原著論文

Fuzzy Classification Function of Fuzzy c -Means Algorithms for Data with Tolerance

Yuchi KANZAWA *1 • Yasunori ENDO *2 • Sadaaki MIYAMOTO *2

In this paper, two fuzzy classification functions of fuzzy c -means for data with tolerance are proposed. First, two clustering algorithms for data with tolerance are introduced. One is based on the standard method and the other is on the entropy-based one. Second, the fuzzy classification function for fuzzy c -means without tolerance is discussed as the solution of a certain optimization problem. Third, two optimization problems are shown so that the solutions are the fuzzy classification function values for fuzzy c -means algorithms with respect to data with tolerance, respectively. Fourth, Karush-Kuhn-Tucker conditions of two objective functions are considered, and two iterative algorithms are proposed for the optimization problems, respectively. Through some numerical examples, the proposed algorithms are discussed.

Keywords : Clustering, Fuzzy c -Means, Data with Tolerance, Fuzzy Classification Function

1. Introduction

Fuzzy c -means (FCM) [1] is one of the well-known fuzzy clusterings and many FCM variants have been proposed after FCM. In these variants, FCM algorithm based on the concept of regularization by entropy has been proposed by one of the authors [2]. This algorithm is called regularized entropy FCM (eFCM) and is discussed not only for its usefulness but also for its mathematical relations with other techniques.

There are many cases that data have some errors in clustering. Many clustering algorithms have been proposed to classify the data with error. In these algorithms, two FCM algorithms considering the inner point in the region of data have been proposed by some of the authors [3]. The one is derived from the standard FCM and the other is from eFCM. Because these algorithms interpret the error of the data as an arbitrary point in a given region called “tolerance”, these are called FCMs for data with tolerance (FCM-Ts). In FCM-Ts, the tolerance is regarded as another decision variable composing the optimization problem like the membership grades and the cluster centers, and is determined as minimizing the objective function. The features of FCM-Ts have not been less clarified yet compared with FCM without tolerance.

Fuzzy classification functions for the standard FCM (sFCM) and eFCM without tolerance have been proposed [2], respectively, in order to show how prototypical an arbitrary point in the data space is to a cluster by extending the membership to the whole space. Fuzzy classification function can be used to classify a brandnew data to a cluster after fuzzy clustering (sFCM or eFCM) is done for the initially given data. It is a kind of supervised classification. Fuzzy classification function is also valid to investigate the features of the corresponding clustering algorithm since it clarify the classifying situation in whole space than only memberships for finite number of data. Actually, some features of FCM have been shown by its fuzzy classification function [2]. In this sense, fuzzy classification function should be recognized as a pair of tool with its original clustering algorithm than the adjunct of the one. In order to investigate the features of FCM-T, the one for FCM-T is essential, but it has not been proposed yet.

In this paper, we propose two fuzzy classification functions for sFCM-T and eFCM-T, respectively. Fuzzy classification for FCM without tolerance can be obtained by replacing the original datum in optimal solutions for memberships by the brand-new datum, but such the replacing cannot be applied to FCM-T. Therefore, ones for FCM-T are calculated by iterative algorithms obtained from certain optimization

*1 Shibaura Institute of Technology

*2 University of Tsukuba

problems, respectively, which are derived from that fuzzy classification function of FCM without tolerance can be interpreted as the solution of an optimization problem. By fuzzy classification function proposed in this paper, the features of FCM-T are expected to be clarified. We also show some numerical examples of such fuzzy classification functions.

The contents of this paper are the followings. In the second section, we define some notation, introduce sFCM and eFCM with their fuzzy classification functions, and also introduce sFCM-T and eFCM-T. In the third section, we propose two fuzzy classification functions, one of which is for sFCM-T and the other is for eFCM-T, by two iterative algorithms, respectively, which are led from two optimization problems based on that fuzzy classification function value for FCM without tolerance can be interpreted as the solution of a certain optimization problem. In the fourth section, some numerical examples are shown. In the last section, we conclude this paper.

2. Preliminaries

In this section, we define some notations, introduce sFCM and eFCM with their fuzzy classification function, and also introduce sFCM-T and eFCM-T. In the first subsection, we define some notations which are the data for clustering, the membership by which the each datum belongs to the each cluster, the cluster centers, the tolerance for the data and the maximum tolerance for the data. In the second subsection, we introduce sFCM and eFCM with their fuzzy classification functions. In the third subsection, we introduce sFCM-T and eFCM-T which are the basis of our main theme.

2.1 Notations

In this subsection, we define some notations which are the data for clustering, the membership by which the each datum belongs to the each cluster, the cluster centers, the tolerance for the data and the maximum tolerance for the data.

The data set $x = \{x_i | x_i \in \mathbb{R}^p, i \in \{1, \dots, N\}\}$ is given. The membership by which x_i belongs to the j -th cluster is denoted by $u_{i,j} (i \in \{1, \dots, N\}, j \in \{1, \dots, C\})$ and the set of $u_{i,j}$ is denoted by $u \in \mathbb{R}^{N \times C}$ called the partition matrix. The constraint for u is

$$\sum_{j=1}^C u_{i,j} = 1 \quad (0 \leq u_{i,j} \leq 1). \quad (1)$$

The cluster center set is denoted by $v = \{v_j | v_j \in \mathbb{R}^p, j \in \{1, \dots, C\}\}$. The tolerance for the data x is denoted by $\varepsilon = \{\varepsilon_i | \varepsilon_i \in \mathbb{R}^p, i \in \{1, \dots, N\}\}$. The maximum tolerance is denoted by $\kappa = \{\kappa_i | \kappa_i \in \mathbb{R}_+, i \in \{1, \dots, N\}\}$.

2.2 FCM and Fuzzy Classification Function

In this subsection, we introduce two types of FCM, which are sFCM[1] and eFCM[2] and also introduce their fuzzy classification functions.

sFCM is the algorithm obtained by solving the following optimization problem :

$$\underset{u,v}{\text{minimize}} J_m(u, v) \quad (2)$$

$$\text{under} \quad \sum_{j=1}^C u_{i,j} = 1, \quad (3)$$

where

$$J_m(u, v) = \sum_{i=1}^N \sum_{j=1}^C u_{i,j}^m \|x_i - v_j\|^2. \quad (4)$$

m is the fuzzifier parameter satisfying $m > 1$. In this paper, the symbol $\|\cdot\|$ stands for Euclidean norm. The optimal solutions u and v are obtained by the following algorithm.

Algorithm 1 (sFCM[1])

Step 1 Give the number of cluster C and the fuzzifier parameter m . Set the initial cluster centers v .

Step 2 Calculate u such that

$$u_{i,j} = 1 / \sum_{k=1}^C \left(\frac{d_{i,j}}{d_{i,k}} \right)^{\frac{1}{m-1}}, \quad (5)$$

Where

$$d_{i,j} = \|x_i - v_j\|^2. \quad (6)$$

Step 3 Calculate v such that

$$v_j = U_j^{-1} \sum_{i=1}^N u_{i,j}^m x_i, \quad (7)$$

Where

$$U_j = \sum_{i=1}^N u_{i,j}^m. \quad (8)$$

Step 4 Check the stopping criterion for (u, v) . If the criterion is not satisfied, go back to **Step 2**.

eFCM is the algorithm obtained by solving the following optimization problem :

$$\text{minimize}_{u,v} J_\lambda(u, v) \quad (9)$$

$$\text{under } \sum_{j=1}^C u_{i,j} = 1, \quad (10)$$

where

$$J_\lambda(u, v) = \sum_{i=1}^N \sum_{j=1}^C u_{i,j} \|x_i - v_j\|^2 + \lambda^{-1} \sum_{i=1}^N \sum_{j=1}^C u_{i,j} \log(u_{i,j}). \quad (11)$$

The second term of the right-hand side in Eq. (11) is for regularization by entropy, where λ is regularization parameter satisfying $\lambda > 0$. The optimal solutions u and v are obtained by the following algorithm.

Algorithm 2 (eFCM[2])

Step 1 Give the number of cluster C , and the regularization parameter λ . Set the initial cluster centers v .

Step 2 Calculate u such that

$$u_{i,j} = \left(\sum_{k=1}^C \exp(-\lambda d_{i,k}) \right)^{-1} \exp(-\lambda d_{i,j}), \quad (12)$$

Where

$$d_{i,j} = \|x_i - v_j\|^2. \quad (13)$$

Step 3 Calculate v such that

$$v_j = U_j^{-1} \sum_{i=1}^N u_{i,j} x_i, \quad (14)$$

Where

$$U_j = \sum_{i=1}^N u_{i,j}. \quad (15)$$

Step 4 Check the stopping criterion for (u, v) . If the criterion is not satisfied, go back to **Step 2**.

Fuzzy classification functions [2] are available in FCMs which show how prototypical an arbitrary point in the data space is to a cluster by extending the

membership $u_{i,j}$ to the whole space. Fuzzy classification function $\tilde{u}_j(\tilde{x})$ for sFCM with respect to a brand-new datum $\tilde{x} \in \mathbb{R}^p$ is defined as

$$\tilde{u}_j(\tilde{x}) = 1 / \sum_{k=1}^C \left(\frac{\tilde{d}_j}{\tilde{d}_k} \right)^{\frac{1}{m-1}}, \quad (16)$$

Where

$$\tilde{d}_j = \|\tilde{x} - v_j\|^2 \quad (17)$$

and where v_j is the j -th cluster center obtained from sFCM. Similarly, fuzzy classification function $\tilde{u}_j(\tilde{x})$ for eFCM with respect to a brand-new datum $\tilde{x} \in \mathbb{R}^p$ is defined as

$$\tilde{u}_j = \left(\sum_{k=1}^C \exp(-\lambda \tilde{d}_k) \right)^{-1} \exp(-\lambda \tilde{d}_j), \quad (18)$$

where

$$\tilde{d}_j = \|\tilde{x} - v_j\|^2 \quad (19)$$

and where v_j is the j -th cluster center obtained from eFCM. We can see that Eq. (16) - (19) are obtained by replacing the original datum x_i in Eq. (5), (6), (12) and (13) by the brand-new datum $\tilde{x}$. Fuzzy classification function is valid to investigate the features of FCM since it clarify the classifying situation in whole space than only memberships for finite number of data. Actually, some features of FCM have been shown using its fuzzy classification function [2]. In this sense, fuzzy classification function should be recognized as a pair of tool with its original clustering algorithm than the adjunct of the one.

2.3 FCM for data with Tolerance

In this subsection, we introduce two types of FCM-T, which are sFCM-T and eFCM-T[3]. These algorithms are the basis of our main theme.

sFCM-T is the algorithm obtained by solving the following optimization problem :

$$\text{minimize}_{u,\varepsilon,v} J_{m,t}(u, \varepsilon, v) \quad (20)$$

$$\text{under } \begin{cases} \sum_{j=1}^C u_{i,j} = 1, \\ \|\varepsilon_i\|^2 \leq \kappa_i^2 \quad (\kappa_i > 0), \end{cases} \quad (21)$$

where

$$J_{m,t}(u, \varepsilon, v) = \sum_{i=1}^N \sum_{j=1}^C u_{i,j}^m \|x_i + \varepsilon_i - v_j\|^2. \quad (22)$$

m is the fuzzifier parameter satisfying $m > 1$. In this paper, the symbol $\|\cdot\|$ stands for Euclidean norm. The optimal solutions u , ε and v are obtained by the following algorithm.

Algorithm 3 (sFCM-T[3])

Step 1 Give the number of cluster C , the maximum tolerance set κ for data x and the value m . Set the initial values of ε and v .

Step 2 Calculate u such that

$$u_{i,j} = 1 / \sum_{k=1}^C \left(\frac{d_{i,j}}{d_{i,k}} \right)^{\frac{1}{m-1}}, \quad (23)$$

Where

$$d_{i,j} = \|x_i + \varepsilon_i - v_j\|^2. \quad (24)$$

Step 3 Calculate ε such that

$$\varepsilon_i = -\alpha_i \left(\sum_{j=1}^C u_{i,j}^m (x_i - v_j) \right), \quad (25)$$

Where

$$\alpha_i = \min \left\{ \kappa_i \left\| \sum_{j=1}^C u_{i,j}^m (x_i - v_j) \right\|^{-1}, \left(\sum_{j=1}^C u_{i,j}^m \right)^{-1} \right\}. \quad (26)$$

Step 4 Calculate v such that

$$v_j = U_j^{-1} \sum_{i=1}^N u_{i,j}^m (x_i + \varepsilon_i), \quad (27)$$

Where

$$U_j = \sum_{i=1}^N u_{i,j}^m. \quad (28)$$

Step 5 Check the stopping criterion for (u, ε, v) . If the criterion is not satisfied, go back to **Step 2**.

If $\kappa_i = 0$ is set, ε_i is always zero and Algorithm 3 is equivalent to sFCM without tolerance.

eFCM-T is the algorithm obtained by solving the following optimization problem :

$$\text{minimize } J_{\lambda,t}(u, \varepsilon, v) \quad (29)$$

$$\text{under } \begin{cases} \sum_{j=1}^C u_{i,j} = 1, \\ \|\varepsilon_i\|^2 \leq \kappa_i^2 \quad (\kappa_i > 0), \end{cases} \quad (30)$$

where

$$J_{\lambda,t}(u, \varepsilon, v) = \sum_{i=1}^N \sum_{j=1}^C u_{i,j} \|x_i + \varepsilon_i - v_j\|^2 + \lambda^{-1} \sum_{i=1}^N \sum_{j=1}^C u_{i,j} \log(u_{i,j}). \quad (31)$$

The second term of the right-hand side in Eq. (31) is for regularization by entropy. The optimal solutions u , ε and v are obtained by the following algorithm.

Algorithm 4 (eFCM-T[3])

Step 1 Give the number of cluster C , the maximum tolerance set κ for data x and the value λ . Set the initial values of ε and v .

Step 2 Calculate u such that

$$u_{i,j} = \left(\sum_{k=1}^C \exp(-\lambda d_{i,k}) \right)^{-1} \exp(-\lambda d_{i,j}), \quad (32)$$

Where

$$d_{i,j} = \|x_i + \varepsilon_i - v_j\|^2. \quad (33)$$

Step 3 Calculate ε such that

$$\varepsilon_i = -\alpha_i \left(x_i - \sum_{j=1}^C u_{i,j} v_j \right), \quad (34)$$

Where

$$\alpha_i = \min \left\{ \kappa_i \left\| x_i - \sum_{j=1}^C u_{i,j} v_j \right\|^{-1}, 1 \right\}. \quad (35)$$

Step 4 Calculate v such that

$$v_j = U_j^{-1} \sum_{i=1}^N u_{i,j} (x_i + \varepsilon_i), \quad (36)$$

Where

$$U_j = \sum_{i=1}^N u_{i,j}. \quad (37)$$

Step 5 Check the stopping criterion for (u, ε, v) . If the criterion is not satisfied, go back to **Step 2**.

If $\kappa_i = 0$ is set, ε_i is always zero and Algorithm 4 is equivalent to eFCM without tolerance.

The features of FCM-T have not been less clarified yet compared with FCM without tolerance. In order to investigate the features of FCM-T, fuzzy classification function for FCM-T is essential in the sense that some important features of FCM without tolerance have been shown using its fuzzy classification function [2]. But the one for FCM-T has not been proposed yet. It is not so easy to define fuzzy classification function of FCM-T. We cannot apply Eq. (16) or (18) directly to FCM-T because we do not know the value of tolerance for brand-new datum $\tilde{x} \in \mathbb{R}^p$ yet.

3. Fuzzy Classification Function of FCM-T

In this section, we propose two fuzzy classification functions for sFCM-T (Algorithm 3) and eFCM-T (Algorithm 4), respectively.

In the first subsection, we interpret fuzzy classification function for FCM without tolerance as the solution of an optimization problem. In the second subsection, we define fuzzy classification function for sFCM-T as a new optimization problem which contains the tolerance of the data in the one derived from sFCM, and construct a new algorithm to obtain the fuzzy classification function value for sFCM-T by solving the optimization problem. In the third subsection, we do for eFCM-T.

3.1 Optimization Problem for Fuzzy Classification Function of FCM without Tolerance

In this subsection, we interpret fuzzy classification function for FCM without tolerance as the solution of an optimization problem. This interpretation is used for considering fuzzy classification functions of FCM-Ts.

Fuzzy classification function $\tilde{u}_j(\tilde{x})$ for sFCM with respect to a brand-new datum $\tilde{x} \in \mathbb{R}^p$ can be interpreted as the solution $\tilde{u}_j$ of the following optimization problem :

$$\underset{\tilde{u}}{\text{minimize}} \sum_{j=1}^C \tilde{u}_j^m \|\tilde{x} - v_j\|^2 \quad (38)$$

$$\text{under} \quad \sum_{j=1}^C \tilde{u}_j = 1 \quad (39)$$

because its Lagrange function

$$L_{m,\text{FCF}}(\tilde{u}) = \sum_{j=1}^C \tilde{u}_j^m \|\tilde{x} - v_j\|^2 + \gamma \left(\sum_{j=1}^C \tilde{u}_j - 1 \right) \quad (40)$$

has KKT condition as Eq. (16), where γ is Lagrange multiplier. Similarly, fuzzy classification function $\tilde{u}_j(\tilde{x})$ for eFCM with respect to a brand-new datum $\tilde{x} \in \mathbb{R}^p$ can be interpreted as the solution $\tilde{u}_j$ of the following optimization problem :

$$\underset{\tilde{u}}{\text{minimize}} \sum_{j=1}^C \tilde{u}_j \|\tilde{x} - v_j\|^2 + \lambda^{-1} \sum_{j=1}^C \tilde{u}_j \log(\tilde{u}_j) \quad (41)$$

$$\text{under} \quad \sum_{j=1}^C \tilde{u}_j = 1 \quad (42)$$

because its Lagrange function

$$L_{\lambda,\text{FCF}}(\tilde{u}) = \sum_{j=1}^C \tilde{u}_j \|\tilde{x} - v_j\|^2 + \lambda^{-1} \sum_{j=1}^C \tilde{u}_j \log(\tilde{u}_j) + \gamma \left(\sum_{j=1}^C \tilde{u}_j - 1 \right) \quad (43)$$

has KKT condition as Eq. (18), where γ is Lagrange multiplier.

Based on the above interpretation, fuzzy classifications function for two FCM-Ts are proposed by solving the corresponding optimization problems with tolerance in the next two subsections, respectively.

3.2 Fuzzy Classification Function for sFCM-T

In this subsection, we propose a fuzzy classification function for sFCM-T (Algorithm 3) as an iterative algorithm. First, we define an optimization problem whose component of the solution is the fuzzy classification function value for sFCM-T with respect to a brand-new datum. Secondly, we solve the optimization problem with KKT conditions. Finally, we construct a new algorithm from the solutions of the optimization problem.

We define the following optimization problem whose component $\tilde{u}$ of solution $(\tilde{u}, \tilde{\varepsilon})$ is the fuzzy classification value for sFCM-T with respect to a

brand-new datum $\bar{x} \in \mathbb{R}^p$:

$$\underset{\tilde{u}, \tilde{\varepsilon}}{\text{minimize}} J_{m,t,\text{FCF}}(\tilde{u}, \tilde{\varepsilon}) \quad (44) \quad \sum_{j=1}^C \tilde{u}_j = 1 \quad (54)$$

$$\text{under} \begin{cases} \sum_{j=1}^C \tilde{u}_j = 1, \\ \|\tilde{\varepsilon}\|^2 \leq \tilde{\kappa}^2 \quad (\tilde{\kappa} > 0), \end{cases} \quad (45)$$

Where

$$J_{m,t,\text{FCF}}(\tilde{u}, \tilde{\varepsilon}) = \sum_{j=1}^C \tilde{u}_j^m \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2 \quad (46)$$

and where v_j is the j -th cluster center obtained from sFCM-T. $\tilde{\varepsilon} \in \mathbb{R}^p$ is the tolerance for $\bar{x}$ and $\tilde{\kappa} \in \mathbb{R}_+$ is the maximum tolerance for $\bar{x}$. Its Lagrange function $L_{m,t,\text{FCF}}$ is as below :

$$L_{m,t,\text{FCF}}(\tilde{u}, \tilde{\varepsilon}) = \sum_{j=1}^C \tilde{u}_j^m \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2 + \gamma \left(\sum_{j=1}^C \tilde{u}_j - 1 \right) + \delta (\|\tilde{\varepsilon}\|^2 - \tilde{\kappa}^2), \quad (47)$$

where γ and δ are Lagrange multipliers. Karush-Kuhn-Tucker conditions of $L_{m,t,\text{FCF}}$ are below :

$$\frac{\partial L_{m,t,\text{FCF}}}{\partial \tilde{u}_j} = 0, \quad (48)$$

$$\frac{\partial L_{m,t,\text{FCF}}}{\partial \tilde{\varepsilon}} = 0, \quad (49)$$

$$\frac{\partial L_{m,t,\text{FCF}}}{\partial \gamma} = 0, \quad (50)$$

$$\delta \frac{\partial L_{m,t,\text{FCF}}}{\partial \delta} = 0, \quad (51)$$

$$\delta \geq 0. \quad (52)$$

The condition (48) implies

$$\tilde{u}_j = \left(-\frac{\gamma}{m} \right)^{1/(m-1)} \left(\frac{1}{\|\tilde{x} + \tilde{\varepsilon} - v_j\|^2} \right)^{1/(m-1)}. \quad (53)$$

From this with

derived from the condition (50), we have

$$\left(-\frac{\gamma}{m} \right)^{1/(m-1)} = 1 / \sum_{k=1}^C \left(\frac{1}{\|\tilde{x} + \tilde{\varepsilon} - v_k\|^2} \right)^{1/(m-1)}, \quad (55)$$

hence we obtain the optimal solution $\tilde{u}_j$ as

$$\tilde{u}_j = 1 / \sum_{k=1}^C \left(\frac{\tilde{d}_j}{\tilde{d}_k} \right)^{\frac{1}{m-1}}, \quad (56)$$

where

$$\tilde{d}_j = \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2. \quad (57)$$

The condition (49) implies

$$\tilde{\varepsilon} = -\tilde{\alpha} \sum_{j=1}^C \tilde{u}_j^m (\tilde{x} - v_j), \quad (58)$$

where

$$\tilde{\alpha} = \left(\sum_{j=1}^C \tilde{u}_j^m + \delta \right)^{-1}. \quad (59)$$

The condition (51) implies $\delta=0$ or $\|\tilde{\varepsilon}\|^2 = \tilde{\kappa}^2$. In the case of $\delta=0$, we have

$$\tilde{\alpha} = \left(\sum_{j=1}^C \tilde{u}_j^m \right)^{-1}. \quad (60)$$

In the case of $\|\tilde{\varepsilon}\|^2 = \tilde{\kappa}^2$, we have

$$\tilde{\alpha} = \frac{\tilde{\kappa}}{\left\| \sum_{j=1}^C \tilde{u}_j^m (\tilde{x} - v_j) \right\|}. \quad (61)$$

Therefore we have the optimal solution of $\tilde{\varepsilon}$ as

$$\tilde{\varepsilon} = -\tilde{\alpha} \left(\sum_{j=1}^C \tilde{u}_j^m (\tilde{x} - v_j) \right), \quad (62)$$

where

$$\tilde{\alpha} = \min \left\{ \tilde{\kappa} \left\| \sum_{j=1}^C \tilde{u}_j^m (\tilde{x} - v_j) \right\|^{-1}, \left(\sum_{j=1}^C \tilde{u}_j^m \right)^{-1} \right\}. \quad (63)$$

From the above discussion, we obtain the following iterative algorithm.

Algorithm 5

Step 1 Inherit m , v from Algorithm 3. Give the maximal tolerance value of $\tilde{\kappa}$ for $\tilde{x}$ and set $\tilde{\varepsilon} = 0$.

Step 2 Calculate $\tilde{u}$ such that

$$\tilde{u}_j = 1 / \sum_{k=1}^C \left(\frac{\tilde{d}_j}{\tilde{d}_k} \right)^{\frac{1}{m-1}}, \quad (64)$$

Where

$$\tilde{d}_j = \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2. \quad (65)$$

Step 3 Calculate $\tilde{\varepsilon}$ such that

$$\tilde{\varepsilon} = -\tilde{\alpha} \left(\sum_{j=1}^C \tilde{u}_j^m (\tilde{x} - v_j) \right), \quad (66)$$

Where

$$\tilde{\alpha} = \min \left\{ \tilde{\kappa} \left\| \sum_{j=1}^C \tilde{u}_j^m (\tilde{x} - v_j) \right\|^{-1}, \left(\sum_{j=1}^C \tilde{u}_j^m \right)^{-1} \right\}. \quad (67)$$

Step 4 Check the stopping criterion for $(\tilde{u}, \tilde{\varepsilon})$. If the criterion is satisfied, $\tilde{u}_j$ is the fuzzy classification function value with respect to $\tilde{x}$. Otherwise, go back to Step 2.

3.3 Fuzzy Classification Function for eFCM-T

In this subsection, we propose a fuzzy classification function for eFCM-T as an iterative algorithm. First, we define an optimization problem whose part of solutions is the fuzzy classification function value for eFCM-T with respect to a brand-new datum. Secondly, we solve the optimization problem with KKT conditions. Finally, we construct a new algorithm from the solutions of the optimization problem.

We define the following optimization problem whose component $\tilde{u}$ of solution $(\tilde{u}, \tilde{\varepsilon})$ is the fuzzy classification

value for eFCM-T with respect to a brand-new datum $\tilde{x} \in \mathbb{R}^p$:

$$\underset{u, \varepsilon}{\text{minimize}} J_{\lambda, t, \text{FCF}}(\tilde{u}, \tilde{\varepsilon}) \quad (68)$$

$$\text{under} \quad \begin{cases} \sum_{j=1}^C \tilde{u}_j = 1, \\ \|\tilde{\varepsilon}\|^2 \leq \tilde{\kappa}^2 \quad (\tilde{\kappa} > 0), \end{cases} \quad (69)$$

where

$$J_{\lambda, t, \text{FCF}}(\tilde{u}, \tilde{\varepsilon}) = \sum_{j=1}^C \tilde{u}_j \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2 + \lambda^{-1} \sum_{j=1}^C \tilde{u}_j \log(\tilde{u}_j), \quad (70)$$

and where v_j is the j -th cluster center obtained from eFCM-T. $\tilde{\varepsilon} \in \mathbb{R}^p$ is the tolerance for $\tilde{x}$ and $\tilde{\kappa} \in \mathbb{R}_+$ is the maximum tolerance for $\tilde{x}$. Its Lagrange function $L_{\lambda, t, \text{FCF}}$ is as below:

$$L_{\lambda, t, \text{FCF}} = \sum_{j=1}^C \tilde{u}_j \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2 + \lambda^{-1} \sum_{j=1}^C \tilde{u}_j \log(\tilde{u}_j) + \gamma \left(\sum_{j=1}^C \tilde{u}_j - 1 \right) + \delta (\|\tilde{\varepsilon}\|^2 - \tilde{\kappa}^2), \quad (71)$$

where γ and δ are Lagrange multipliers. Karush-Kuhn-Tucker conditions of $L_{\lambda, t, \text{FCF}}$ are below:

$$\frac{\partial L_{\lambda, t, \text{FCF}}}{\partial \tilde{u}_j} = 0, \quad (72)$$

$$\frac{\partial L_{\lambda, t, \text{FCF}}}{\partial \tilde{\varepsilon}} = 0, \quad (73)$$

$$\frac{\partial L_{\gamma, t, \text{FCF}}}{\partial \gamma} = 0, \quad (74)$$

$$\delta \frac{\partial L_{\gamma, t, \text{FCF}}}{\partial \delta} = 0, \quad (75)$$

$$\delta \geq 0. \quad (76)$$

The condition (72) implies

$$\tilde{u}_j = \exp(-\lambda \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2) \exp(-\lambda \gamma - 1). \quad (77)$$

From this with

$$\sum_{j=1}^C \tilde{u}_j = 1 \quad (78)$$

derived from the condition (74), we have

$$\exp(-\lambda\gamma - 1) = 1 / \sum_{k=1}^C \exp(-\lambda \|\tilde{x} + \tilde{\varepsilon} - v_k\|^2), \quad (79)$$

hence we obtain the optimal solution $\tilde{u}_j$ as

$$\tilde{u}_j = \left(\sum_{k=1}^C \exp(-\lambda \tilde{d}_k) \right)^{-1} \exp(-\lambda \tilde{d}_j), \quad (80)$$

where

$$\tilde{d}_j = \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2. \quad (81)$$

The condition (73) implies

$$\tilde{\varepsilon} = -\tilde{\alpha} \sum_{j=1}^C \tilde{u}_j (\tilde{x} - v_j), \quad (82)$$

where

$$\tilde{\alpha} = (1 + \delta)^{-1}. \quad (83)$$

The condition (75) implies $\delta=0$ or $\|\tilde{\varepsilon}\|^2 = \tilde{\kappa}^2$. In the case of $\delta=0$, we have

$$\tilde{\alpha} = 1. \quad (84)$$

In the case of $\|\tilde{\varepsilon}\|^2 = \tilde{\kappa}^2$, we have

$$\tilde{\alpha} = \tilde{\kappa} \left\| \sum_{j=1}^C \tilde{u}_j (\tilde{x} - v_j) \right\|^{-1}. \quad (85)$$

Therefore we have the optimal solution of $\tilde{\varepsilon}$ as

$$\tilde{\varepsilon} = -\tilde{\alpha} \left(\sum_{j=1}^C \tilde{u}_j (\tilde{x} - v_j) \right), \quad (86)$$

where

$$\tilde{\alpha} = \min \left\{ \tilde{\kappa} \left\| \sum_{j=1}^C \tilde{u}_j (\tilde{x} - v_j) \right\|^{-1}, 1 \right\}. \quad (87)$$

From the above discussion, we obtain the following iterative algorithm.

Algorithm 6

Step 1 Inherit λ, v from Algorithm 4. Give the maximum tolerance set $\tilde{\kappa}$ for $\tilde{x}$ and set $\tilde{\varepsilon}=0$.

Step 2 Calculate $\tilde{u}$ such that

$$\tilde{u}_j = \left(\sum_{k=1}^C \exp(-\lambda \tilde{d}_k) \right)^{-1} \exp(-\lambda \tilde{d}_j), \quad (88)$$

Where

$$\tilde{d}_j = \|\tilde{x} + \tilde{\varepsilon} - v_j\|^2. \quad (89)$$

Step 3 Calculate $\tilde{\varepsilon}$ such that

$$\tilde{\varepsilon} = -\tilde{\alpha} \left(\tilde{x} - \sum_{j=1}^C \tilde{u}_j v_j \right), \quad (90)$$

Where

$$\tilde{\alpha} = \min \left\{ \tilde{\kappa} \left\| \tilde{x} - \sum_{j=1}^C \tilde{u}_j v_j \right\|^{-1}, 1 \right\}. \quad (91)$$

Step 4 Check the stopping criterion for $(\tilde{u}, \tilde{\varepsilon})$. If the criterion is satisfied, $\tilde{u}_j$ is the fuzzy classification function value with respect to $\tilde{x}$. Otherwise, go back to Step 2.

4. Numerical Examples

In this section, we show some examples of fuzzy classification functions (Algorithm 5 and Algorithm 6) of FCM-T. In each example, 100 trials for each FCM-T algorithm (Algorithm 3 or Algorithm 4) with different initial cluster centers are tested and the solution with the minimal objective function value is selected as the final classification result. Using the obtained cluster centers v_j , fuzzy classification function values $\tilde{u}$ are calculated for points $\tilde{x}$ around the initially given data x by Algorithm 5 or Algorithm 6.

The Classified data set shown in Fig.1 is constructed by 300 elements in the two dimensional Euclidean space. We classify the data set into four clusters.

At the first, we show the cases of $\kappa_i = \tilde{\kappa} = 0$, which is equivalent to the established FCM without tolerance. Algorithm 3 (sFCM) with $m=2$ produced the classification result shown in Fig.2. Inheriting the cluster centers, Algorithm 5 produced the fuzzy classification functions $\tilde{u}_1(\tilde{x})$ and $\tilde{u}_2(\tilde{x})$ for the first and the second clusters shown in Fig.3 and Fig.4. The contour

Fig.1 Data

Fig.2 Classification Results of sFCM-T(Algorithm 3)
with $\kappa_i=0$ Fig.3 Fuzzy Classification Function $\bar{u}_1(\bar{x})$ for sFCM-T
(Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0$ Fig.4 Fuzzy Classification Function $\bar{u}_2(\bar{x})$ for sFCM-T
(Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0$

maps of them are shown in Fig.5 and Fig.6, respectively. Algorithm 4 (eFCM) with $\lambda=20$ produced the classification result shown in Fig.7. Inheriting the cluster centers, Algorithm 6 produced the fuzzy classification functions $\bar{u}_1(\bar{x})$ and $\bar{u}_2(\bar{x})$ for the first and the second clusters shown in Fig.8 and Fig.9. The contour maps of them are shown in Fig.10 and Fig.11,

Fig.5 Contour Map of Fuzzy Classification Function
 $\bar{u}_1(\bar{x})$ for sFCM-T(Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0$ Fig.6 Contour Map of Fuzzy Classification Function
 $\bar{u}_2(\bar{x})$ for sFCM-T(Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0$ Fig.7 Classification Results of eFCM-T(Algorithm 4)
with $\kappa_i=0$

Fig.8 Fuzzy Classification Function $\bar{u}_1(\bar{x})$ for eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0$

Fig.9 Fuzzy Classification Function $\bar{u}_2(\bar{x})$ for eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0$

Fig.10 Contour Map of Fuzzy Classification Function $\bar{u}_1(\bar{x})$ for eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0$

Fig.11 Contour Map of Fuzzy Classification Function $\bar{u}_2(\bar{x})$ for eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0$

Fig.12 Classification Results of sFCM-T (Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0.04$

respectively.

Next, we show the cases of positive values of $\kappa_i = \tilde{\kappa}$ with the same m or λ . sFCM-T with $\kappa = 0.04$ produced the classification result shown in Fig.12 and the contour map of fuzzy classification functions $\bar{u}_1(\bar{x})$ and $\bar{u}_2(\bar{x})$ for the first and the second clusters shown in Fig.13 and Fig.14, respectively, with tolerance vector $\bar{\varepsilon}$ for each grid point $\bar{x}$. eFCM-T with $\kappa = 0.03$ produced the classification result shown in Fig.15 and the contour map of fuzzy classification functions $\bar{u}_1(\bar{x})$ and $\bar{u}_2(\bar{x})$ for the first and the second clusters shown in Fig.16 and Fig.17, respectively, with tolerance vector $\bar{\varepsilon}$ for each grid point $\bar{x}$.

We can see that sFCM without tolerance ($\kappa_i = \tilde{\kappa} = 0$) and the one with tolerance ($\kappa_i = \tilde{\kappa} = 0.04$) classified some data into different clusters each other as shown in Fig.2 and Fig.12. Similarly, eFCM without tolerance ($\kappa_i = \tilde{\kappa} = 0$) and the one with tolerance

Fig.13 Contour Map of Fuzzy Classification Function $\bar{u}_1(\bar{x})$ for sFCM-T (Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0.04$

Fig.16 Contour Map of Fuzzy Classification Function $\bar{u}_1(\bar{x})$ for eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0.03$

Fig.14 Contour Map of Fuzzy Classification Function $\bar{u}_2(\bar{x})$ for sFCM-T (Algorithm 3) with $\kappa_i = \tilde{\kappa} = 0.04$

Fig.17 Contour Map of Fuzzy Classification Function $\bar{u}_2(\bar{x})$ for eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0.03$

Fig.15 Classification Results of eFCM-T (Algorithm 4) with $\kappa_i = \tilde{\kappa} = 0.03$

($\kappa_i = \tilde{\kappa} = 0.04$) classified some data into different clusters each other shown in Fig.7 and Fig.15. Such the difference can be seen without fuzzy classification functions.

Furthermore, we can observe how the classification for a brand-new datum changes by the uncertainty of the tolerance using fuzzy classification function values obtained by Algorithm 5 or Algorithm 6. (Compare Fig.13 with Fig.5, and Fig.14 with Fig.6, Fig.16 with Fig.10, and, Fig.17 with Fig.11.) Especially, we can see the two interesting phenomena as the following two paragraphs, respectively.

In Fig.13, we can find the region of $\bar{u}_1=1$ around v_1 , and the regions of $\bar{u}_0=0$ around v_2, v_3 and v_4 , while we cannot find in Fig.5. Similarly, in Fig.14, we can find the region of $\bar{u}_1=1$ around v_2 , and the regions of $\bar{u}_0=0$

around v_1 , v_3 and v_4 , while we cannot find in Fig.6. This phenomenon can be described as below. On each point $\bar{x}$ in the sphere with the center of each cluster center and with the radius $\tilde{\kappa}$, the corresponding tolerance vector $\bar{\varepsilon}$ heads to the center of the sphere, and $\bar{x} + \bar{\varepsilon}$ coincides with the cluster center. Remarking that fuzzy classification function value of sFCM without tolerance is 0 or 1 only on cluster centers and that the fuzzy classification function value (64) for sFCM-T at $\bar{x}$ can be interpreted as the one (16) for sFCM at $\bar{x} + \bar{\varepsilon}$, i.e., the contour map of fuzzy classification function for sFCM-T can be interpreted as moving the one for sFCM with $-\bar{\varepsilon}$, we can obtain the contour line on which the fuzzy classification function value is 0 or 1 as the spheres with the center of each cluster center and with the radius of $\tilde{\kappa}=0.04$ shown in Fig.13 and Fig.14.

In Fig.16 and Fig.17, we can find the distortion of the contour line around the midpoints of neighboring cluster center pairs $((v_1, v_2), (v_2, v_3)$ and $(v_2, v_4))$, compared with Fig.10 and Fig.11, respectively. In order to describe this phenomenon, Fig.18 is used as the model around $(v_1 + v_2)/2$ in Fig.17. In this figure, on perpendicular bisector of two cluster centers, the tolerance vector $\bar{\varepsilon}$ heads to the midpoint of two cluster centers. On the line segment connecting with two cluster centers, the tolerance vector $\bar{\varepsilon}$ heads to the closer cluster center. Considering these directions of tolerance vector together, the tolerance vectors around the midpoint of two cluster centers are described as the arrows in Fig.18. Remarking that the contour lines of fuzzy classification function for eFCM without tolerance intersect orthogonally with the line segment connecting two cluster centers as the dotted lines in Fig.18 and that the fuzzy classification function value (88) for eFCM-T at $\bar{x}$ can be interpreted as the one (18) for eFCM at $\bar{x} + \bar{\varepsilon}$, i.e., the contour map of fuzzy classification function for eFCM-T can be interpreted as moving the one for eFCM with $-\bar{\varepsilon}$, the contour lines of fuzzy classification function for eFCM-T are described as the solid lines in Fig.18. Hence, we can find that the distortion of the contour line around the midpoint of two cluster centers is produced by tolerance vectors. We can find the similar distortion around $(v_1 + v_2)/2$, $(v_2 + v_3)/2$ and $(v_2 + v_4)/2$ in Fig.17; and around $(v_1 + v_2)/2$ in Fig.16. We should note that the above description cannot always be applied with any cases of eFCM-T.

Fig.18 Model Describing that Contour Map of Fuzzy Classification Function for eFCM-T is not longer Convex

5. Conclusions

In this paper, we proposed two new fuzzy classification functions for sFCM-T and eFCM-T, respectively, which calculate the membership value for an arbitrary point in the data space and also calculate the tolerance for it. Because we cannot obtain these directly from the original optimal solutions of memberships, we proposed iterative algorithms respectively for that by solving certain optimizations, which are derived from that fuzzy classification function of FCM without tolerance can be interpreted as the solution of an optimization problem. Now we got the tool of fuzzy classification function in order to investigate the features of FCM-T essentially.

We showed some numerical examples of fuzzy classification functions for FCM-T. In sFCM-T, it was shown that fuzzy classification function generated the regions where the function values are 0 or 1 around cluster centers. In eFCM-T, it was described with the simplified model that the contour lines of fuzzy classification function are distorted around the midpoints of neighboring cluster center pairs.

As the further study, we will analyze FCM-T mathematically using the fuzzy classification function. Especially we will theoretically investigate the shape of contour map for fuzzy classification function of eFCM-T with any cases of data. We will also investigate the shape of cluster regions for FCM-T, i.e. whether cluster regions for FCM-T are Voronoi shaped as same as FCM without tolerance or not. The case that $\kappa_i \neq \kappa_k$ ($i \neq k$) will be also investigated.

References

- [1] Bezdek, J.C. : Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum, New York (1981).
- [2] Miyamoto, S. and Umayahara, K. : "Methods in Hard and Fuzzy Clustering", in : Liu, Z.-Q. and Miyamoto, S.(eds), Soft computing and humancentered machines, Springer-Verlag Tokyo, (2000).
- [3] Murata, R., Endo, Y., Haruyama, H. and Miyamoto, S. : "On fuzzy c -means for data with tolerance", J. Advanced Computational Intelligence and Intelligent Informatics, vol.10, no.5, pp.673-681, (2006).

(2008年12月15日 受付)

(2009年 5月22日 採録)

[Contact Address]

3-7-5, Toyosu, Koto-ku, Tokyo

Shibaura Institute of Technology

Yuchi Kanzawa

TEL : 03-5859-8253

E-mail : kanzawa@sic.shibaura-it.ac.jp

Information about Author

Yuchi KANZAWA [non-member]

Yuchi Kanzawa was born in Japan, on August 9, 1969. He received the B.E., M.E. and Dr. E. degrees in electronics and communication engineering, from Waseda university, Tokyo, Japan, in 1993, 1995 and 1998, respectively. From 1997 to 2000, he had been a research assistant of the department of information and computer science, Waseda university. In 2000, he joined the faculty of engineering, Shibaura institute of technology. Since then, as a faculty member, he has engaged in education and research in the fields of nonlinear science, applied mathematics and numerical analysis. Currently, he is an associate professor of the department of communication engineering of the institute of technology. He received the best paper award in 1998 from IEICE.

Yasunori ENDO [member]

Yasunori Endo was born in Japan, on January 26, 1967. He received the B.E., M.E. and D.E. degrees in electronics and communication engineering, from Waseda University, Tokyo, Japan, in 1990, 1992 and 1995, respectively. From 1994 to 1997, he had been a research assistant of the school of science and engineering, Waseda University. From 1997 to 2001, he had been an assistant professor of the school of engineering, Tokai University. Currently, he is an associate professor of the department of risk engineering, faculty of systems and information engineering, University of Tsukuba. His research interests are in the analysis of uncertainty by the methodology of soft computing including clustering or fuzzy set theory.

Sadaaki MIYAMOTO [member]

Sadaaki Miyamoto was born in Osaka, Japan, in 1950. He received the B.S., M.S. and the Dr. Eng. degrees in Applied Mathematics and Physics Engineering from Kyoto University, Japan, in 1973, 1975, and 1978, respectively. Prior to his professorship at the University of Tsukuba, he was an Assistant Professor from 1980 to 1987 and Associate Professor from 1987 to 1990 in the University of Tsukuba. He was a Professor with the Faculty of Engineering, the University of Tokushima, where he was working from 1990 to 1994. His current research interests include methodology for fuzzy systems, uncertainty modeling. In particular he has been working on data clustering algorithms and related classification methods, image indexing and retrieval, multisets, rough sets, metaheuristic optimization, and algorithms for data mining. He is a member of the Society of Instrumentation and Control, Engineers of Japan, Information Processing Society of Japan, the Japan Society of Fuzzy Theory and Systems, and the Institute of Systems, Control, Information Engineers of Japan, and IEEE. He has served a number of international conferences as chairs, co-chairs and committee members. He received excellent paper awards from the Japan Society of Fuzzy Theory and Systems in 1994 and 1999. Dr. Miyamoto has published 2 books of which one is in English and the other is in Japanese. He also has published 1 edited book and over 200 research papers.