Frontiers in Educational Research

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

The Application Of Fuzzy Cluster Analysis In The Evaluation And Classification Of Urban Happiness

Jing Peng , Hean Liu

College of Science, Hunan City University, Yiyang, China

ABSTRACT. *Urban happiness is a kind of value embodiment of residents' life satisfaction, and the evaluation and classification of urban happiness is an important basis for the formulation of urban development direction and strategy. But city happiness too much consideration indicators, here we are in the first three lines of ten cities as evaluation unit, and in such aspects as economy, housing, employment, education has chosen 13 indexes to establish the original matrix, using the translation effect of the standardization of original data matrix, then using the hamming distance method to calculate the distance of each city or similar degree, and finally select a threshold lambda, according to the actual situation and needs to adjust lambda value, get the selected area under different level of lambda different classification, finally concluded that happiness will cities for classification and discrimination.*

KEYWORDS: *Urban Development; Fuzzy Clustering; Urban Happiness*

1. Introduction

Happiness is a kind of physical and mental experience. It is not only an objective judgment of living conditions, but also an effective judgment of subjective satisfaction with life. As the most important non-economic factor, it is the barometer of social operation and living conditions of residents. It is also the “vane” of urban development and popular aspiration [1-5]. Therefore, the study of urban well-being is conducive to locating the city and formulating the urban development strategy scientifically, which is of great significance to the direction and shape of urban development. However, urban happiness is affected by many factors, which are vague and uncertain. The result of classification and discrimination by traditional clustering will have some limitations, so we use fuzzy clustering to discriminate, and classify the characteristics of an objective thing, the degree of familiarity and alienation and the degree of similarity, which can more objectively reflect the series of urban comprehensive well-being, and help people to have a clearer and more numerical understanding of the well-being of each city. It will help people to have a deeper study of urban well-being, and also provide a certain reference for the future direction of urban development and the formulation of its strategy [6-9].

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

2. Fuzzy Sets and Fuzzy Matrix

2.1 Fuzzy Sets

Given a universe U, a mapping from U to unit interval [0,1] is called a fuzzy set on U, or a fuzzy subset of U. Represents that a fuzzy set can be recorded as A. The mapping (function) u (?) or abbreviated as A (?) is called the attribution function of fuzzy set A. For each x < U, (x) is called the degree of attribution of element X to fuzzy set A.

2.2 Fuzzy Matrix

A matrix in which all elements in a matrix are valued in the [0,1] closed interval is called a fuzzy matrix. Let R=(), 0 < 1, call R a fuzzy matrix. When only 0 or 1 is taken, R is called Boole matrix; when only 0 is taken, R is called zero matrix; when only 1 is taken, R is called full matrix; when the elements on the diagonal line of fuzzy matrix R= (m*n) are all 1, R is called fuzzy unit matrix and I.

3. 3 General steps of fuzzy clustering analysis

3.1 Data Matrix

Let domain $U=\{x_1, x_2, \cdots, x_n\}$ be classified as each element, and the attributes of each element are represented by m indices.

$x_i=\{x_{i1}, x_{i2}, \cdots, x_{im}\}$ $(i=1,2,\cdots,n)$, then, the original data matrix is obtained as follows :

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}$$

Among them, x_{nm} is the initial data of n index of m classification element.

3.2 Data standardization

When dealing with practical problems, it is often necessary to transform data appropriately. However, even so, the data obtained are usually not in the region [10-13]. Therefore, data standardization is to meet the requirements of fuzzy matrix and

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

transform data into intervals. There are usually the following transformations:

Translational and Standard Deviation Transform

$$x'_{ik} = \frac{x_{ik} - \overline{x}_k}{s_k} \quad (i = 1, 2, \cdots, n; k = 1, 2, \cdots, m)$$

among $\overline{x}_k = \frac{1}{n}\sum_{i=1}^{n} x_{ik}$, $s_k = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ik} - \overline{x}_k)^2}$.

After conversion, the mean value of each variable is 0, the standard deviation is 1, and the dimension effect is removed. However, the re-use of x'_{ik} may not necessarily be in the area.

(2) translation and range transformation

$$x''_{ik} = \frac{x'_{ik} - \min_{1\le i\le n}\{x'_{ik}\}}{\max_{1\le i\le n}\{x'_{ik}\} - \min_{1\le i\le n}\{x'_{ik}\}}, (k = 1, 2, \cdots, m)$$

Obviously there is $0 \le x''_{ik} \le 1$, and it also eliminates the influence of dimension.

(3) Logarithmic transformation

$$x'_{ik} = \lg x_{ik} \ (i = 1, 2, \cdots, n; k = 1, 2, \cdots, m)$$

Take logarithms to reduce the order of magnitude between variables.

3.3 Establishment of Fuzzy Similarity Matrix

Let universe $U = \{x_1, x_2, \cdots, x_n\}$ and $x_i = \{x_{i1}, x_{i2}, \cdots, x_{im}\}$ determine similarity coefficient according to traditional clustering method, and construct fuzzy similarity matrix, similarity degree of x_i and x_j. In this paper, the method of determining $r_{ij} = R(x_i, x_j)$ is Hamming distance method.

Hamming distance

$$d(x_i, x_j) = \sum_{k=1}^{m} \left| x_{ik} - x_{jk} \right|.$$

3.4 Determination of the Best Threshold

In the fuzzy clustering analysis, different classifications can be obtained for

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

different classifications. Many practical problems need to choose a threshold. The methods are as follows.

Table 1 Original data table

Sample this	*Refers to the standard*					
	1	*2*	$\cdots$	*k*	$\cdots$	*m*
x_1	x_{11}	x_{12}	$\cdots$	x_{1k}	$\cdots$	x_{1m}
x_2	x_{21}	x_{22}	$\cdots$	x_{2k}	$\cdots$	x_{2m}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_i	x_{i1}	x_{i2}	$\cdots$	x_{ik}	$\cdots$	x_{im}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	$\cdots$	x_{nm}
$\overline{x}$	$(\overline{x_1}$	$\overline{x_2}$	$\cdots$	$\overline{x_k}$	$\cdots$	$\overline{x_m})$

Let the number of classifications corresponding to λ value be r, the number of samples in class j and the number of samples in class j be recorded as:

$x_1^{(j)}, x_2^{(j)}, \cdots, x_{n_j}^{(j)}$, the clustering center of the class is vector j, where $\overline{x}^{(j)} = (\overline{x}_1^{(j)}, \overline{x}_{12}^{(j)}, \cdots, \overline{x}_m^{(j)})$ is the average value of the first feature, that is, vector $\overline{x}_k^{(j)}$.

$$\overline{x}_k^{(j)} = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{ik}^{(j)},\ (k = 1, 2, \cdots, m),$$

Make Fstatistics

$$F = \frac{\sum_{j=1}^{r} n_j \left\| \overline{x}^{(j)} - \overline{x} \right\| \Big/ (r-1)}{\sum_{j=1}^{r}\sum_{i=1}^{n_j} \left\| x_i^{(j)} - \overline{x}^{(j)} \right\| \Big/ (n-r)},$$

Among $$\left\| \overline{x}^{(j)} - \overline{x} \right\| = \sqrt{\sum_{k=1}^{m} (\overline{x}_k^{(j)} - \overline{x}_k)^2}$$

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

4. Urban Happiness

4.1 Selection of Index of Urban Happiness

Because there are too many indicators to measure urban well-being, it is too cumbersome to collect and analyze them. Here we select some data and indicators that are closely related to life and residents are more concerned about. The selected indicators are as follows:

Table 2 index table of city happiness

Urban per capita economic income	*x1:Per capita monthly income level/yuan* *x2:Residents' average monthly consumption level/yuan* *x3:Urban GDP growth rate* *x4:Residents' annual disposable income/yuan*
Urban education	*x5:Per 100,000 people in higher education* *x6:Per capita education cost per yuan from primary school to university graduation*
Urban employment	*x7:Employment per 100,000 adults* *x8:Per capita sleep per hour* *x9:The number of people per 100,000 who are satisfied with their jobs*
Population mobility	*x10:Permanent floating population / 10,000 people*
The housing situation	*x11:Per capita living area/square meter* *x12:Average urban housing price per yuan* *x13:Average monthly rent per yuan*

4.2 Urban Selection

In this paper, the choice of cities is an important part, so we should select some representative cities, and include the first, second, third-tier cities and southern and Northern cities. The hierarchical selection is more representative and persuasive. It is also convenient for us to analyze the different levels of cities in the economic environment, living environment, life rhythm and other different circumstances for the city to live in. The impact of people's happiness is also easy to classify and distinguish.

Table 3 City selection table

New first-tier cities	*Changsha*	*Chengdu*	*Hangzhou*
Second-tier cities	*Xiamen*	*Wenzhou*	
Third line cities	*Sanya*	*Zhuzhou*	
First-tier cities	*Beijing*	*Shanghai*	*Shenzhen*

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

4.3 Calculation of Fuzzy Clustering of Urban Happiness

Table 4 Raw data

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
Beijing	10749	4255	11.86	62631	63073	1203960	96348	7.6	54354	794.3	31.96	55090	3400
Shanghai	11250	5204	13.52	69108	51028	1003481	94234	7.51	50278	972.69	27.07	52268	3780
Shenzhen	15583	3200	13.39	52938	47610	984230	91783	7.5	43298	818.11	40.02	50494	2980
Chengdu	7250	1836	8.2	36142	43184	764320	89230	7.46	23795	437	36.85	12669	2670
Hangzhou	7608	3451	11.7	54348	46870	832567	88482	7.47	23181	450.44	36.4	21458	2745
Changsha	6660	3567	6.36	44647	61863	667830	88678	7.38	19863	200	45.5	10106	2350
Xiamen	7452	3319	11.82	46630	15300	534628	84328	7.5	18630	221.02	43.5	47071	2890
Wenzhou	5050	2601	7.8	43185	48600	432000	81259	7.36	17962	297	44.01	19462	2210
Sanya	5467	1260	10.72	33638	38800	412080	88321	7.9	18920	20	32.47	31704	1980
Zhuzhou	5453	2172	-0.67	32867	19722	342590	85623	7.88	18230	-0.07	46.3	6077	1350

Here we use X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 to represent Beijing, Shanghai, Shenzhen, Chengdu, Hangzhou, Changsha, Xiamen, Wenzhou, Sanya, Zhuzhou, X1 x 3 x 4 x 5 x 6 x 7 x 8 x 9 x 10 x 11 x 12 x 13 respectively. Because of the complexity of the calculation, we use MATLAB to calculate the original data and get the fuzzy moments after standardization. Matrix and similarity matrix; the fuzzy similarity matrix obtained by MATLAB operation is as follows;

Table 5 Fuzzy similarity matrix table

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
X1	1.0000	0.8000	0.7000	0.4300	0.5900	0.4200	0.4300	0.3100	0.3100	0.5000
X2	0.8000	1.0000	0.7100	0.4000	0.5600	0.3400	0.4100	0.2800	0.2600	0
X3	0.7000	0.7100	1.0000	0.5700	0.7100	0.5100	0.6000	0.4800	0.3900	0.2000
X4	0.4300	0.4000	0.5700	1.0000	0.8300	0.7600	0.6600	0.7500	0.7000	0.5600
X5	0.5900	0.5600	0.7100	0.8300	1.0000	0.7500	0.7400	0.7100	0.6200	0.4200
X6	0.4200	0.3400	0.5100	0.7600	0.7500	1.0000	0.6800	0.8100	0.5800	0.5700
X7	0.4300	0.4100	0.6000	0.6600	0.7400	0.6800	1.0000	0.7100	0.6100	0.5600
X8	0.3100	0.2800	0.4800	0.7500	0.7100	0.8100	0.7100	1.0000	0.6600	0.6200
X9	0.3100	0.2600	0.3900	0.7000	0.6200	0.5800	0.6100	0.6600	1.0000	0.6800
X10	0.05000	0.2000	0.5600	0.4200	0.5700	0.5600	0.6200	0.6800	1.0000	

4.4 Fuzzy Dynamic Clustering of Urban Happiness

Similar Fuzzy Matrix is analyzed, and different values are taken to see the classification of cities.

When = 1.0, cities are classified as

{X1} {X2} {X3} {X4} {X5} {X3} {X6} {X7} {X8} {X9} {X10}

When = 0.83, X4 is similar to X5, where cities are classified as

{X1} {X2} {X3} {X4 X5} {X6} {X7} {X8} {X9} {X10}

When = 0.81, X6 is similar to X8, and cities are classified as

{X1} {X2} {X3} {X4 X5} {X6 X8} {X7} {X9} {X10}

When = 0.80, we can get the similarity between X1 and X2. At this time, we can

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

get the classification of cities.

{X1 X2} {X3} {X4 X5} {X6 X8} {X7} {X9} {X10}

When = 0.76, X4 is similar to X6, and cities are classified as

{X1 X2} {X3} {X4 X5 X6 X8} {X7} {X9} {X10}

When = 0.74, you can get the similarity between X5 and X7, where cities are classified as

{X1 X2} {X3} {X4 X5 X6 X7 X8} {X9} {X10}

When taking = 0.68, we can get the similarity between X9 and X10. At this time, cities are classified as

{X1 X2 X3 X4 X5 X6 X7 X8 X9 X10}

The domain U={X1 X2......... X10} has been synthesized into a class and clustering has been completed.

Use F statistics to determine values.

By formula

$$F = \frac{\sum_{j=1}^{r} n_j \left\| \overline{x}^{(j)} - \overline{x} \right\| \Big/ (r-1)}{\sum_{j=1}^{r} \sum_{i=1}^{n_j} \left\| x_i^{(j)} - \overline{x}^{(j)} \right\| \Big/ (n-r)}$$

The bigger the F value obtained, the bigger the direct gap between the proof class and the class, and the more detailed the classification, but the specific value should be determined according to the actual situation.

When cities are classified into nine categories, (Chengdu, Hangzhou) are classified into one category and the rest are classified into different categories.

When dividing cities into eight categories, (Chengdu, Hangzhou) is one category, (Changsha, Wenzhou) is one category, and the rest are different.

When cities are divided into seven categories, (Beijing, Shanghai) is one category, (Chengdu, Hangzhou) is one category, (Changsha, Wenzhou) is one category, and the rest are different.

When dividing cities into six categories, (Beijing, Shanghai) is one category (Chengdu, Hangzhou, Changsha, Wenzhou), and the rest are different.

When cities are divided into five categories, (Beijing, Shanghai) into one category, (Chengdu, Changsha). Xiamen and Wenzhou) belong to one group, the others are different.

When dividing cities into three categories, Wenzhou is one, Zhuzhou is one, and the rest is one.

The dynamic clustering diagram is as follows:

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

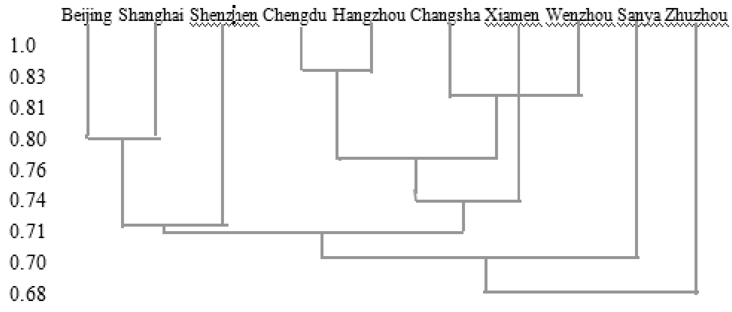

Fig.1 Dynamic clustering diagram

5. Conclusion

In the diversified economy and culture, each city has its own characteristics and culture. In order to create a city with a great sense of happiness, it is very important to find its own position. Considering many indicators, it is difficult to classify and distinguish it by using traditional clustering. Therefore, this paper uses fuzzy clustering to classify different cities and selects 10 cities. Thirteen indicators classify and discriminate the well-being of urban residents. According to different values, cities are classified differently, which better reflects the differences and similarities between cities, facilitates mutual learning and reference between cities, and provides a theoretical basis for the development direction and strategy of cities and the formulation of their policies.

Acknowledgment

The first author is a student from Hunan City University. This paper is directed by Professor Hean Liu.

Reference :

[1] Xie Jijian, Liu Chengping. Fuzzy Mathematics Method and Its Application. Huazhong University of Science and Technology Press, 200
[2] Yang Lun Biao, Gao Yingyi. Principle and Application of Fuzzy Mathematics. South China University of Technology Press, 200
[3] Wang Peizhuang. Fuzzy set theory and its application. Shanghai Science and Technology Publishing House, 1983:78-92
[4] Luo Chengzhong. Introduction to Fuzzy Sets. Beijing Normal University Press, 1989:56-78
[5] Zhang Yue et al. Fuzzy mathematics method and its application. Coal Industry

ISSN 2522-6398 Vol. 2, Issue 4: 51-59, DOI: 10.25236/FER.034017

Press, 1992,4

[6] Yang Lunbiao, Gao Yingyi. Principle and Application of Fuzzy Mathematics. South China University of Technology Press, 200 [7] Tang Yicheng. A Class of Cluster Analysis Method . Practice and Knowledge of Mathematics, 1988, (2): 99-108

[7] Li Hongxing and Yen V.C. Fuzzy sets and fuzzy decision making. CRC Press: FL, 1995

[8] Li Hongxing and Yen V.C. Fuzzy sets and fuzzy decision making . CRC Press: FL, 1995

[9] Mo Zhiwen, Shulan, Xu Biao. Comments on Fuzzy Mathematics Theory and Its Application . Journal of Sichuan Normal University, 1998

[10] Fu Yunpeng, Ma Shucai. Clustering of Urban Comprehensive Competitiveness Based on Fuzzy Clustering Analysis . Journal of Hebei Normal University of Science and Technology, 2015

[11] Zhang Jiehui. Improving the Comprehensive Competitiveness of Cities . Economic Daily, 2007

[12] Hi-tech Wave, Xie Weixin. Research Progress in Development and Application of Fuzzy Clustering Theory . Scientific Bulletin, 1999

[13] He Zhongxiong. Fuzzy Mathematics and Its Application. Tianjin: Tianjin Science and Technology Press, 1983.