Fuzzy Clustering with New Separable Criterion

HSIANG-CHUAN LIU Department of Bioinformatics Asia University No. 500, Lioufeng Rd., Wufeng , Taichung County, 41354 TAIWAN lhc@asia.edu.tw

DER-BANG WU Graduate Institute of Educational Measurement Taichung University No. 140, Ming-Sheng Rd., Taichung, 40306 TAIWAN wudb@hotmail.com

HSIU-LAN MA Department of Business Administration Ling Tung, Taiwan No. 1, Ling Tung Rd., Taichung, 40852 TAIWAN hlma@mail.ltu.edu.tw

Abstract: - Fuzzy clustering has been used widely in education, statistics, engineering, communication...etc. The fuzzy partition clustering algorithms are most based on Euclidean distance function, which can only be used to detect spherical structural clusters. Extending Euclidean distance to Mahalanobis distance, Gustafson-Kessel (GK) clustering algorithm and Gath-Geva (GG) clustering algorithm were developed to detect non-spherical structural clusters, but these two algorithms fail to consider the relationships between cluster centers in the objective function. Yin-Tang-Sun-Sun (YTSS) clustering algorithm solved the relationships between cluster centers question, unfortunately, they did not consider the distance between the center of all data points and the center of each cluster. This problem was solved and presented in this paper.

In this paper, a new fuzzy clustering algorithm (LWM) was developed based on the conventional fuzzy cmeans (FCM) to obtain better quality clustering results with new separable criterion and better initial value. It is different from YTSS cluster algorithm. The improved equations for the membership and the cluster center were derived from the alternating optimization algorithm. Ten fuzzy scattering matrices in the objective function assure the compactness between data points and cluster centers, and also strengthen the separation between cluster centers in terms of a new separable criterion. The conclusions were drawn from this study as follows. (a) The distance between the center of all points and the center of each cluster was considered to obtain more accurate clustering results. (b) The singular problem was solved by using LWM algorithm. (c) The β value would not be required. The β value could be replaced by the best value, produced by the new LWM algorithm, developed by the authors of this paper. (d) Numerical data show that the LWM clustering algorithm gave more accurate clustering results than the FCM and YTSS clustering algorithm.

Key-Words: - GG clustering algorithm, GK clustering algorithm, YTSS clustering algorithm.

1. Introduction

Clustering groups the data into classes or clusters so that the data objects within a cluster have high similarity in comparison to one another, but are very dissimilar to those data objects in other clusters. Each point is assigned a membership to represent the degree of belonging to a certain class in fuzzy clustering. The fuzzy c-means (FCM) method is well known fuzzy clustering methods [2, 3]. The objective function is the trace of a within-cluster scatter matrix with spherical clusters obtained by minimizing the objective function with alternative optimization in FCM.

The Gustafson-Kessel (GK) [4] and Gath-Geva (GG) clustering algorithm [5] were developed to detect non-spherical structural clusters, but these two algorithms fail to consider the links between cluster centers in the objective function. The compatibilities of points with the cluster centers were guaranteed in a possibilistic c-means (PCM) method presented by Krish-napuram and Keller [6]. However, their algorithm exhibited bad behavior because there were no relationships among clusters [7]. The concepts of regularization were used by Ozdemir and Akarun [8] in an inter-cluster separation (ICS) algorithm and by Yang et al. [9] in a fuzzy compactness and separation (FCS) algorithm. But, the ICS algorithm had different objective functions for different cluster centers with the regulating term only considered as a perturbation. The FCS algorithm has hard kernel boundaries which depend on experiments and all the data points in one kernel could not be discriminated because they had identical member-ship values.

Yin et al. [1] described an extended objective function consisting of a fuzzy within-cluster scatter matrix and a new between-cluster centers scattering matrix. The corresponding fuzzy clustering algorithm assures the compactness between data points and cluster centers and also strengthens the separation between cluster centers based on the separation criterion. They clustering algorithm solved the relationships between cluster centers question, but they did not consider the distance between the center of all points and the center of each cluster. This problem was solved and presented in this paper. Moreover, In this paper, a new fuzzy clustering algorithm was developed based on the conventional fuzzy c-means (FCM) to obtain better quality clustering results with new separable criterion and better initial value. The improved equations for the membership and the cluster center were derived from the alternating optimization algorithm. Ten fuzzy scattering matrices in the objective function assure the compactness between data points and cluster centers,

and also strengthen the separation between cluster centers in terms of a new separable criterion. The distance between the center of all points and the center of each cluster was considered, the singular problem was solved, as well as, the β value would not be required. The β value will be replaced by the best value produced by the new algorithm, developed by the authors of this paper. Numerical data show that the new clustering algorithm gave more accurate clustering results than the FCM method and YTSS clustering algorithm.

2. New Objective Function

The clustering optimization was based on objective functions. The choice of an appropriate objective function is the point to the success of the cluster analysis. Let $\{z_l, z_2, , z_3, ..., z_n\}$ be a set of *n* data points represented by p-dimensional feature vectors $z_j = [z_{1j}, z_{2j}, ..., z_{pj}]^T \in \mathbb{R}^p$. The $p \times n$ data matrix *Z* has the cluster center matrix $V = [v_{l}, ..., v_c]$ (1<c<n) and the membership matrix $U = [\mu_{ij}]_{cxn}$, where μ_{ij} is the membership value of z_j belonging to v_i . $\hat{U} = [\hat{\mu}_{ik}]_{cxc}$ expresses the weighting matrix, and $\hat{\mu}_{ik}$ is the weighting value between v_i and v_k . The fuzzy exponent *m* is greater than 1 [10]. Thus, the proposed objective function is

$$\overline{J} = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^{m} \left\| z_{j-v_{i}} \right\|^{2} - \frac{1}{c(c-1)} \sum_{i=1}^{c} \sum_{k=1}^{c} \widehat{\mu}_{ik}^{m} \left\| v_{i} - v_{k} \right\|^{2}$$
(1)

Such that $\mu_{ij} \in [0,1], \sum_{i=1}^{c} \mu_{ij} = 1, \forall j,$

$$0 < \sum_{j=1}^n \mu_{ij} < n, \forall i,$$

where $\hat{\mu}_{ik}$ is defined as

$$\hat{\mu}_{ij} = \frac{\left[\left\|\vec{z}-v_{j}\right\|^{2} + \left\|\vec{z}-v_{j}\right\|^{2}\right] - \min_{1 \le r, s \le c} \left[\left\|\vec{z}-v_{r}\right\|^{2} + \left\|\vec{z}-v_{s}\right\|^{2}\right]}{\max_{1 \le r, s \le c} \left[\left\|\vec{z}-v_{r}\right\|^{2} + \left\|\vec{z}-v_{s}\right\|^{2}\right] - \min_{1 \le r, s \le c} \left[\left\|\vec{z}-v_{r}\right\|^{2} + \left\|\vec{z}-v_{s}\right\|^{2}\right]}$$
(2)
where $\overline{z} = \left(\sum_{i=1}^{n} z_{j}\right)/n$ is the data set mass center.

The following Eq. (3) of YTSS algorithm was replaced by Eq. (2) LWM algorithm. In this change, there a lot of problems were improved.

$$\hat{\mu}_{ik} = \beta \frac{\min\{\|\bar{z} - v_i\|^2, \|\bar{z} - v_k\|^2\}}{\max\{\|\bar{z} - v_i\|^2, \|\bar{z} - v_k\|^2\}}, 0 \le \beta \le 1$$
(3)

3. New Fuzzy Clustering Algorithm

The goal of the clustering algorithm is to identify the cluster centers and the membership values by solving an optimization problem. Alternating optimization is a popular mathematical tool for the regular objective function-based fuzzy clustering algorithms. The optimal update equations can be obtained using the Lagrange method by setting the partial derivative of the Lagrange with respect to v_i and with respect to μ_{ij} equal to zero. Setting $\partial J / \partial \mu_{ij}$ equal to zero gives the update equation for μ_{ij} . The new fuzzy clustering algorithm can be summarized in the following steps:

Step 1: Determining the number of cluster; c and m-

value (let m=2), given converging error, $\varepsilon > 0$ (such as $\varepsilon = 0.001$), appropriately choose the initial membership matrix, such that the memberships are not all equal;

$$U^{(0)} = \begin{bmatrix} \mu_{11}^{(0)} & \mu_{12}^{(0)} \dots & \mu_{1n}^{(0)} \\ \mu_{21}^{(0)} & \mu_{22}^{(0)} \dots & \mu_{2n}^{(0)} \\ \dots & \dots & \dots \\ \mu_{c1}^{(0)} & \mu_{c2}^{(0)} \dots & \mu_{cn}^{(0)} \end{bmatrix} = \begin{bmatrix} \mu_{1}^{(0)}(\underline{x}_{1}) & \mu_{1}^{(0)}(\underline{x}_{2}) \dots & \mu_{1}^{(0)}(\underline{x}_{n}) \\ \mu_{2}^{(0)}(\underline{x}_{1}) & \mu_{2}^{(0)}(\underline{x}_{2}) \dots & \mu_{2}^{(0)}(\underline{x}_{n}) \\ \dots & \dots & \dots \\ \mu_{c}^{(0)}(\underline{x}_{1}) & \mu_{c}^{(0)}(\underline{x}_{2}) \dots & \mu_{c}^{(0)}(\underline{x}_{n}) \end{bmatrix}$$

Step 2: Calculate the wighting value \hat{u}_{ij} between the cluster centers v_i and v_j ,

$$\hat{\mu}_{ij} = \frac{\left[\left\|\bar{z} - v_i\right\|^2 + \left\|\bar{z} - v_j\right\|^2\right] - \min_{1 \le r, s \le c} \left[\left\|\bar{z} - v_r\right\|^2 + \left\|\bar{z} - v_s\right\|^2\right]}{\max_{1 \le r, s \le c} \left[\left\|\bar{z} - v_r\right\|^2 + \left\|\bar{z} - v_s\right\|^2\right] - \min_{1 \le r, s \le c} \left[\left\|\bar{z} - v_r\right\|^2 + \left\|\bar{z} - v_s\right\|^2\right]}$$

Step 3: Find the cluster centers v_i

$$v_i^{(k)} = \frac{\sum_{j=1}^n \left[\mu_{ij}^{(k-1)} \right]^m z_j - \frac{1}{c(c-1)} \sum_{k=1}^c u_{ik} v_k}{\sum_{j=1}^n \left[\mu_{ij}^{(k-1)} \right]^m - \frac{1}{c(c-1)} \sum_{k=1}^c u_{ik}} \quad i = 1, 2, \dots, c$$

Step 4: Calculate the membership value

$$\mu_{ij}^{(k)} = \left[\sum_{s=1}^{c} \left[\frac{\left\|z_{j} - v_{i}\right\|^{2}}{\left\|z_{j} - v_{s}\right\|^{2}}\right]^{\frac{1}{m-1}}\right]$$

Step 5: Increment k; until

$$\max_{1\leq i\leq c} \left\|v_i^{(k)}-v_i^{(k-1)}\right\|<\varepsilon$$

Singularities could not occur in these update equations.

4. Numerical Example

A numerical example using the data set of sample size 493 grades three students from elementary schools was selected. These data included the independent variables, test scores of four mathematics concepts (division, ordering, multiplication, and place value) and 10 questions.

At first, the main factors of 493 data were calculated by using factor analysis. Next, according to the main factors, the samples were assigned to 4 clusters based on the clustering analysis using the k-mean clustering of SPSS for Windows 10.0. The results were shown in Table 1.

Table 1

The characteristics of 4 clusters

Cluster	Samples	mathematics	average distance from	
		concepts	center of cluster	
1	115	division	1.2576760	
2	128	ordering	1.2968550	
3	168	multiplication	1.1244569	
4	82	place value	1.7861002	

From Cluster 1, 15 samples randomly were selected, 15 from cluster 2, 15 from cluster 3, and 5 from cluster 4.

A comparison of different clustering algorithms to compute cluster's values with distinct adjustable factor was shown in Table 2.

Table 2

A comparison of different clustering algorithms

	00				
Clustering Algorithms	Cluster's Values				
LWM clustering	Classification Accuracies,				
algorithm	RMSE,				
	OFV,				
	NI				
YTSS clustering	Classification Accuracies,				
algorithm	RMSE,				
_	OFV,				
	NI				
	Classification Accuracies,				
FCM clustering	RMSE,				
algorithm	OFV,				
	NI				
Adjustable Factor					
$\beta = 0.1, 0.2, 0.3, \dots, 0.9, 1$					

The FCM algorithm, YTSS algorithm, and LWM clustering algorithm are compared using the following criteria for the cluster center locations: the root mean square error (RMSE, Eq. 4) of the centers, the objective function values (OFV), and the number of

iterations (NI). The clustering accuracies of different clustering algorithms are displayed in Table 3.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{c} \|\hat{v}_{i} - v_{i}\|^{2}}{c}}$$
(4)

where \hat{v}_i is the computed center and v_i the true center.

Table 3

Clustering accuracies of different clustering algorithms

M T	β	C Ac. (%)	RMSE	OFV	NI
L W M	1.0	44%	2.2940314	5487.3052	29
	0.9	44%	2.2884425	5471.2785	30
	0.8	44%	2.2828645	5455.1279	30
	0.7	44%	2.2732821	5434.0554	25
	0.6	42%	2.2717919	5429.8446	31
	0.5	42%	2.2668733	5412.5695	32
	0.4	42%	2.2620397	5390.6402	32
	0.3	42%	2.2575088	5374.5353	33
	0.2	40%	2.2522548	5357.4788	33
	0.1	40%	2.2482919	5341.8294	34
T Y S S	1.0	38%	1.8917183	5406.2721	37
	0.9	38%	1.8880599	5406.1152	34
	0.8	38%	1.8819993	5402.1438	34
	0.7	38%	1.8760029	5394.445	34
	0.6	38%	1.8661021	5371.584	34
	0.5	38%	1.8600801	5362.0084	33
	0.4	42%	1.8661021	5369.1405	32
	0.3	42%	2.2417878	5359.9644	34
	0.2	40%	2.241293	5347.2138	34
	0.1	40%	2.2429972	5336.4053	34
F			2.2448315	5325.6135	
С	0	40%			34
М					

Note: MT means Methods; C Ac. Means clustering accuracies

From the data shown in Table 3, no mater what the β values were picked up, numerical data show that the LWM clustering algorithm gave more accurate clustering results than the FCM and YTSS clustering algorithm. The LWM algorithm presented the best clustering accuracies, up to 44%.

5. Conclusions

An improved fuzzy clustering algorithm, LWM algorithm, is developed to obtain better quality of fuzzy clustering results. Comparison of LWM, YTSS, and FCM clustering algorithm shows that the LWM clustering algorithm will increase the cluster

compactness and the separation between clusters. Finally, the conclusions were drawn from this study as follows. (a) The distance between the center of all points and the center of each cluster was considered to obtain more accurate clustering results. (b) The singular problem was solved by using LWM algorithm. (c) The β value would not be required. The β value could be replaced by the best value, produced by the new LWM algorithm, developed by the authors of this paper. (d) Numerical data show that the LWM clustering algorithm gave more accurate clustering results than the FCM and YTSS clustering algorithm.

6. References

- Z. H. Yin, Y. G. Tang, F. C. Sun and Z. Q. Sun, "Fuzzy Clustering with Novel Separable Criterion," *Tsinghua science and technology*, 11(1), pp. 50-53, Feb. 2006.
- [2] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact wellseparated clusters," *J. Cybernetics*, 3(3), pp. 32-57, 1973.
- J. C. Bezdek, *Pattern recognition with fuzzy* objective function algorithms. New York: Plenum Press, pp. 65-70, 1 981.
- [4] D. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," *Proceedings of IEEE Conference Decision Control.* San Diego, CA, pp. 761-766, 1979.
- [5] I. Gath, and A. B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, 773-781, Jul. 1989.
- [6] R. Krishnapuram and J. M. Keller, "A possibilistic approach to clustering," *IEEE Trans. Fuzzy Systerms*, 1(2), pp. 98-110, 1993.
- [7] M. Barni, V. Cappellini and A. Mecocci, Comments on "A possibilistic approach to clustering," *IEEE Trans. Fuzzy Systerms*, 4(3), pp. 393-396, 1996.
- [8] D. Ozdemir and L. Akarun, "Fuzzy algorithms for combined quantization and dithering," *IEEE Trans. Image Processing*, 10(6), pp. 923-9312001.
- [9] M. S. Yang, K. L. Wu and J. Yu, "A novel fuzzy clustering algorithm," *Computational Intelligence in Robotics and Auto mation IEEE International Symposium*, 2, pp. 647-652, 2003.
- [10] J. Yu, Q. Cheng and H. Huang, "Analysis of the weighting exponent in the FCM," *IEEE Trans. Systems, Man and Cybernetics, Part B*, 34(1), 634-639, 2004.