Asian Agricultural Research 2014, 6(2): 1 –4, 8

Comprehensive Evaluation of the Economic Development Level of Guanzhong – Tianshui Economic Zone Using Principal Component Cluster Analysis

Chunxiang ZHAO *, **Xinrong YAN**, **Xu GUO**
College of Science, Engineering University of CAPF, Xi'an 710086, China

Abstract Based on 10 years of statistics concerning economic development in Xi'an as the main part of Guanzhong – Tianshui Economic Zone, this article builds the main indicator system to reflect economic development. Using two mathematical methods (principal component analysis and cluster analysis), we carry out comprehensive evaluation analysis of the main economic indicators, point out the distribution differences in the economic development level in this region, and make classification, in order to provide a scientific basis for the decision – making body to lay down the relevant economic development strategies in accordance with the economic development level and geographical location.
Key words Regional economy, Principal component analysis, Cluster analysis

In June 2009, the National Development and Reform Commission issued Guanzhong – Tianshui Economic Zone Development Plan, bringing a new development opportunity for Shaanxi and Gansu. Guanzhong – Tianshui Economic Zone includes Xi'an City, Tongchuan City, Baoji City, Xianyang City, Weinan City, Shangluo City (some districts and counties) and Yangling Demonstration Zone in Shaanxi Province, and the administrative regions under the jurisdiction of Tianshui City, Gansu Province, covering 80 000 square kilometers. At the end of 2010, the permanent population was 28.107 million. The directly involved regions include Hanzhong and Ankang in southern Shaanxi, Yan'an and Yulin in northern Shaanxi, Pingliang, and Qingyang in Gansu.

Guanzhong – Tianshui Economic Zone, located in the center of the Eurasian Continental Bridge, at the strategic point connecting east and west, north and south, is the region with a good economic foundation, superior natural conditions, long human history and great development potential in China's western regions. Du Yueping[1] made a comprehensive and systematic analysis of development of Guanzhong – Tianshui Economic Zone, and he proposed that it was necessary to seize the opportunity, work hard, and make every effort to build Xi'an into an international metropolis. Yan Xinrong[2] discussed on how to change the mode of economic development. In this paper, using the principal component cluster analysis method, we carry out statistical analysis of the economic development level.

Regional economy is a comprehensive geographical concept of economic development, and also a comprehensive economic concept arising from the interaction between internal factors of economic development and external conditions in a certain area. It is restricted and influenced by the natural conditions, development and utilization of resources, socio-economic conditions, economic policies, and many other factors in the region. Regional economic development not only means the expansion of the size of the region's national economy, but also means the improvement of economic and social living standards. The regional economic effects are not merely reflected in the economic indicators, and there is a need to consider the overall social economic benefits and regional ecological benefits. Therefore, the study of economic development is of very important significance to the future development.

1 Data collection[3-4]

Xi'an City administers 13 districts and counties, and the differences in the resource endowments between regions lead to great disparities the economic development. In order to scientifically, objectively, accurately and reasonably measure the level of economic development of the district and county, according to the status of economic development in Xi'an City, this article selects seven key indicators that reflect the economic situation, and establishes the corresponding evaluation system.

The specific indicators are as follows: per capita GNP (X_1), the indicator reflecting the level of economic growth; per capita total industrial output value (X_2); per capita total agricultural output value (X_3); per capita total tertiary industry output value (X_4); per capita revenue (X_5), the indicator reflecting the government's financial strength; total retail sales of social consumer goods (X_6), the indicator reflecting the level of social purchasing power; fixed asset investment (X_7), the indicator reflecting the level of investment.

According to the total amount of economic indicators and population in the districts and counties of Xi'an City, we calculate the above indicators, as shown in Table 1.

Received: October 21, 2013 Accepted: November 28, 2013
Supported by Shaanxi Natural Science Fundamental Research Foundation (2011JM1019).
* Corresponding author. E-mail: yxr663311@163.com

Table 1 The major economic indicators for 13 districts and counties of Xi'an City in 2011

	Per capita GNP (X_1) yuan	Per capita total industrial output value (X_2) // yuan	Per capita total agricultural output value (X_3) // yuan	Per capita total tertiary industry output value (X_4) yuan	Per capita revenue (X_5) yuan	Total retail sales of social consumer goods (X_6) // 10^4 yuan	Fixed asset investment (X_7) 10^4 yuan
Xincheng District	64 138.11	15 254.09	0	38 308.29	3 472.126	3 419 069	2 884 809
Beilin District	70 585.1	7 026.022	0	52 738	4 325.036	3 420 722	291 1287
Lianhu District	62 338.51	19 623.56	0	33 894.2	4 011.065	2 816 163	3 764 003
Baqiao District	34 586.6	17 185.57	2 406.882	12 047.77	2 196.426	464 033	1 865 078
Weiyang District	58 115.6	21 541.78	284.6931	27 395.86	2 151.565	2 782 585	4 498 254
Yanta District	61 722.65	11 643.32	220.2903	39 544.23	2 232.326	3 800 048	7 677 355
Yanliang District	42 759.73	17 429.49	6 287.04	15 126.74	2321.421	232 089	1 209 231
Lintong District	28 000.91	14 145.2	4 284.632	8 252.501	948.833	485 405	1 178 946
Chang'an District	29 806.44	11 785.16	2 657.554	12 465.83	1 652.839	1 100 538	3 143 523
Lantian County	15 864.47	3 552.759	3 967.086	6 193.611	410.919 7	354 500	770 932
Zhouzhi County	11 710.55	2 371.444	3 587.206	5 025.623	274.076 7	23 7439	684 630
Huxian County	23 120.86	11 067.14	3 803.044	6 633.841	801.325	388 564	1 098 101
Gaoling County	56 158.66	41 216.82	5 049.21	6 078.139	2 088.607	158 619	1 776 472

2 Evaluation methods

Due to differences in resource endowments between the districts and counties, there are great differences in the economic development, and also in the rate of contribution to overall economic development of Xi'an. In order to objectively and comprehensively understand the specific circumstance of economic development in the districts and counties of Xi'an City to identify the degree of difference between them, this article will use both principal component analysis and cluster analysis.

Using the principal component analysis, the principal component scores are calculated, and the economic development of cities is reflected by the scores; using the cluster analysis to divide the 13 cities into several categories, the appropriate countermeasures are proposed according to their respective economic development level and geographical location.

Principal component analysis (PCA)[5-6], proposed by Hotelling in 1933, is a statistical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (i. e., uncorrelated with) the preceding components. Principal components are guaranteed to be independent if the data set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables.

Hierarchical clustering[5-6] is a method of cluster analysis which seeks to build a hierarchy of clusters. Strategies for hierarchical clustering generally fall into two types: Agglomerative: This is a "bottom up" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. Divisive: This is a "top down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy. In general, the merges and splits are determined in a greedy manner. The results of hierarchical clustering are usually presented in a dendrogram. Hierarchical clustering has the distinct advantage that any valid measure of distance can be used. In fact, the observations themselves are not required: all that is used is a matrix of distances.

2.1 The evaluation process using principal component analysis Assuming there are P observation objects, each with n indicators, the raw data of n indicators are derived as follows:

$X=(x_{ij})_{p\times n}$.

(i) Standardization of raw data. In order to make the comprehensive evaluation results objective and reasonable, it is necessary to eliminate the impact of dimension and magnitude. The standardization processing is commonly used:

$y_{ij}=(x_{ij}-\bar{x}_j)/s_j,\ i=1,2,\cdots,p,\ j=1,2,\cdots,n$

where $\bar{x}_j$, s_j are the sample mean and sample standard deviation of indicator j, respectively.

(ii) Calculation of correlation coefficient matrix of indicator sample.

$R=(r_{ij})_{n\times n}$

where r_{ij} is the correlation coefficient between indicator x_i and indicator x_j.

$$x_{ij}=\frac{\sum_{k=1}^{p}(x_{ki}-\bar{x}_i)(x_{ki}-\bar{x}_j)}{\sqrt{i\sum_{k=1}^{p}(x_{ki}-\bar{x}_i)^2(x_{ki}-\bar{x}_j)^2}}\ i,j=1,2,\cdots,n.$$

(iii) Calculation of eigenvalues and eigenvectors of the correlation coefficient matrix. The eigenvalues and eigenvectors of the correlation coefficient matrix are calculated. n eigenvalues of R are denoted as $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\geqslant 0$, and the corresponding eigenvectors are as follows:

$u_i=(u_{i1},u_{i2},\cdots,u_{in})\quad i=1,2,\cdots,n$

(iv) Calculation of the contribution rate of eigenvalues and extraction of principal component in accordance with the criterion of cumulative contribution rate. When the cumulative contribution rate of variance reaches 85%, the first m principal components are taken:

$$Y_i = u_{i1}y_1 + u_{i2}y_2 + \cdots + u_{in}y_n \quad i = 1, 2, \cdots, m$$

(v) Calculation of comprehensive evaluation values based on multi – indicator and weighted comprehensive evaluation model. The weighted sum of m principal components is calculated to get the final evaluation value, and the weight is the variance contribution rate of each principal component.

$$Y = \sum_{i=1}^{m} \frac{\lambda_i}{\sum_{i=1}^{m}\lambda_i} Y_i$$

2.2 Principal component analysis According to the data in Table 1, using SPSS software[7], we can derive the correlation coefficient matrix of X_1, X_2, X_3, X_4, X_5, X_6 and X_7 as follows:

$$R = \begin{pmatrix} 1 & & & & & & \\ 0.452 & 1 & & & & & \\ -0.640 & 0.151 & 1 & & & & \\ 0.846 & -0.089 & -0.831 & 1 & & & \\ 0.899 & 0.275 & -0.637 & 0.847 & 1 & & \\ 0.802 & -0.075 & -0.917 & 0.947 & 0.725 & 1 & \\ 0.655 & 0.097 & -0.713 & 0.679 & 0.452 & 0.810 & 1 \end{pmatrix}$$

The eigenvalues, contribution rate and cumulative contribution rate of the correlation matrix R is shown in Table 2.

From the above table, we can find that when the number of principal component is 2, the cumulative contribution rate reaches 88.01% (> 85%), containing the vast majority of the needed information. Two principal components will be enough, so here we take the first two as the first principal component and second principal component, respectively.

The original value of each indicator is standardized:

$$y_{ij} = (x_{ij} - \bar{x}_j)/s_j$$

where $\bar{x}_J$ and s_j are the sample mean and sample standard deviation of indicator j. The standardized matrix is as follows:

Table 2 Eigenvalues and contribution rate

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	% of Variance	Cumulative//%	Total	% of Variance	Cumulative//%
X_1	4.831	69.020	69.020	4.831	69.020	69.020
X_2	1.329	18.990	88.010	1.329	18.990	88.010
X_3	0.574	8.202	96.212			
X_4	0.186	2.663	98.876			
X_5	0.069	0.988	99.864			
X_6	0.009	0.132	99.995			
X_7	0.000	0.005	100.000			

Table 3 The standardized matrix

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
Xincheng District	1.04	0.03	−1.14	1.12	1.09	1.30	0.16
Beilin District	1.36	−0.80	−1.14	2.02	1.75	1.30	0.17
Lianhu District	0.96	0.48	−1.14	0.85	1.51	0.89	0.61
Baqiao District	−0.42	0.23	−0.04	−0.51	0.10	−0.71	−0.36
Weiyang District	0.75	0.68	−1.01	0.44	0.06	0.87	0.99
Yanta District	0.93	−0.33	−1.04	1.20	0.13	1.56	2.62
Yanliang District	−0.01	0.26	1.72	−0.32	0.20	−0.87	−0.70
Lintong District	−0.74	−0.08	0.81	−0.75	−0.87	−0.70	−0.72
Chang'an District	−0.65	−0.32	0.07	−0.49	−0.32	−0.28	0.29
Lantian County	−1.34	−1.16	0.67	−0.88	−1.29	−0.79	−0.92
Zhouzhi County	−1.55	−1.28	0.49	−0.95	−1.39	−0.87	−0.97
Huxian County	−0.98	−0.39	0.59	−0.85	−0.98	−0.77	−0.76
Gaoling County	0.65	2.68	1.16	−0.88	0.02	−0.92	−0.41

Table 4 Factor loading matrix

	Per capita GNP//yuan	Per capita total industrial output value//yuan	Per capita total agricultural output value//yuan	Per capita total tertiary industry output value//yuan	Per capita revenue//yuan	Total retail sales of social consumer goods//10^4 yuan	Fixed asset investment 10^4 yuan
The first principal component	0.911	0.113	−0.877	0.958	0.856	0.966	0.797
The second principal component	0.385	0.965	0.317	−0.151	0.274	−0.210	−0.087

Table 5 Principal component eigenvectors

	Per capita GNP//yuan	Per capita total industrial output value//yuan	Per capita total agricultural output value//yuan	Per capita total tertiary industry output value//yuan	Per capita revenue//yuan	Total retail sales of social consumer goods//10^4 yuan	Fixed asset investment 10^4 yuan
The first principal component	0.414	0.051	−0.399	0.436	0.389	0.439	0.363
The second principal component	0.334	0.837	0.275	−0.131	0.238	−0.182	−0.075

From the factor loading matrix, we know that the standardized eigenvector corresponding to the eigenvalue of R is (j = 1, 2,) in the column j of this matrix.

$$l_j = a_{ij}/\sqrt{\lambda_j}$$

Assuming Y_1 and Y_2 represent the first principal component and second principal component, then:

$$Y_1 = 0.414X_1^* + 0.051X_2^* - 0.399X_3^* + 0.436X_4^* + 0.389X_5^* + 0.439X_6^* + 0.363X_7^*$$

$$Y_2 = 0.334X_1^* + 0.837X_2^* + 0.275X_3^* - 0.131X_4^* + 0.238X_5^* - 0.182X_6^* - 0.075X_7^*$$

X_j^* (i = 1, 2, , 7) in the formula is the standardized indicator of X_{ij}^*, namely:

$$x_i^* = \frac{x_{ij} - \bar{x}_i}{S_i}$$

where $\bar{x}_i$ is the mean of x_{ij}; S_i is the standard deviation of x_{ij}.

The first principal component Y_1 is mainly determined by X_1, X_3, X_4, X_5, X_6, X_7, and their coefficients are much larger than the coefficients of the other variables. Y_1, is the integrated expression of X_1, X_3, X_4, X_5, X_6, X_7. F_2 is mainly determined by X_2, so the second principal component Y_2 mainly reflects the industrial sector.

Thus, each principal component has a particular substantive significance. From two perspectives, two principal components reflect the economic development level of the districts and counties, so it is appropriate to use them to analyze and assess the economic development level of the districts and counties. After the score of principal components is calculated, then it is multiplied by the variance contribution rate of principal component, thereby we get the composite score for each city:

$$Y = 0.7843Y_1 + 0.2157Y_2$$

Table 6 Principal component scores and ranking

Sample	Xincheng District	Beilin District	Lianhu District	Baqiao District	Weiyang District	Yanta District	Yanliang District
Y_1	2.428 1	3.171 0	2.446 9	−0.772 0	1.704 7	2.992 8	−1.374 9
Y_2	−0.076 9	−0.626 3	0.449 2	0.288 1	0.266 0	−0.858 3	0.987 6
Y	1.887 6	2.351 7	2.015 9	−0.543 3	1.394 3	2.162 0	−0.865 2
Rank	4	1	3	7	5	2	9

Sample	Lintong District	Chang'an District	Lantian County	Zhouzhi County	Huxian County	Gaoling County
Y_1	−1.8677	−0.6691	−2.4475	−2.5904	−2.0268	−0.9857
Y_2	−0.0188	−0.4485	−1.2132	−1.4128	−0.4163	3.0975
Y	−1.4688	−0.6215	−2.1812	−2.3364	−1.6793	−0.1047
Rank	10	8	12	13	11	6

Based on the weighted composite value Y, we derive the ranking of economic development indicators of 13 districts and counties in the whole city. No. 1 is Beilin District, No. 2 is Yanta District, and No. 3 is Lianhu District. The development scale of the tertiary industry in these areas is large, occupying a large proportion of the entire national economy, thus ranking forward. Zhouzhi County, Lantian County and Huxian County have no advantages in this area, so they rank behind.

2.3 Cluster analysis According to the component scores of principal component, we carry out cluster analysis of the regions, and divide the 13 districts and counties of Xi'an City into three categories.

The first category: Yanta District, Beilin District, Weiyang District, Xincheng District and Lianhu District; The second category: Gaoling County, Yanliang District, Baqiao District, Chang'an District; The third category: Lantian County, Zhouzhi County, Lintong District, Huxian County (Fig. 1).

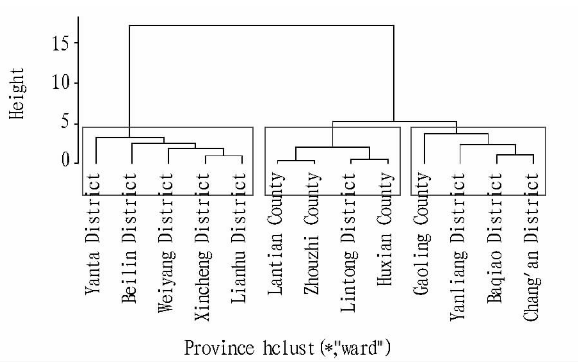

Fig. 1 Cluster analysis of the regions

The first category is the inner city of Xi'an; the second category is the north and south of Xi'an City; the third category is the east and west of Xi'an City. The inner city is developed rapidly, with high economic level, and especially the tertiary industry is booming, followed by urban north and south; east and west of city are lagging behind. For the districts in the first category, the per capita total agricultural output value is 100.82 yuan; for the districts in the third category, the per capita total agricultural output value is 3910.25 yuan, indicating that in the economically backward areas, there are more people engaged in agriculture, and the development of industry and tertiary industry is relatively backward. Thus, it can be seen that there are great differences in the economic development level between the regions of Xi'an City. How to bridge this gap? The development of secondary and tertiary industries is backward in the economically backward areas, so there is a need to vigorously develop industry and tertiary industry, and accelerate urbanization. The more developed regions can give technical help to backward areas, and the backward areas can also provide the labor resources needed by the developed regions.

3 Conclusions and recommendations

In this paper, using principal component analysis and cluster analysis, we carry out comprehensive evaluation analysis of the economic development level of 13 districts and counties of Xi'an City. The results of these two analyses are basically the same, and the results show that there are great differences in the economic development level between the regions of Xi'an. Yanta District, Beilin District, Weiyang District, Xincheng District and Lianhu District are the center of Xi'an City. They should give full play to their advantages in science, education, commerce and industry, focus on developing high-tech industries and equipment manufacturing to constantly improve the city's service function, and strive to keep the economic development ahead. The development of secondary

(To page 8)

D_2 ("complete foreclosure").

5 Conclusions

Based on repeated game, this article considers the "good" or "bad" type of enterprises, and studies the impact of vertical integration between one upstream enterprise and one downstream enterprise on the independent downstream enterprises in a duopoly market. Meanwhile the suppliers can invest in reverse engineering, which will also affect the downstream enterprises' innovation and profits.

The study results show that when the cost is higher than the threshold value $\beta(\Pi^I - \Pi^S)\theta(\Delta - 2\delta)$, there will be no enterprise investing in reverse engineering; when the cost is lower than the threshold value, the integrated enterprises will invest in reverse engineering; in the game of Time period 2, the downstream enterprises' innovation investment and profits increase with the increase in the "good" probability of suppliers before vertical integration; after vertical integration, the independent downstream enterprises' innovation investment and profits decline, but the integrated enterprises' innovation investment and profits increase; in the game of Time period 1, the vertical integration makes the independent downstream enterprises "totally foreclosed".

References

[1] XIA Y, NIU RF. Agricultural industrial integration and international experience[J]. Issues in Agricultural Economy, 1996, 12:2 -7. (in Chinese).
[2] NIU RF. Study on agricultural industrial integrated operation[J]. Issues in Agricultural Economy, 1997 (2):18 -24. (in Chinese).
[3] Spengler JJ. Vertical integration and antitrust policy[J]. The Journal of Political Economy, 1950, 58(4):347 -352.
[4] Stigler GJ. The division of labor is limited by the extent of the market[J]. The Journal of Political Economy, 1951, 59(3):185 -193.
[5] Arrow KJ. Vertical integration and communication[J]. The Bell Journal of Economics, 1975:173 -183.
[6] Williamson OE. The vertical integration of production:market failure considerations[J]. The American Economic Review, 1971, 61(2):112 -123.
[7] Riordan MH, Williamson OE. Asset specificity and economic organization [J]. International Journal of Industrial Organization, 1985, 3(4):365 -378.
[8] ZHANG DS, CHEN HM. Analysis on effects of vertical integration[J]. Journal of Systems Engineering, 2001, 16(6):425 -449. (in Chinese).
[9] HUO PJ, XUAN GL. Vertical integration with research and development (R&D) investment[J]. Journal of Shanghai Jiaotong University, 2000, 34 (11):1562 -1565. (in Chinese).
[10] HUO PJ, XUAN GL. Comparisons of the upstream R&D expenditures between vertical integration and non integration[J]. Journal of Systems Engineering, 2001, 16(1):36 -38. (in Chinese).
[11] HUO PJ, XUAN GL. The effect of vertical integration on R&D expenditure of the downstream firm[J]. Journal of Industrial Engineering and Engineering Management, 2002, 16(1):44 -46. (in Chinese).
[12] JING H, XUAN ZJ. Endogenous vertical integration of successive oligopolists[J]. Journal of Industrial Engineering and Engineering Management, 2007, 21(2):10 -14. (in Chinese).
[13] Hart O, Tirole J, Carlton D W, *et al*. Vertical integration and market foreclosure[J]. Brookings Papers on Economic Activity. Microeconomics, 1990:205 -286.

(From page 4)

developing high - tech industries and equipment manufacturing to constantly improve the city's service function, and strive to keep the economic development ahead. The development of secondary and tertiary industries is balanced in Gaoling County, Yanliang District, Baqiao District and Chang'an District, so the overall economic strength is also great. They should give full play to their strengths to make the central city facilities serve Chang'an and Gaoling, speed up the construction of Chang'an Aerospace Industrial Base, Jinghe Industrial Park, Yanliang Aviation Industrial Base, and take active support policies and measures, to promote a new round of rapid development of the economy. The secondary and tertiary industries are underdeveloped in Lantian County, Zhouzhi County, Lintong District, and Huxian County. Based on their own characteristics, they should rely on the industrial park, undertake the relocation of the central city businesses, and form "supporting industry + featured agriculture" functional area by providing support services for large enterprises. In summary, it is reasonable to use principal component analysis and cluster analysis to carry out comprehensive evaluation of the overall economic strength of districts and counties in Xi'an City. In fact, it is also consistent with the actual situation of Xi'an. Through the division of the economic conditions of 13 districts and counties, it provides data support for further economic development and policy establishment in Xi'an City.

References

[1] DU YP, ZHAO W, DENG JR. The development of Western Chinese and Guanzhong - Tianshui Economic Region[M]. Xidian University Press, 2012. (in Chinese).
[2] YAN XR. Study on economic development pattern of Guanzhong - Tianshui Economic Region[J]. Journal of Guizhou College of Finance and Economics, 2013: 58 -59(in Chinese).
[3] DING YX, ZHANG XG. Shaanxi statistical yearbook, 2012[M]. Shaanxi Statistics Press, 2012. (in Chinese).
[4] ZHANG XG, CHEN B. Statistical yearbook of Guanzhong - Tianshui Economic Region[M]. Shaanxi Provincial Bureau of Statistics, 2012. (in Chinese).
[5] WU Y, LI YL, HU QJ. Applied mathematical statistics[M]. National University of Defense Technology Publish House, 2008. (in Chinese).
[6] XUE Y, CHEN LP. Statistical modeling and R software[M]. Beijing: Tsinghua University Press, 2012. (in Chinese).
[7] YU JY, HE XH. Data statistics analysis and SPSS application[M]. Beijing: Posts & Telecom Press, 2003. (in Chinese).