Compare the Principal Component Analysis and Linear Weighting Decision Method

Consumption of the urban residents in China

Keying Zhang
School of Management
Xi'an Polytechnic University
Shanxi, Xi'an, China
keerbin@163.com

Mi Tian
School of Management
Xi'an Polytechnic University
Shanxi, Xi'an, China
382744952@qq.com

Xiaoli Zhang
School of Management
Xi'an Polytechnic University
Shanxi, Xi'an, China
Xoly2008@xpu.edu.cn

***Abstract*—Based on the statistical information, this paper studies on the comprehensive evaluation of our consumption of the urban residents, compares principal component analysis and the linear weighting method. Compared the rank of them find their same point.**

Keywords-principal component analysis; composite scores; urban residents

I. INTRODUCTION

After 30 years of reform and opening up, China's economy continue to soar, the country's economic development level of each region has also been improved with each passing day that may be because the region by its own geographical and historical reasons, the impact of various factors, regional economic development uneven across the income of the residents are very different, the growing gap between rich and poor regions significantly. The structure of consumption is changing with the economic development. The consumption more focus on the spiritual level, therefore, the analysis results can be related to urban construction investment and provide reference information.

II. THE PRINCIPAL COMPONENT ANALYSIS OF CHINA URBAN HOUSEHOLD AVERAGE ANNUAL CONSUMPTION EXPENDITURE

Principal component analysis is to be correlated with the original number of target re-assembled into a new comprehensive index has nothing to do with each other instead of the original target of statistical methods. In this paper, the principal component analysis is based on the data for statistical analysis in "China Statistical Yearbook 2009" [1]:"the area of urban residents' per capita annual household consumption expenditure (2008)".

A. Establishment of Evaluation Index

Because of urban households the average annual consumer spending, so we select the "China Statistical Yearbook 2009" in "all areas of urban residents per capita annual household consumption expenditure (2008)" of data, data in the index with food, clothing, household equipment and services, health care, transportation and communication, education, culture and entertainment services, housing, miscellaneous goods and services [1]. Evaluation index of the contents is in Table 1.

TABLE I. CONSUMPTION EXPENDITURE INDEX NAME AND CODE

number	Name of indexes	number	Name of indexes	number	Name of indexes
1	Foods	4	Health Care	7	Live
2	Clothing	5	Transportation and Communications	8	Miscellaneous goods and services
3	Household equipment and services	6	Education, Culture and Entertainment		

B. Principal Component Analysis of Data

We analyze the data of "all areas of urban resident per capita annual household consumption expenditure (2008)," in "China Statistical Yearbook 2009" using SPSS software for data processing.

By the factor (FACTOR) analysis in SPSS, we can test the suitability of the index system, which is to test its suitability for factor analysis. KMO statistic test usually spherical Bartlett's test. Generally speaking, we can believe that as long as the KMO is greater than 0.5, which shows suitable for factor analysis, the KMO test value is 0.844. Bartlett's ball-type test

for testing whether the correlation matrix unit matrix, that is, whether the variables are independent. Test value is less than 0.01 to illustrate the independent variables [2]; this Bartlett's ball test is 0.000, so the index system is suitable for factor analysis.

TABLE II. TOTAL VARIANCE EXPLAINED

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.600	70.005	70.005	5.600	70.005	70.005	4.466	55.820	55.820
2	1.114	13.926	83.931	1.114	13.926	83.931	2.249	28.110	83.931
3	.527	6.583	90.514						
4	.284	3.545	94.059						
5	.203	2.537	96.596						
6	.162	2.024	98.621						
7	.068	.844	99.465						
8	.043	.535	100.000						

TABLE III. COMPONENT SCORE COEFFICIENT MATRIX

	Component	
	1	2
Food	.330	-.269
Clothing	-.187	.541
Household equipment and service	.150	.039
Health Care	.149	.064
Transportation and Communication	-.193	.557
Education,Culture and Entertainment	.288	-.169
Live	.167	.057
Miscellaneous goods and service	.156	.054

According to Figure 1, draw two main components, foods and clothing that can replace the other factors. The first two principal components analysis include information

on all the indicators that we have no need to calculate the other principal components .The total contribution rate of the first two components is 83.931%.

From the Figure 2, there are the two main components of the formula:

$$F_1 = 0.33 * X_1 - 0.187 * X_2 + 0.15 * X_3 + 0.149 * X_4$$
$$-0.193 * X_5 + 0.288 * X_6 + 0.167 * X_7 + 0.156 * X_8$$
$$F_2 = -0.269 * X_1 + 0.541 * X_2 + 0.039 * X_3 + 0.064$$
$$* X_4 + 0.557 * X_5 - 0.169 * X_6 + 0.057 * X_7 + 0.054 * X_8$$

The main component in all cities and composite scores is in Table 2.From this Table, we can get the rank of these areas. The areas in the front of the rank are the developed cities in eastern. Therefore, the economic development level of these areas has influent their consumption of the urban residents.

TABLE IV. CONSUMPTION EXPENDITURE INDEX NAME AND CODE

Area	**F1**	**F2**	**zf**	**rank**
Beijing	1.10739	2.64835	1.36	2
Tianjin	0.67238	1.02287	0.66	5
Hebei	-0.7164	0.37165	-0.3	17
Shanxi	-0.84729	0.33893	-0.38	23
Inner Mongolia	-0.56176	1.61355	0.14	9
Liaoning	0.03033	0.43061	0.14	9
Jilin	-0.68083	0.95363	-0.11	12
Heilongjiang	-1.13023	0.73554	-0.42	25
Shanghai	3.52295	0.0848	1.99	1
Jiangsu	0.47485	0.152	0.31	7
Zhejiang	1.16527	0.86538	0.89	4
Anhui	-0.24761	-0.49132	-0.28	16
Fujian	1.14619	-0.87057	0.4	6
Jiangxi	-0.3088	-0.878	-0.42	25
Shandong	-0.12028	0.84988	0.17	8
Henan	-0.79797	0.41304	-0.33	20
Hubei	-0.46239	-0.24482	-0.33	20
Hunan	-0.30088	0.07489	-0.15	13
Guangdong	2.33466	-0.70342	1.11	3
Guangxi	0.2076	-1.49247	-0.3	17
Hainan	0.44053	-2.15344	-0.36	22
Chongqing	-0.15812	0.67317	0.1	11
Sichuan	-0.14358	-0.77744	-0.3	17
Guizhou	-0.38312	-1.23214	-0.56	29
Yunnan	-0.61836	-0.93256	-0.61	30
Xizang	-0.41572	-1.55521	-0.67	31
Shanxi	-0.39858	0.27103	-0.15	13
Gansu	-0.7757	-0.24716	-0.5	27
Qinghai	-0.68381	-0.57248	-0.54	28
Ningxia	-0.53975	0.46989	-0.17	15
Xinjiang	-0.81097	0.18182	-0.4	24

III. MULTI-OBJECTIVE DECISION-MAKING

There are many common features about Multi-objective decision-making .Understand and grasp of these characteristics. The study of multi-objective decision theory and method, and Decision-making practice more flexible and effective decision making is important.

A. The Method of Linear Weighting

First, determine the decision matrix A, and then there are normalized, get the matrix B, then the weights obtained by factor analysis; we can weighted composite score of various cities. Linear weighting method steps:

- Determine the decision matrix G.
- The standardization decision matrix A, which transform into all the target values are dimensionless quantity, and they are in the range of (0, 1). This transformation is nonlinear. Transforming the property of the maximum and minimum value that are not uniform minimum which is not 0, the maximum value is not 1. It is sometimes not easy to compare between the objectives. The decision matrix in the actual index value is not often comparable. The matrix element of standardization and the index value will change to (0,1); and weight formula as follows:

$$b_{ij} = a_{ij} \Big/ \sqrt{\sum_{i=1}^{m} a_{ij}^2} \quad . \tag{1}$$

$$B = \begin{bmatrix}
0.23718795 & 0.243385521 & 0.208925412 & 0.292441823 & 0.35508002 & 0.288942722 & 0.318625201 & 0.295259031 \\
0.213456531 & 0.178645412 & 0.248224803 & 0.217931923 & 0.277349049 & 0.197548709 & 0.215084576 & 0.218220171 \\
0.134571154 & 0.176099661 & 0.178242456 & 0.153302805 & 0.183748402 & 0.133849088 & 0.126510588 & 0.12757216 \\
0.126867239 & 0.176175538 & 0.203167587 & 0.125783293 & 0.174868561 & 0.117345851 & 0.139280888 & 0.095813283 \\
0.151548426 & 0.250325943 & 0.166999673 & 0.179384871 & 0.197566787 & 0.150151987 & 0.184948112 & 0.215000261 \\
0.186718436 & 0.183871633 & 0.206429001 & 0.135317381 & 0.207430247 & 0.163255794 & 0.153123289 & 0.223586687 \\
0.141042545 & 0.195053424 & 0.208756494 & 0.136141447 & 0.207734646 & 0.120323187 & 0.143276537 & 0.178310897 \\
0.133406867 & 0.188459869 & 0.152878789 & 0.131874442 & 0.196471856 & 0.094378909 & 0.121138052 & 0.134980468 \\
0.303167648 & 0.235467989 & 0.267375866 & 0.315288965 & 0.171574684 & 0.425015677 & 0.384263981 & 0.392933824 \\
0.19381931 & 0.180697188 & 0.16925895 & 0.216937177 & 0.180511315 & 0.171100439 & 0.240587746 & 0.192062595 \\
0.235525535 & 0.239470887 & 0.216619296 & 0.190231063 & 0.211968983 & 0.301466938 & 0.293501678 & 0.21841303 \\
0.166542145 & 0.156493975 & 0.160491463 & 0.154569572 & 0.144006063 & 0.11601531 & 0.155085689 & 0.136603001 \\
0.216602239 & 0.171158366 & 0.211163574 & 0.192593916 & 0.122811663 & 0.223906261 & 0.194258814 & 0.219620496 \\
0.154941919 & 0.150140439 & 0.138244655 & 0.166191826 & 0.109938281 & 0.1099422 & 0.126458454 & 0.141671844 \\
0.15777246 & 0.215879336 & 0.202545514 & 0.215043694 & 0.181683481 & 0.177714081 & 0.170764831 & 0.155968579 \\
0.131347827 & 0.176802685 & 0.156507275 & 0.168898707 & 0.179657178 & 0.115303421 & 0.132201279 & 0.135852527 \\
0.170432486 & 0.170206032 & 0.148495046 & 0.161186096 & 0.15340838 & 0.112153467 & 0.13865661 & 0.109317619 \\
0.169330038 & 0.16889909 & 0.156057369 & 0.179971584 & 0.179675351 & 0.122350497 & 0.148397757 & 0.157901364 \\
0.250211336 & 0.150989022 & 0.283937938 & 0.252697342 & 0.189997683 & 0.33050321 & 0.25885223 & 0.249228574 \\
0.174130911 & 0.119588335 & 0.144770732 & 0.161036751 & 0.120251525 & 0.173377225 & 0.144578564 & 0.12160191 \\
0.180268369 & 0.076161919 & 0.179700997 & 0.150814608 & 0.12185076 & 0.16423858 & 0.124437236 & 0.103590539 \\
0.188432881 & 0.200423657 & 0.178146628 & 0.224575115 & 0.199506767 & 0.131587421 & 0.169374576 & 0.128125582 \\
0.181487246 & 0.161424431 & 0.133068297 & 0.157481803 & 0.128331748 & 0.141300618 & 0.126594806 & 0.141722155 \\
0.153444552 & 0.131855632 & 0.135871684 & 0.140197768 & 0.107082829 & 0.109763283 & 0.124953235 & 0.109120567 \\
0.182204157 & 0.158954558 & 0.120061621 & 0.088524343 & 0.137856734 & 0.153271702 & 0.097979602 & 0.063064954 \\
0.181798149 & 0.156681345 & 0.103127605 & 0.082731884 & 0.072029155 & 0.121807445 & 0.056090131 & 0.16786296 \\
0.15294088 & 0.162223463 & 0.163668418 & 0.16485572 & 0.195974367 & 0.121905724 & 0.171319597 & 0.167988738 \\
0.135781928 & 0.158353736 & 0.137450416 & 0.145672868 & 0.148751519 & 0.102961903 & 0.12516712 & 0.126364694 \\
0.141417846 & 0.14635588 & 0.130380229 & 0.143622039 & 0.138574572 & 0.099239918 & 0.117751978 & 0.130691448 \\
0.142991127 & 0.182550754 & 0.173652438 & 0.159161936 & 0.185563442 & 0.138134285 & 0.139522846 & 0.169259093 \\
0.137998765 & 0.1927926 & 0.126996999 & 0.142760637 & 0.146175479 & 0.126488276 & 0.108595006 & 0.172579625
\end{bmatrix}$$

TABLE V. THE MAIN COMPONENT IN ALL CITIES AND THEIR SCORE

Area	U	rank
Beijing	0.243176026	2
Tianjin	0.212386894	5
Hebei	0.145014241	25
Shanxi	0.139675498	30
Inner Mongolia	0.169059698	12
Liaoning	0.185754263	9
Jilin	0.154331563	21
Heilongjiang	0.143008975	27
Shanghai	0.292095618	1
Jiangsu	0.190775157	7
Zhejiang	0.23434993	4
Anhui	0.162462462	17
Fujian	0.206653238	6
Jiangxi	0.151202908	23
Shandong	0.17194429	10
Henan	0.141590851	28
Hubei	0.166415591	15
Hunan	0.16784055	14
Guangdong	0.238860114	3
Guangxi	0.162220388	18
Hainan	0.161996992	19
Chongqing	0.189345814	8
Sichuan	0.17181026	11
Guizhou	0.146268262	24
Yunnan	0.168489097	13
Xizang	0.164468503	16
Shanxi	0.156054982	20
Gansu	0.13890443	31
Qinghai	0.140268303	29
Ningxia	0.152178002	22
Xinjiang	0.144978064	26

- Make the linear weighting function of program structure, so that

$$U\left(A_i\right)=\sum_{i=1}^{n}\omega_j b_{ij} \quad i=1,2,\ldots,m. \qquad (2)$$

Based on the above two formulas, we can calculate the score of each area. According to the score, get all cities' composite score and ranking, as the table 3. Similarly, based on the data the top of the ranking is consistent with the previous method, and also developed eastern cities are at the forefront.

IV. CONCLUSION

In summary, principal component analysis and the method of linear weighting are consistent with the conclusions from Table 2 and Table 3. The first seven of the city is the same, but the city after the ranking is only a small difference. The core idea of principal component analysis is the component of rotation matrix, we can find by observing the rotation after the public factor of distribution of each more clearly. And we are more able to represent variables. Such as food is the first principal component, which reflects the number of representatives of urban residents' consumption variables [4]. Therefore, we can use the new extraction of two potential factors (food, clothing) on the sample of 31 cities can comprehensive description of urban household consumption. Then, based on the formula, we can determine the principal component of each city's urban consumer composite score and obtain a comprehensive ranking.

There are many ways for Multi-objective decision-making, for example, a simple linear weighting method, TOPSIS method, AHP method, principal component analysis, Grey relational analysis method[4]. A simple linear weighting method is not simple, in the process of its calculation in the calculation of complex thought has joined them, first of all is the standardization of the original matrix, that is, the original data processing dimensionless, so the data with comparable; the linear transformation sucked into a unity of its dimension[5]. This is very helpful in data operations. In the linear weighting method used in the weights is the weight factor analysis was, therefore, this is a link between the two. By comparison of the two is not difficult to find multi-objective decision making methods and concluded that a uniform consistency.

ACKNOWLEDGMENT

This paper is supported by Soft Science Research Project in Shaanxi Province Department of Education(12JK0052) and Philosophy and Social Science Base Project in Shaanxi Province Department of Education(13JZ026).

REFERENCES

[1] China Statistical Yearbook 2009. Beijing: China Statistics Press, 2009.(in Chinese)

[2] You JiaXing. How to apply a comprehensive evaluation of factor analysis, Statistical Education, 2003 (5):10-11. (in Chinese)

[3] Mikael Skou,Andersen,Ilmo Massa. Ecological Modernization-origins, Dilemmas and Future Directions, Journal of Environmental Policy and Planning. 2002, 2(4):337~345.

[4] Harald Payer, Petra Burger, Sylvia Lorek. Food Consumption in AustriaDriving Forces and Environmental Impacts, OECD Case StudyAustria. 2000,(11): 23~31.

[5] Xia Jianfeng, Zhou Li. Development trend forecast of consumption structure in Wuhan urban residents, Journal of Wuhan Institute of Technology.2010(2):17~21