the common factor fj to the overall index data x are obtained, and then the most important comprehensive factor is selected.

where the sums of squares of row and column elements are:

$$h_i^2 = \sum_{j=1}^{m} a_{ij}^2 g_j^2 = \sum_{i=1}^{p} a_{ij}^2$$

4.4 Determining the Number of Factors

In this paper, three common factors are extracted through the turning point of the gravel diagram and the cumulative contribution rate of the variance explanation table. The variance contribution rates of the three extracted common factors are: 80.839%, 10.029%, and 6.156%. The cumulative variance contribution rate of the factors has reached 97.024%, that is, the first three common factors can contain 97.024% of the information of the original index, so the first three common factors are selected to reflect the overall index data.

4.5 Factor Score

Factor analysis is to express variables as a linear combination of common factors and special factors. In this paper, the common factor can be expressed as a linear combination of the original variables in reverse, and the factor score can be obtained.

According to the score coefficient matrix, the expressions of the three principal components are:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0061

$$f_1 = 0.185x_1 + 0.108x_2 + 0.237x_3 + \cdots + 0.068x_{16}$$
$$f_2 = -0.124x_1 + 0.035x_2 - 0.132x_3 + \cdots + 0.021x_{16}$$
$$f_3 = 0.163x_1 - 0.126x_2 + 0.125x_3 + \cdots + 0.056x_{16}$$

In the first principal component f1, the positive load coefficient of investing funds and holding activities is large, so the first principal component is called government support; in the second principal component f2, the positive load coefficient of the development level and talent contribution of colleges and universities is large, so it is called The second principal component is the influence of the level of science and technology; the third principal component f3 has a larger positive load coefficient of the state's participation, so the third principal component is called the state's contribution. The three principal components reflect the local education system more intuitively from the perspectives of government, technological development level, and national contribution. The data of the three principal components are obtained according to the above expressions.

4.6 Evaluation Quantitative Model based on Linear Regression

Multiple linear regression is used to predict the index coefficients, f1, f2, f3 are independent variables, y is the dependent variable, and the following linear relationship is satisfied:

$$y_i = \beta_0 + \sum \beta_i f_{ij} + u_i, i = 1, 2, \cdots, n; j = 1, 2, \cdots, p$$

where β is the regression coefficient, and u is the unobservable interference term that meets certain conditions.

Based on the above analysis, f1 is defined as the government's support, f2 is the influence of the level of science and technology, and f3 is the degree of national contribution. The constructed regression coefficient model is as follows:

$$y = \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \beta_3 f_3$$

The regression equation obtained by solving the index coefficients is as follows:

$$y = -0.021f_1 + 0.002f_2 + 0.018f_3 + 0.088$$

5. Sensitivity Analysis

Among the index weights determined by the entropy weight method, F1>F2>F4>F3>F6>F5, the comprehensive population component has the greatest impact on the final score, the larger the population base, the greater the corresponding demand for mineral resources, it should be placed in an important position. For example, populous countries such as China and India have higher scores and obtain more mineral resources. The weight coefficient of the comprehensive economic component is second only to the comprehensive population component, and the economically active areas have a higher demand for mineral resources. For example, developed regions such as the United States, the United Kingdom, France, and Spain have higher scores and obtain more mineral resources. The indicators of the integrated mineral composition and the integrated environmental composition have lower weights, so the Argentina and Kenya regions have low scores. For example, Argentina is rich in mineral resources, and the mineral resources are basically self-sufficient, and the demand for mineral resources is not high; the environment in Kenya is poor and lacks the prerequisites for mineral resource smelting.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202207_8(7).0061

Table 1. Positive ideal solution and negative ideal solution

item	*Positive solution*	*negative ideal solution*
ZF1	0.70112	2.78E-05
ZF6	0.612727	1.91E-05
ZF2	0.646239	2.21E-05
ZF5	0.39864	1.29E-05
ZF3	0.598517	2.04E-05
ZF4	0.896471	2.61E-05

The scores of each country and region are in the interval [0,1]. The larger the score S, the closer the evaluation object is to the positive ideal solution; the smaller the score S, the closer the evaluation object is to the negative ideal solution.Equation Chapter 4 Section 4.

The comprehensive evaluation model can be expressed as:

$$y = f(F1, F2, F3, f4, F5, F6)$$

Keep the same observation except F1, let F2=0.64, F3=0.60, F4=0.90, F5=0.40, F6=0.61, the change of comprehensive evaluation model score:

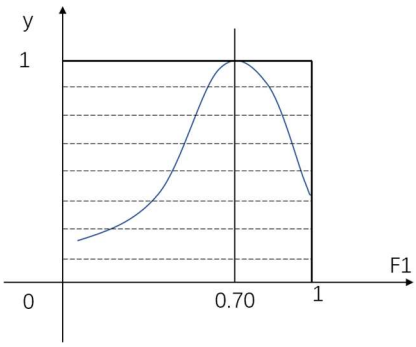

Fig. 1 Change curve of score y and principal component F1

Observing the above figure, you can find that the function is a convex function with a maximum value of 1, the peak value of the function is less than 1, and the defined interval is [0,1]. When F1=0.70, the function achieves the maximum value. The comprehensive evaluation model is a distance optimal problem, so it has an optimal value. The closer F1 is to 0.7, the greater the score y.

References

[1] Thuy B , Gale A S , Reich M . A new echinoderm Lagersttte from the Pliensbachian (Early Jurassic) of t he French Ardennes[J]. swiss journal of palaeontology, 2011, 130(1):173-185.

[2] Gu Cuimei, Li Li, Liu Yaqin, Li Yunpeng, Li Ailing, Wang Liting, Yang Dong. Mathematical Model of Gross Salary Distribution [J]. Chinese and Foreign Entrepreneurs, 2020(14):224-225.

[3] Xu Hui. The American "Deep Space Industry" mining company will launch an asteroid prospecting plan in 2020 [J]. Resources, Environment and Engineering, 2016, 30(06): 1033.

[4] Qiu Chengbo, Sun Yukun, Wang Yamin, Jiang Jun, Chen Xin. Development Status of Near-Earth Asteroid Mining and Defense Program [J]. Journal of Deep Space Exploration, 2019,6(01):63-72.DOI:10. 15982/j .issn.2095-7777.2019.01.010.