Fuzzy Possibility C-Mean Based on Mahalanobis Distance and Separable Criterion

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

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

JENG-MING YIH
Graduate Institute of Educational Measurement
Taichung University
No. 140, Ming-Sheng Rd., Taichung, 40306
TAIWAN

SHIN-WU LIU
Program of Cell and Developmental Biology
Rutgers University
Piscataway, New Jersey 08854
USA

Abstract: - The well known fuzzy partition clustering algorithms are most based on Euclidean distance function, which can only be used to detect spherical structural clusters. Gustafson-Kessel (GK) clustering algorithm and Gath-Geva (GG) clustering algorithm, were developed to detect non-spherical structural clusters, but both of them based on semi-supervised Mahalanobis distance needed additional prior information. An improved Fuzzy C-Mean algorithm based on unsupervised Mahalanobis distance, FCM-M, was proposed by our previous work, but it didn't consider the relationships between cluster centers in the objective function. In this paper, we proposed an improved Fuzzy C-Mean algorithm, FPCM-MS, which is not only based on unsupervised Mahalanobis distance, but also considering the relationships between cluster centers, and the relationships between the center of all points and the cluster centers in the objective function, the singular and the initial values problems were also solved. A real data set was applied to prove that the performance of the FPCM-MS algorithm gave more accurate clustering results than the FCM and FCM-M methods, and the ratio method which is proposed by us is the better of the two methods for selecting the initial values.

Key-Words: - FPCM-MS; FPCM-M, FCM-M; GK algorithms; GG algorithms; Mahalanobis distance

1 Introduction

Fuzzy partition clustering is a branch in cluster analysis, it is widely used in pattern recognition field. The well known ones, such as, C. Bezdek's "Fuzzy C-Mean (FCM)" [1], are all based on Euclidean distance function, which can only be used to detect the data classes with same super spherical shapes.

Extending Euclidean distance to Mahalanobis distance, the well known fuzzy partition clustering algorithms, Gustafson-Kessel (GK) clustering algorithm [2] and Gath-Geva (GG) clustering algorithm [3] were developed to detect non- spherical structural clusters, but these two algorithms fail to consider the relationships between cluster centers in the objective function, GK algorithm must have prior information of shape volume in each data class, otherwise, it can only be considered to detect the data classes with same volume. GG algorithm must have prior probabilities of the clusters. An improved algorithm, based on Mahalanobis distance, "Fuzzy Possibility C-Mean Based on Mahalanobis Distance and Separable Criterion (FPCM-MS)" is proposed by our previous work [4, 5, 6, 7, 8].

Yin et al. [9] 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. Then 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 also solved and presented in this paper. Moreover, In this paper, an improved fuzzy clustering algorithm, denoted FPCM-MS, was developed based on FCM-M 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. The distance between the center of all points and the center of each cluster was considered by the authors of this paper, the singular problem was also solved. A real data set was applied to prove that the performance of the FPCM-MS algorithm gave more accurate clustering results than the FCM-M and FCM methods, and the ratio method which is proposed by us is the better of the two methods for selecting the initial values.

2 Some Exist Algorithms

2.1 Fuzzy c-Mean Algorithm

Fuzzy C-Mean Algorithm (FCM) is the most popular objective function based fuzzy clustering algorithm, it is first developed by Dunn [10] and improved by Bezdek [1].The objective function used in FCM is given by Equation (1).

$$J_{FCM}^{m}(U,A,X)=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}d_{ij}^{2}=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left\|\underline{x}_j-\underline{a}_i\right\|^2 \qquad (1)$$

$\mu_{ij}\in[0,1]$ is the membership degree of data object $\underline{x}_j$ in cluster C_i, and it satisfies the following constraint given by Equation (2).

$$\sum_{i=1}^{c}\mu_{ij}=1,\ \forall j=1,2,\ldots,n \qquad (2)$$

C is the number of clusters, m is the fuzzifier, m>1,which controls the fuzziness of the method. They are both parameters and need to be specified before running the algorithm. $d_{ij}^2=\left\|\underline{x}_j-\underline{a}_i\right\|^2$ is the square Euclidean distance between data object $\underline{x}_j$ to center $\underline{a}_i$.

Minimizing objective function (1) with constraint (2) is a non-trivial constraint nonlinear optimization problem with continuous parameters $\underline{a}_i$ and descrete parameters μ_{ij}. So there is no obvious analytical solution. Therefore an alternating optimization scheme, alternatively optimizing one set of parameters while the other set of parameters are considered as fixed, is used here. Then the updating function for $\underline{a}_i$ and μ_{ij} is obtained as Eq. (3) to Eq. (5).

Step 1: Determining the number of cluster; c and m-value (let m=2), given converging error, $\varepsilon>0$ (such as $\varepsilon=0.001$), randomly choose the initial membership matrix, such that the memberships are not all equal;

$$U^{(0)}=\begin{bmatrix}\mu_{11}^{(0)}\ \mu_{12}^{(0)}\ldots\mu_{1n}^{(0)}\\ \mu_{21}^{(0)}\ \mu_{22}^{(0)}\ldots\mu_{2n}^{(0)}\\ \ldots\ \ldots\ \ldots\\ \mu_{c1}^{(0)}\ \mu_{c2}^{(0)}\ldots\mu_{cn}^{(0)}\end{bmatrix}=\begin{bmatrix}\mu_{1}^{(0)}(\underline{x}_1)\ \mu_{1}^{(0)}(\underline{x}_2)\ldots\mu_{1}^{(0)}(\underline{x}_n)\\ \mu_{2}^{(0)}(\underline{x}_1)\ \mu_{2}^{(0)}(\underline{x}_2)\ldots\mu_{2}^{(0)}(\underline{x}_n)\\ \ldots\ \ldots\ \ldots\\ \mu_{c}^{(0)}(\underline{x}_1)\ \mu_{c}^{(0)}(\underline{x}_2)\ldots\mu_{c}^{(0)}(\underline{x}_n)\end{bmatrix} \qquad (3)$$

Step 2: Find

$$\underline{a}_i^{(k)}=\frac{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m\underline{x}_j}{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m}\quad i=1,2,\ldots,c \qquad (4)$$

$$\mu_{ij}^{(k)}=\left[\sum_{l=1}^{c}\left[\frac{\left(\underline{x}_j-\underline{a}_i^{(k)}\right)'\left(\underline{x}_j-\underline{a}_i^{(k)}\right)}{\left(\underline{x}_j-\underline{a}_l^{(k)}\right)'\left(\underline{x}_j-\underline{a}_l^{(k)}\right)}\right]^{\frac{1}{m-1}}\right]^{-1} \qquad (5)$$

Step 3: Increment k; until $\max_{1\le i\le c}\left\|\underline{a}_i^{(k)}-\underline{a}_i^{(k-1)}\right\|<\varepsilon$.

2.2 FCM-M Algorithm

For improving the above two problems, our previous work [4, 5, 6, 7, 8] proposed the improved algorithm FCM-M which added a regulating factor of covariance matrix, $-\ln\left|+\Sigma_i^{-1}\right|$, to each class in objective function, and deleted the constraint of the determinant of covariance matrices, $|M_i| = \rho_i$, in GK Algorithm as the objective function (6).

Using the Lagrange multiplier method, to minimize the objective function (6) with constraint (7) respect to parameters $\underline{a}_i$, μ_{ij}, Σ_i, we can obtain the solutions as (10), (11), and (13).

To avoid the singular problem and to select the better initial membership matrix , the updating functions for $\underline{a}_i$, μ_{ij}, and Σ_i are obtained as Eq. (8) to (13).

Note that:

(1).Both of FCM and FCM-M can not exploit all of the memberships with the same value.

(2).FCM is a special case of FCM-M, when covariance matrices equal to identity matrices.

$$J^m_{FCM-M}(U, A, \Sigma, X) = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^m\left[(\underline{x}_j - \underline{a}_i)'\Sigma_i^{-1}(\underline{x}_j - \underline{a}_i) - \ln\left|\Sigma_i^{-1}\right|\right] \quad (6)$$

Constraints: membership,

$$\sum_{i=1}^{c}\mu_{ij} = 1, \ \forall j = 1, 2, \ldots, n \quad (7)$$

Where $\Sigma = \{\Sigma_1, \Sigma_2, \ldots, \Sigma_c\}$ is the set of covariance of cluster.

Step 1: Determining the number of cluster; c and m-value (let m=2), given converging error, (such as), randomly choose the initial membership matrix, such that the memberships are not all equal;Step 1: Determining the number of cluster; c and m-value (let m=2), given converging error, $\varepsilon > 0$ (such as $\varepsilon = 0.001$).

Method 1: choose the result membership matrix of FCM algorithm as the initial one.

Method 2: $\underline{a}_i^{(0)}, i = 1, 2, \ldots, c$ let be the result centers of k-mean algorithm, and $d_{ij} = \left\|\underline{x}_j - \underline{a}_i^{(0)}\right\|$ be distances between data object $\underline{x}_j$ to center $\underline{a}_i^{(0)}$.

$$d_i^M = \max_{1\le j\le n} d_{ij} = \max_{1\le j\le n}\left\|\underline{x}_j - \underline{a}_i^{(0)}\right\|, d_i^m = \min_{1\le j\le n} d_{ij} = \min_{1\le j\le n}\left\|\underline{x}_j - \underline{a}_i^{(0)}\right\| \quad (8)$$

$$u_{ij}^{(0)} = \frac{d_i^M - d_{ij}}{d_i^M - d_i^m}, \ i = 1, 2, \ldots, c, \ j = 1, 2, \ldots, n$$

$$U^{(0)} = \begin{bmatrix} \mu_{11}^{(0)}\ \mu_{12}^{(0)} \ldots \mu_{1n}^{(0)} \\ \mu_{21}^{(0)}\ \mu_{22}^{(0)} \ldots \mu_{2n}^{(0)} \\ \ldots \quad \ldots \quad \ldots \\ \mu_{c1}^{(0)}\ \mu_{c2}^{(0)} \ldots \mu_{cn}^{(0)} \end{bmatrix} = \begin{bmatrix} \mu_1^{(0)}(\underline{x}_1)\ \mu_1^{(0)}(\underline{x}_2) \ldots \mu_1^{(0)}(\underline{x}_n) \\ \mu_2^{(0)}(\underline{x}_1)\ \mu_2^{(0)}(\underline{x}_2) \ldots \mu_2^{(0)}(\underline{x}_n) \\ \ldots \quad \ldots \quad \ldots \\ \mu_c^{(0)}(\underline{x}_1)\ \mu_c^{(0)}(\underline{x}_2) \ldots \mu_c^{(0)}(\underline{x}_n) \end{bmatrix} \quad (9)$$

Step 2: Find

$$\Sigma_i^{(k)} = \frac{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m(\underline{x}_j - \underline{a}_i^{(k)})(\underline{x}_j - \underline{a}_i^{(k)})'}{\sum_{j=1}^{n}\left[\mu_{ij}^{(k-1)}\right]^m} \quad (10)$$

$$\Sigma_i^{(k)} = \sum_{s=1}^{p}\lambda_{si}^{(k)}\Gamma_{si}^{(k)}\left(\Gamma_{si}^{(k)}\right)',$$

$$\left[\lambda_{si}^{(-1)}\right]^{(k)} = \begin{cases} \left[\lambda_{si}^{(k)}\right]^{-1} & if\ \lambda_{si}^{(k)} > 0 \\ 0 & if\ \lambda_{si}^{(k)} = 0 \end{cases}$$

$$\left[\Sigma_i^{-1}\right]^{(k)} = \sum_{s=1}^{p}\left[\lambda_{si}^{(-1)}\right]^{(k)}\Gamma_{si}^{(k)}\left(\Gamma_{si}^{(k)}\right)' \quad (11)$$

$$\left|\Sigma_i^{-1}\right|^{(k)} = \prod_{1\le s\le p,\ \lambda_{si}^{(k)}>0}\left[\lambda_{si}^{(-1)}\right]^{(k)} \quad (12)$$

$$\mu_{ij}^{(k)} = \left[\sum_{s=1}^{c}\left[\frac{(\underline{x}_j - \underline{a}_i^{(k)})'\left[\Sigma_i^{-1}\right]^{(k)}(\underline{x}_j - \underline{a}_i^{(k)}) - \ln\left|\left[\Sigma_i^{-1}\right]\right|^{(k)}}{(\underline{x}_j - \underline{a}_s^{(k)})'\left[\Sigma_s^{-1}\right]^{(k)}(\underline{x}_j - \underline{a}_s^{(k)}) - \ln\left|\left[\Sigma_s^{-1}\right]\right|^{(k)}}\right]^{\frac{1}{m-1}}\right]^{-1} \quad (13)$$

Step 3: Increment k; until $\max_{1\le i\le c}\left\|\underline{a}_i^{(k)} - \underline{a}_i^{(k-1)}\right\| < \varepsilon$.

3 New Algorithm - FPCM-MS

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. In FCM-M algorithm, it didn't consider the relationships between cluster centers in the objective function, now, we proposed an improved Fuzzy C-Mean algorithm, FPCM-MS, which is not only based on unsupervised Mahalanobis distance, but also considering the relationships between cluster centers, and the relationships between the center of all points and the cluster centers in the objective function, the singular and the initial values problems were also solved. Let $\{x_1, x_2, x_3, \ldots, x_n\}$ be a set of n data points represented by p-dimensional feature vectors $x_j = (x_{1j}, x_{2j}, \ldots, x_{pj})' \in R^p$. The $p \times n$ data matrix Z has the cluster center matrix $A=[a_1, \ldots, a_c]$, 1<c<n and the membership matrix $U = [\mu_{ij}]_{cxn}$, where μ_{ij} is the membership value of x_j belonging to a_i. $V = [v_{ik}]_{cxc}$ express the weighting matrix, and v_{ik} is the weighting value between v_i and v_k. The fuzzy exponent m is greater than 1 [11]. Thus, we can

obtained the objective function of Fuzzy Possibility c-Mean based on Mahalanobis distance (FPCM-M) as following

$$J^{m}_{FPCM-MS}(U,A,T,\Sigma,X)= \sum_{i=1}^{c}\sum_{j=1}^{n}\left(\mu_{ij}^{m}+t_{ij}^{\delta}\right)\left[\left(\underline{x}_j-\underline{a}_i\right)'\Sigma_i^{-1}\left(\underline{x}_j-\underline{a}_i\right)-\ln\left|\Sigma_i^{-1}\right|\right] - \frac{1}{c(c-1)}\sum_{i=1}^{c}\sum_{l=1}^{c}v_{il}^{m}\left(\underline{a}_i-\underline{a}_l\right)'\left(\underline{a}_i-\underline{a}_l\right) \quad (14)$$

constraints：membership $\sum_{i=1}^{c}\mu_{ij}=1$, $\forall j=1,2,\ldots,n$

typicality $\sum_{j=1}^{n}t_{ij}=1$, $\forall i=1,2,\ldots,c$

$$v_{il}^{(k)}=\frac{w_{il}^{k-1}-\min\limits_{1\le r,s\le c}w_{rs}^{k-1}}{\max\limits_{1\le r,s\le c}w_{rs}^{k-1}-\min\limits_{1\le r,s\le c}w_{rs}^{k-1}}, \quad (15)$$

where $w_{rs}^{k-1}=\left\|\underline{a}-\underline{a}_r^{(k-1)}\right\|^2+\left\|\underline{a}-\underline{a}_s^{(k-1)}\right\|^2$

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 $\varepsilon=0.001$).

Method 1: choose the result membership matrix of FCM algorithm as the initial one.

Method 2: let $\underline{a}_i^{(0)}, i=1,2,\ldots,c$ be the result centers of k-mean algorithm, and be distances between data object $\underline{x}_j$ to center $\underline{a}_i^{(0)}$.

$$z_{ij}^{(0)}=\begin{cases}\dfrac{d_M-d_{ij}}{d_M-d_m}, & if\ d_M>d_m\\ r_{ij}\sim U[0,1],\sum_{i=1}^{c}\left(r_{ij}-\bar{r}_j\right)^2>0, & if\ d_M=d_m\end{cases} \quad (16)$$

$i=1,2,\ldots,c$; $j=1,2,\ldots,n$

$d_M=\max\limits_{1\le i\le c,1\le j\le n}d_{ij}, d_m=\min\limits_{1\le i\le c,1\le j\le n}d_{ij}, \bar{r}_j=\frac{1}{c}\sum_{i=1}^{c}r_{ij}$

$$\mu_{ij}^{(0)}=\frac{z_{ij}}{\sum_{s=1}^{c}z_{sj}},\ i=1,2,\ldots,c,\ j=1,2,\ldots,n \quad (17)$$

$$t_{ij}^{(0)}=\frac{z_{ij}}{\sum_{s=1}^{n}z_{is}},\ i=1,2,\ldots,c,\ j=1,2,\ldots,n$$

$$\underline{a}_i^{(0)}=\left(\sum_{j=1}^{n}\left[\mu_{ij}^{(0)}\right]\underline{x}_j\right)\left(\sum_{j=1}^{n}\left[\mu_{ij}^{(0)}\right]\right)^{-1},\quad i=1,2,\ldots,c \quad (18)$$

$$\Sigma_i^{(0)}=\left(\sum_{j=1}^{n}\left[\mu_{ij}^{(0)}\right]^m\left(\underline{x}_j-\underline{a}_i^{(0)}\right)\left(\underline{x}_j-\underline{a}_i^{(0)}\right)'\right)\left(\sum_{j=1}^{n}\left[\mu_{ij}^{(0)}\right]^m\right)^{-1} \quad (19)$$

$$U^{(0)}=\begin{bmatrix}\mu_{11}^{(0)}\ \mu_{12}^{(0)}\ldots\mu_{1n}^{(0)}\\ \mu_{21}^{(0)}\ \mu_{22}^{(0)}\ldots\mu_{2n}^{(0)}\\ \ldots\ \ldots\ \ldots\\ \mu_{c1}^{(0)}\ \mu_{c2}^{(0)}\ldots\mu_{cn}^{(0)}\end{bmatrix}=\begin{bmatrix}\mu_1^{(0)}(\underline{x}_1)\ \mu_1^{(0)}(\underline{x}_2)\ldots\mu_1^{(0)}(\underline{x}_n)\\ \mu_2^{(0)}(\underline{x}_1)\ \mu_2^{(0)}(\underline{x}_2)\ldots\mu_2^{(0)}(\underline{x}_n)\\ \ldots\ \ldots\ \ldots\\ \mu_c^{(0)}(\underline{x}_1)\ \mu_c^{(0)}(\underline{x}_2)\ldots\mu_c^{(0)}(\underline{x}_n)\end{bmatrix} \quad (20)$$

$$T^{(0)}=\begin{bmatrix}t_{11}^{(0)}\ t_{12}^{(0)}\ldots t_{1n}^{(0)}\\ t_{21}^{(0)}\ t_{22}^{(0)}\ldots t_{2n}^{(0)}\\ \ldots\ \ldots\ \ldots\\ t_{c1}^{(0)}\ t_{c2}^{(0)}\ldots t_{cn}^{(0)}\end{bmatrix}=\begin{bmatrix}t_1^{(0)}(\underline{x}_1)\ t_1^{(0)}(\underline{x}_2)\ldots t_1^{(0)}(\underline{x}_n)\\ t_2^{(0)}(\underline{x}_1)\ t_2^{(0)}(\underline{x}_2)\ldots t_2^{(0)}(\underline{x}_n)\\ \ldots\ \ldots\ \ldots\\ t_c^{(0)}(\underline{x}_1)\ t_c^{(0)}(\underline{x}_2)\ldots t_c^{(0)}(\underline{x}_n)\end{bmatrix} \quad (21)$$

$$A^{(0)}=\left[\underline{a}_1^{(0)}\ \underline{a}_2^{(0)}\ldots\underline{a}_c^{(0)}\right] \quad (22)$$

$$\Sigma^{(0)}=\left[\Sigma_1^{(0)},\Sigma_2^{(0)},\ldots,\Sigma_c^{(0)}\right] \quad (23)$$

$$\Sigma^{(0)}=\left[\Sigma_1^{(0)},\Sigma_2^{(0)},\ldots,\Sigma_c^{(0)}\right] \quad (24)$$

Step 2: Find

$$v_{il}^{(k)}=\frac{w_{il}^{(k-1)}-\min\limits_{1\le r,s\le c}w_{rs}^{(k-1)}}{\max\limits_{1\le r,s\le c}w_{rs}^{(k-1)}-\min\limits_{1\le r,s\le c}w_{rs}^{(k-1)}}, \quad (25)$$

where

$$w_{rs}^{(k-1)} = \left\| \underline{a} - \underline{a}_r^{(k-1)} \right\|^2 + \left\| \underline{a} - \underline{a}_s^{(k-1)} \right\|^2$$

$$a_i^{(k)} = \left[F^{(k)}\right]^{-1} \left[\sum_{j=1}^{n} \left(\left[\mu_{ij}^{(k-1)}\right]^m + \left[t_{ij}^{(k-1)}\right]^\delta \right) \left[\Sigma_i^{(k-1)}\right]^{-1} x_j - d^{(k)} \right] \quad (26)$$

$$F^{(k)} = \left[\sum_{j=1}^{n} \left(\left[\mu_{ij}^{(k-1)}\right]^m + \left[t_{ij}^{(k-1)}\right]^\delta \right) \left[\Sigma_i^{(k-1)}\right]^{-1} - d^{(k)} I \right] \quad (27)$$

$$where \quad d^{(k)} = \frac{1}{c(c-1)} \sum_{l=1}^{c} \left[v_{il}^{(k)}\right]^m \underline{a}_l^{(k-1)} \quad , i = 1, 2, \ldots, c$$

$$\mu_{ij}^{(k)} = \left[\sum_{s=1}^{c} \left[\frac{\left(\underline{x}_j - \underline{a}_i^{(k)}\right)' \left[\Sigma_i^{(k)}\right]^{-1} \left(\underline{x}_j - \underline{a}_i^{(k)}\right) - \ln\left|\Sigma_i^{(k)}\right|^{-1}}{\left(\underline{x}_j - \underline{a}_s^{(k)}\right)' \left[\Sigma_s^{(k)}\right]^{-1} \left(\underline{x}_j - \underline{a}_s^{(k)}\right) - \ln\left|\Sigma_s^{(k)}\right|^{-1}} \right]^{\frac{1}{m-1}} \right]^{-1} \quad (28)$$

$$t_{ij}^{(k)} = \left[\sum_{s=1}^{n} \left[\frac{\left(\underline{x}_j - \underline{a}_i^{(k)}\right)' \left[\Sigma_i^{(k)}\right]^{-1} \left(\underline{x}_j - \underline{a}_i^{(k)}\right) - \ln\left|\Sigma_i^{(k)}\right|^{-1}}{\left(\underline{x}_s - \underline{a}_i^{(k)}\right)' \left[\Sigma_i^{(k)}\right]^{-1} \left(\underline{x}_s - \underline{a}_i^{(k)}\right) - \ln\left|\Sigma_i^{(k)}\right|^{-1}} \right]^{\frac{1}{\delta-1}} \right]^{-1} \quad (21)$$

where

$$\Sigma_i^{(k)} = \frac{\sum_{j=1}^{n} \left(\left[\mu_{ij}^{(k)}\right]^m + \left[t_{ij}^{(k)}\right]^\delta \right) \left(\underline{x}_j - \underline{a}_i^{(k)}\right) \left(\underline{x}_j - \underline{a}_i^{(k)}\right)'}{\sum_{j=1}^{n} \left(\left[\mu_{ij}^{(k)}\right]^m + \left[t_{ij}^{(k)}\right]^\delta \right)}$$

$$i = 1, 2, \ldots, c$$

$$\Sigma_i^{(k)} = \sum_{s=1}^{p} \lambda_{si}^{(k)} \Gamma_{si}^{(k)} \left(\Gamma_{si}^{(k)}\right)' ,$$

$$\left[\lambda_{si}^{-1}\right]^{(k)} = \begin{cases} \left[\lambda_{si}^{(k)}\right]^{-1} & if \ \lambda_{si}^{(k)} > 0 \\ 0 & if \ \lambda_{si}^{(k)} = 0 \end{cases}$$

$$\left[\Sigma_i^{(k)}\right]^{-1} = \sum_{s=1}^{p} \left[\lambda_{si}^{(-1)}\right]^{(k)} \Gamma_{si}^{(k)} \left(\Gamma_{si}^{(k)}\right)'$$

$$\left|\Sigma_i^{(k)}\right|^{-1} = \prod_{1 \le s \le p, \lambda_{si}^{(k)} > 0} \left[\lambda_{si}^{(k)}\right]^{-1}$$

Step 3: Increment k; until $\max_{1 \le i \le c} \left\| \underline{a}_i^{(k)} - \underline{a}_i^{(k-1)} \right\| < \varepsilon$.

4 Experiment

A real data set of students with sample size 493 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 the 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

Clu ster	sample size	mathematics concepts	average distance of the points from center of cluster
1	100	division	1.2879
2	82	ordering	1.7861
3	173	multiplication	1.1402
4	138	place value	1.2890

Each 15 sample points were randomly drawn from Cluster 1, cluster 2, and cluster 3, respectively, and 5 from cluster 4.

How to select the better initial value to improve the cluster accuracy is an important issue. In order to test the FCM-M algorithm, developed by the authors of this paper, the four .25 were selected as initial value. After calculating, the results were found that the memberships were all equal to .25 too. This evidence displayed that the FCM-M algorithm was work correctly.

There were 2 methods (*Ratio, Random*) to calculate the Normalized initial number which satisfied the Equation (2).

The steps of *Ratio Method* were as follows.

Step 1: The distance between observing value and every cluster center of every Point, say d. Compute the Average Distance of Clustering Result Marking Group.

$$cd_i^{(0)} = \sum_{j=1}^{n_i} d_j n_i^{-1} \quad n_i, number\ of\ \text{Result}\ Marking\ Group\ i$$

Step 2: Compute the Difference of d and the Average Distance of Clustering Result Marking Group

$$l_{ij} = \left| d_j - cd_i \right| \ j = 1, 2, \ldots, n_i, \ i = 1, 2, \ldots, c$$

Step 3: Find the values of maximum and minimum

$$M = \max\left\{ l_{ij} \ j = 1, 2, \ldots, n_i, \ i = 1, 2, \ldots, c \right\},$$

$$m = \min\left\{ l_{ij} \ j = 1, 2, \ldots, n_i, \ i = 1, 2, \ldots, c \right\}$$

Step 4: Compute the initial membership Difference of every Point $\mu_{ij} = (M - l_{ij})(M - m)^{-1}$.

The steps of *Random Method* were as follows. Choose any 4 random numbers, r1, r2, r3, and r4.

Step 1: Choose any 4 random numbers, r1, r2, r3, and r4, such that $0<r_i<1$, i=1, 2, 3, 4.

Step2: Let S= $r_1 + r_2 + r_3 + r_4$, then r_1/S, r_2/S, r_3/S, r_4/S were calculated as one set of initial value.

The classification accuracies of testing samples were shown in Table 2.

Table 2: Classification Accuracies of Testing Samples

choosing the initial membership	computing distance	Classification Accuracies (%)
Ratio	FCM	30
	PCM	32
	FPCM	36
	FPCM-M	44
	FPCM-MS	56
Random	FCM	12
	PCM	14
	FPCM	30
	FPCM-M	44
	FPCM-MS	54

From the data of Table 2, we found that the FPCM-MS algorithms of *Ratio Method* could obtain the best results, up to 56%. Next, the FPCM-MS algorithms of *Random Method* could obtain the second better results, up to 54%.

5 Conclusion

An improved new fuzzy clustering algorithm, FPCM-MS, is developed to obtain better quality of fuzzy clustering results. The objective function includes a fuzzy within-cluster scatter matrix, a new between-prototypes scatter matrix, the regulating terms about the covariance matrices, and the regulating terms about the relationships between cluster centers, the relationships between the center of all points and the cluster centers. The update equations for the memberships and the cluster centers and the covariance matrices are directly derived from the Lagrange's method, which is different from the GK and GG algorithms. The singular problem and the selecting initial values problem are improved by the Eigenvalue method and the Ratio method. Finally, a numerical example shows that the new fuzzy clustering algorithm FPCM-MS gives more accurate clustering results than the FCM and FCM-M algorithms for a real data set, the ratio method which is proposed by us is the better of the two methods for selecting the initial values.

References:

[1] I. Gath, and A. B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, Jul. 1989, pp. 773-781.

[2] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, NY. 1981.

[3] D. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," *Proc. IEEE Conf. Decision Contr.*, San Diego, CA, 1979.

[4] H.-C. Liu, J.-M. Yih, and S.-W. Liu, "Fuzzy C-mean Algorithm Based On Mahalanobis Distances and better initial values," *International Conference on Fuzzy Theory & Technology, JCIS*, Salt Lake City, Utah, U.S.A. July 18-24, 2007.

[5] H.-C. Liu, D.-B. Wu, and H.-L. Ma, "Fuzzy Clustering with New Separable Criterion," *International Conference on Fuzzy Theory & Technology, JCIS*, Salt Lake City, Utah, U.S.A. July 18-24, 2007.

[6] H.-C. Liu, J.-M. Yih, T.-W. Sheu, and S.-W. Liu, "A new Fuzzy Possibility Clustering Algorithms Based On Unsupervised Mahalanobis Distances," *International Conference on Machine Learning and Cybernetics 2007*, Hong Kong, China, Aug. 19-22, 2007. (Accepted)

[7] H.-C. Liu, D.-B. Wu, H.-L. Ma, and C.-C. Chen, "Fuzzy Possibility C-Mean with New Separable Criterion," *International Conference on Machine Learning and Cybernetics 2007*, Hong Kong, China, Aug. 19-22, 2007. (Accepted)

[8] H.-C. Liu, J.-M.Yih, D.-B. Wu, and C.-C. Chen, "Fuzzy C-Mean Algorithm Based on Mahalanobis Distance and New Separable Criterion," *International Conference on Machine Learning and Cybernetics 2007*, Hong Kong, China, Aug. 19-22, 2007. (Accepted)

[9] Z. H. Yin, Y. G. Tang, F. C. Sun and Z. Q. Sun, "Fuzzy Clustering with Novel Separable Criterion", Tsinghua science and technology, Vol. 11, No. 1, Feb. 2006, pp. 50-53.

[10] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact, well-separated clusters," *J. Cybern*, Vol. 3, No. 3, 1973, pp. 32-57.

[11] J. Yu, Q. Cheng and H. Huang, "Analysis of the weighting exponent in the FCM," *IEEE Trans. Systems, Man and Cybernetics*, Part B, 2004, Vol.34, No. 1, pp. 634-639.