Fuzzy c-Lines for Data with Tolerance

Yuchi Kanzawa[1] Yasunori Endo[2] Sadaaki Miyamoto[2]

1.Shibaura Institute of Technology, Japan
2.University of Tsukuba, Japan
Email: kanzawa@sic.shibaura-it.ac.jp

Abstract— *This paper presents a new clustering algorithm ,which is based on fuzzy c-lines, can treat data with some errors. First, the tolerance is formulated and introduce into optimization problem of clustering. Next, the problem is solved using Karush-Kuhn-Tucker conditions. Last, the algorithm is constructed based on the results of solving the problem. Some numerical examples for the proposed method are shown.*

Keywords— Data with Tolerance, Fuzzy c-Lines, Fuzzy Clustering

1 Introduction

Fuzzy c-means (FCM) [1] is one of the well-known fuzzy clustering and many FCM variants have been proposed after FCM. In these variants, an algorithm allowing cluster prototypes to be linear varieties instead of points has been proposed [1]. This algorithm is called fuzzy c-varieties (FCV) and regarded as the combination of clustering and principal component analysis. If linear varieties of the prototypes are limited to one dimensional, FCV is called fuzzy c-lines(FCL).

In general, any data that is represented by numeric have some errors. In order to classify such data, an algorithm has been proposed by some of the authors [2]. This algorithm is called the FCM for data with tolerance (FCM-T). In FCM-T, the tolerance is regarded as another decision variable composing the optimization problem like membership grades and the cluster centers, and is determined as minimizing the objective function.

In this paper, on the analogy of the story from FCM to FCM-T, we propose fuzzy c-lines for data with tolerance (FCL-T).

The contents of this paper are the followings. In the second section, we define some notations and introduce FCM-T and FCL. In the third section, we propose FCL-T. 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 fuzzy c-lines (FCL), and also introduce fuzzy c-means with tolerance (FCM-T) derived from conventional fuzzy c-means (FCM). In the first subsection, we define some notations which are the data for clustering, the membership by which the each data belongs to the each cluster, the cluster prototypes of one dimensional linear manifold. In the second subsection, we introduce FCL, whose prototypes are one dimensional linear manifolds. In the third subsection, we define the tolerance for the data and the maximum tolerance for the data, and introduce FCM-T, which is the story that clustering algorithms were oriented for data with some errors.

2.1 Notations

In this subsection, we define some notations which are the data for clustering, the membership by which the each data belongs to the each cluster, the cluster prototypes of one dimensional linear manifold.

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

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

The cluster prototype set is denoted by $V = \{V_1, \cdots, V_C\}$, where

$$V_j = \{y_j \in \mathbf{R}^p \mid y_j = v_j + t_j s_j, t_j \in \mathbf{R}\}$$

with $v = \{v_j \mid v_j \in \mathbf{R}^p, j \in \{1, \dots, C\}\}$ and $s = \{s_j \in \mathbf{R}^p \mid \|s_j\| = 1, j \in \{1, \cdots, C\}\}$.

2.2 Fuzzy c-Lines

In this subsection, we introduce FCL (FCV), whose prototypes are one dimensional linear manifolds.

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

$$\underset{u,v,s}{\text{minimize}}\ J_\ell(u, v, s) \tag{1}$$

$$\text{subject to} \sum_{j=1}^{C} u_{i,j} = 1, \quad \|s_j\| = 1, \tag{2}$$

where

$$J_\ell(u, v, s) = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{i,j}^m \left(\|x_i - v_j\|^2 - ((x_i - v_j)s_{j,\ell})^2\right). \tag{3}$$

The optimal solutions u, v and s are obtained by the following algorithm.

Algorithm 1 (FCL)

Step 1 *Give the number of clusters C and the fuzzifier parameter m. Set the initial values of v and s.*

Step 2 *Calculate u such that*

$$u_{i,j} = 1 / \sum_{k=1}^{C} \left(\frac{d_{i,j}}{d_{i,k}}\right)^{1/(m-1)}, \tag{4}$$

where

$$d_{i,j} = \|x_i - v_j\|^2 - ((x_i - v_j)s_{j,\ell})^2. \tag{5}$$

ISBN: 978-989-95079-6-8

Step 3 *Calculate v such that*

$$v_j = U_j^{-1} \sum_{i=1}^{N} u_{i,j}^m x_i, \tag{6}$$

where

$$U_j = \sum_{i=1}^{N} u_{i,j}^m. \tag{7}$$

Step 4 *Calculate s as the eigenvector corresponding to the maximal eigenvalue of the following matrix:*

$$\sum_{i=1}^{N} u_{i,j}^m (x_i - v_j)(x_i - v_j)^{\mathrm{T}}. \tag{8}$$

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

2.3 Fuzzy c-Means for Data with Tolerance

In this subsection, we define the tolerance for the data and the maximum tolerance for the data, and introduce fuzzy c-means for data with tolerance (FCM-T) [2]. This algorithm is oriented toward data with error. The tolerance for the data x is denoted by $\varepsilon = \{\varepsilon_i \mid \varepsilon_i \in \mathbf{R}^p, i \in \{1, \ldots, N\}\}$. The maximum tolerance is denoted by $\kappa = \{\kappa_i \mid \kappa_i \in \mathbf{R}_+, i \in \{1, \ldots, N\}\}$.

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

$$\underset{u,\varepsilon,v}{\text{minimize}}\; J_{\mathrm{t}}(u, \varepsilon, v)$$

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

where

$$J_{\mathrm{t}}(u, \varepsilon, v) = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{i,j}^m \|x_i + \varepsilon_i - v_j\|^2. \tag{9}$$

The optimal solutions u, ε and v are obtained by the following algorithm.

Algorithm 2 (FCM-T)

Step 1 *Give the number of clusters C, the fuzzifier parameter m and the maximum tolerance set κ for data x. Set the initial values of ε and v.*

Step 2 *Calculate u such that*

$$u_{i,j} = 1 / \sum_{k=1}^{C} \left(\frac{d_{i,j}}{d_{i,k}} \right)^{1/(m-1)}, \tag{10}$$

where

$$d_{i,j} = \|x_i + \varepsilon_i - v_j\|^2. \tag{11}$$

Step 3 *Calculate ε such that*

$$\varepsilon_i = -\alpha_i \left(x_i - \sum_{j=1}^{C} u_{i,j}^m v_j \right), \tag{12}$$

where

$$\alpha_i = \min \left\{ \kappa_i \left\| x_i - \sum_{j=1}^{C} u_{i,j}^m v_j \right\|^{-1}, \left(\sum_{j=1}^{C} u_{i,j}^m \right)^{-1} \right\}. \tag{13}$$

Step 4 *Calculate v such that*

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

where

$$U_j = \sum_{i=1}^{N} u_{i,j}^m. \tag{15}$$

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

If the maximal tolerance κ_i is set to zero, this algorithm is coincides with FCM without tolerance. This algorithm is executing FCM without tolerance for data moving within a region as shown in Fig. 1.

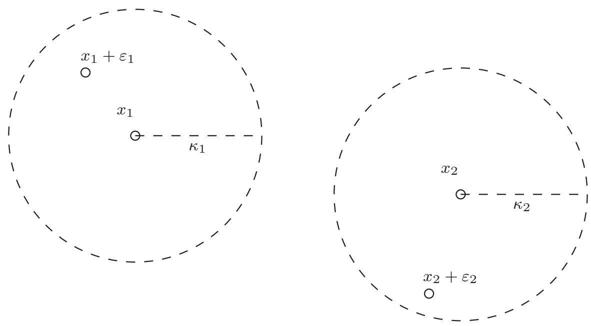

Figure 1: Data with Tolerance

3 Fuzzy c-Lines for Data with Tolerance

In this section, fuzzy c-lines for data with tolerance (FCL-T) is proposed. First, the tolerance is formulated and introduce into optimization problem of FCL. Next, the problem is solved using Karush-Kuhn-Tucker conditions. Last, the algorithm is constructed based on the results of solving the problem.

3.1 Optimization Problem and Its KKT conditions

In this subsection, a new optimization problem is proposed such that the tolerance is formulated and introduce into optimization problem of FCL, and its Karush-Kuhn-Tucker conditions are considered.

Let us consider the following optimization problem:

$$\underset{u,\varepsilon,v,s}{\text{minimize}}\; J_{\ell,\mathrm{t}}(u, \varepsilon, v, s)$$

$$\text{subject to} \quad \begin{cases} \sum_{j=1}^{C} u_{i,j} = 1, \\ \|s_j\| = 1, \\ \|\varepsilon_i\|^2 \le \kappa_i^2, \end{cases} \tag{16}$$

ISBN: 978-989-95079-6-8

where

$$J_{\ell,\mathrm{t}}(u,\varepsilon,v,s) = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{i,j}^m \left(\|x_i+\varepsilon_i-v_j\|^2 - ((x_i+\varepsilon_i-v_j)s_j)^2\right). \tag{17}$$

Lagrange function of this problem $L_{\ell,\mathrm{t}}(u,\varepsilon,v,s)$ is described as below:

$$\begin{aligned} L_{\ell,\mathrm{t}}(u,\varepsilon,v,s) =& J_{\mathrm{s},\ell,\mathrm{t}}(u,\varepsilon,v,s) + \sum_{i=1}^{N}\gamma_i\left(\sum_{j=1}u_{i,j}-1\right) \\ &+\sum_{j=1}^{C}\xi_j\left(s_j^{\mathrm{T}}s_j-1\right) \\ &+\sum_{i=1}^{N}\delta_i\left(\|\varepsilon_i\|^2-\kappa_i^2\right). \end{aligned} \tag{18}$$

where $\gamma = (\gamma_1,\cdots,\gamma_N)$, $\xi = (\xi_1,\cdots,\xi_C)$ and $\delta = (\delta_1,\cdots,\delta_N)$ are Karush-Kuhn-Tucker vectors. Karush-Kuhn-Tucker (KKT) conditions are described as below:

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial u_{i,j}} =0, \tag{19}$$

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial v_j} =0, \tag{20}$$

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial s_j} =0, \tag{21}$$

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial \varepsilon_i} =0, \tag{22}$$

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial \gamma_i} =0, \tag{23}$$

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial \xi_j} =0, \tag{24}$$

$$\frac{\partial L_{\ell,\mathrm{t}}}{\partial \delta_i} =0, \tag{25}$$

$$\delta_i\frac{\partial L_{\ell,\mathrm{t}}}{\partial \delta_i} =0, \tag{26}$$

$$\delta_i \geq 0. \tag{27}$$

3.2 Optimal Solution of $u_{i,j}$, v_j and s_j

In this subsection, we derive the optimal solution of $u_{i,j}$, v_j and s_j from the above KKT conditions, and we show that these solutions are just substituting $x_i+\varepsilon_i$ to x_i in FCL (Algorithm 1).

From KKT conditions (19) and (23), we have the optimal solution of $u_{i,j}$ as

$$u_{i,j} = 1/\sum_{k=1}\left(\frac{d_{i,j}}{d_{i,k}}\right)^{1/(m-1)}, \tag{28}$$

where

$$d_{i,j} = \|x_i+\varepsilon-v_j\|^2 - (s_j^{\mathrm{T}}(x_i+\varepsilon_i-v_j))^2. \tag{29}$$

From KKT conditions (20), we have

$$(E-s_js_j^{\mathrm{T}})\sum_{i=1}^{N}u_{i,j}^m(x_i+\varepsilon_i-v_j)=0, \tag{30}$$

where E is the p-dimensional unit matrix. The null space of the matrix $E-s_js_j^{\mathrm{T}}$ include s_j since

$$(E-s_js_j^{\mathrm{T}})s_j = s_j-(s_j^{\mathrm{T}}s_j)s_j=0. \tag{31}$$

For the p-dimensional vector w such as $w^{\mathrm{T}}s_j = 0$ and as $w^{\mathrm{T}}w=1$ and for a scalar value α, we have

$$\begin{aligned}(E-s_js_j^{\mathrm{T}})(\alpha w) =& \alpha(w-(s^{\mathrm{T}}w)s_j) \\ =&\alpha w, \end{aligned} \tag{32}$$

which is equal to 0 if and only if $\alpha = 0$, hence, the null space of $E-s_js_j^{\mathrm{T}}$ is just s_j. Therefore, Eq.(30) indicates that

$$\sum_{i=1}^{N}u_{i,j}^m(x_i+\varepsilon_i-v_j)=as_j \tag{33}$$

for an arbitrary value of a. From the above, we have the optimal solution of v_j as

$$v_j = \frac{\sum_{i=1}^{N}u_{i,j}^m(x_i+\varepsilon_i)-as_j}{\sum_{i=1}^{N}u_{i,j}^m}, \tag{34}$$

in which we may set $a=0$ and have

$$v_j = \frac{\sum_{i=1}^{N}u_{i,j}^m(x_i+\varepsilon_i)}{\sum_{i=1}^{N}u_{i,j}^m}. \tag{35}$$

From KKT conditions (21) and (24), we have

$$\sum_{i=1}^{N}u_{i,j}^m(x_i+\varepsilon_i-v_j)(x_i+\varepsilon_i-v_j)^{\mathrm{T}}s_j = -\xi_js_j, \tag{36}$$

from which we can obtain the optimal solution of s_j as the eigenvector for the maximal eigenvalue of the matrix

$$\sum_{i=1}^{N}u_{i,j}^m(x_i+\varepsilon_i-v_j)(x_i+\varepsilon_i-v_j)^{\mathrm{T}}. \tag{37}$$

Comparing between Eq. (28) and (4), Eq. (29) and (5), Eq. (35) and (6), and Eq. (37) and (8), we can see that optimal solutions in FCL-T is obtained just by substituting $x_i+\varepsilon_i$ to x_i in FCL (Algorithm 1).

3.3 Optimal Solution of ε_i

In this subsection, we obtain the optimal solution of ε_i from the KKT conditions.

From KKT conditions (22), (25), (26) and (27), we have

$$\begin{aligned}&\left(\delta_iE+\sum_{j=1}^{C}u_{i,j}^m(E-s_js_j^{\mathrm{T}})\right)\varepsilon_i+ \\ &\sum_{j=1}^{C}u_{i,j}^m(E-s_js_j^{\mathrm{T}})(x_i-v_j)=0, \end{aligned} \tag{38}$$

$$\delta_i(\|\varepsilon_i\|^2-\kappa_i^2)=0. \tag{39}$$

ISBN: 978-989-95079-6-8

In order to solve this equation, we discuss in each case that the matrix $\left(\sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})\right)$ is invertible or not as the followings.

The matrix $\left(\sum_{j=1}^{C} u_{i,j}^m \left(E - s_j s_j^{\mathrm{T}}\right)\right)$ is not invertible if $s_j = s_{\tilde{j}}$ for any $j \neq \tilde{j}$ or if there exists j^* such as $u_{i,j^*} = 1$, though the proof is omitted because of the sake of the pages. Here, we write $s = s_j$ in the case of $s_j = s_{\tilde{j}}$ for any $j \neq \tilde{j}$ and also write $s = s_{j^*}$ in the case that there exists j^* such as $u_{i,j^*} = 1$. In these cases, Eq. (38) is rewritten by setting $\varepsilon_i = \alpha s + \beta w$ with certain values α and β, and a p-dimensional vector w such as $s^{\mathrm{T}} w = 0$ and as $w^{\mathrm{T}} w = 1$, as

$$\beta(\sum_{j=1}^{C} u_{i,j}^m) w + \delta_i(\alpha s + \beta w) + (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j) = 0. \tag{40}$$

Multiplying s^{T} to Eq.(40), we have $\delta_i \alpha = 0$, which indicates $\delta_i = 0$ or $\alpha = 0$. $\alpha = 0$ means that ε_i is orthogonal to s. In the case of $\delta_i = 0$, we have

$$(E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i + \varepsilon_i - v_j) = 0. \tag{41}$$

which implies

$$\sum_{j=1}^{C} u_{i,j}^m (x_i + \varepsilon_i - v_j) = as \tag{42}$$

with an arbitrary value a. Hence, we have the optimal solution of ε_i in the case that $\left(\sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})\right)$ is not invertible with $\delta_i = 0$, as

$$\varepsilon_i = \frac{as - \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j)}{\sum_{j=1}^{C} u_{i,j}}. \tag{43}$$

If we set a such that ε_i is orthogonal to s, we can discuss both the case of $\delta_i \alpha = 0$ and the base of $\alpha = 0$ as the followings. Since $s^{\mathrm{T}} \varepsilon_i = 0$, Eq.(38) is rewritten as

$$\left(\sum_{j=1}^{C} u_{i,j}^m\right) \varepsilon_i + \delta_i \varepsilon_i + (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j) = 0, \tag{44}$$

from which we have

$$\varepsilon_i = -\frac{(E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j)}{\sum_{j=1}^{C} u_{i,j}^m + \delta_i}. \tag{45}$$

If $\delta_i = 0$, we have

$$\varepsilon_i = -\frac{(E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j)}{\sum_{j=1}^{C} u_{i,j}^m}. \tag{46}$$

If $\delta_i \neq 0$, then $\|\varepsilon_i\| = \kappa_i^2$ from Eq. (39), by which we have

$$\|\varepsilon_i\|^2 = \left\| \frac{(E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j)}{\sum_{j=1}^{C} u_{i,j}^m + \delta_i} \right\|^2 = \kappa_i^2. \tag{47}$$

By solving Eq.(47) as

$$\sum_{j=1}^{C} u_{i,j}^m + \delta_i = \frac{\left\| (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j) \right\|}{\kappa_i}, \tag{48}$$

we have

$$\varepsilon_i = -\kappa_i \left\| (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j) \right\|^{-1} \cdot (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j). \tag{49}$$

Getting Eq.(46) and (49) together, we have the optimal solution of ε_i in the case that the matrix $\left(\sum_{j=1}^{C} u_{i,j}^m \left(E - s_j s_j^{\mathrm{T}}\right)\right)$ is not invertible

$$\varepsilon_i = -\min \left\{ \kappa_i \left\| (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j) \right\|^{-1}, \left(\sum_{j=1}^{C} u_{i,j}^m\right)^{-1} \right\} (E - ss^{\mathrm{T}}) \sum_{j=1}^{C} u_{i,j}^m (x_i - v_j). \tag{50}$$

In the case that $\left(\sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})\right)$ is invertible, we cannot obtain the explicit form of the optimal solution ε_i, so, we apply some numerical method for Eq. (38) and (39) as the followings. First, we set $\delta_i = 0$ and solve Eq. (38) as

$$\varepsilon_i = -\left(\sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})\right)^{-1} \cdot \sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})(x_i - v_j). \tag{51}$$

If the obtained value of ε_i satisfies $\|\varepsilon_i\|^2 \leq \kappa_i^2$, we adopt this value as the optimal solution of ε_i. Otherwise, we solve Eq. (38) with $\|\varepsilon_i\|^2 = \kappa_i^2$ numerically.

ISBN: 978-989-95079-6-8

3.4 Algorithm

In this subsection, we propose an iterative algorithm of fuzzy c-lines for data with tolerance from the above discussion.

Algorithm 3 (FCL-T)

Step 1 *Give the values of the fuzzifier parameter m and the maximal tolerance set κ. Set the initial cluster center v, s and ε.*

Step 2 *Calculate $u_{i,j}$ such that*

$$u_{i,j} = 1/\sum_{k=1}^{C}\left(\frac{d_{i,j}}{d_{i,k}}\right)^{1/(m-1)}, \tag{52}$$

where

$$d_{i,j} = \|x_i + \varepsilon - v_j\|^2 - (s_j^{\mathrm{T}}(x_i + \varepsilon_i - v_j))^2. \tag{53}$$

Step 3 *Calculate v_j such that*

$$v_j = \frac{\sum_{i=1}^{N} u_{i,j}^m (x_i + \varepsilon_i)}{\sum_{i=1}^{N} u_{i,j}^m}. \tag{54}$$

Step 4 *Obtain s_j as the eigenvector with the maximal eigenvalue of the matrix*

$$\sum_{i=1}^{N} u_{i,j}^m (x_i + \varepsilon_i - v_j)(x_i + \varepsilon_i - v_j)^{\mathrm{T}}. \tag{55}$$

Step 5 *If $s_j = s_{\tilde{j}}$ for any $j \neq \tilde{j}$ or if there exists j such that $u_{i,j} = 1$, obtain ε_i as*

$$\varepsilon_i = -\min\left\{\kappa_i \left\|(E - ss^{\mathrm{T}})\sum_{j=1}^{C} u_{i,j}^m (x_i - v_j)\right\|^{-1}, \left(\sum_{j=1}^{C} u_{i,j}^m\right)^{-1}\right\}(E - ss^{\mathrm{T}})\sum_{j=1}^{C} u_{i,j}^m (x_i - v_j), \tag{56}$$

where

$$s = \begin{cases} s_j & \text{for} s_j = s_{\tilde{j}} \quad (j \neq \tilde{j}), \\ s_{j^*} & \text{for } u_{i,j^*} = 1, \end{cases} \tag{57}$$

and go to Step 7. *Otherwise, go to* Step 6.

Step 6 *Calculate*

$$\varepsilon_i = -\left(\sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})\right)^{-1} \cdot \sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})(x_i - v_j). \tag{58}$$

If $\|\varepsilon_i\|^2 \leq \kappa_i^2$ holds, adopt this ε_i and go to Step 7. *Otherwise, solve the equation*

$$\left(\delta_i E + \sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})\right)\varepsilon_i + \sum_{j=1}^{C} u_{i,j}^m (E - s_j s_j^{\mathrm{T}})(x_i - v_j) = 0, \tag{59}$$

$$\|\varepsilon_i\|^2 = \kappa_i^2 \tag{60}$$

Step 7 *Check the stopping criterion. If the criterion is not satisfied, go back to* Step 2.

If the maximal tolerance κ_i is set to zero, this algorithm is coincides with FCL without tolerance (Algorithm 1). Remark that Eq. (56) in Algorithm 3 and Eq. (13) for obtaining the tolerance in Algorithm 2 are essentially the same with each other, because Eq. (56) means that the tolerance is obtaining in the orthogonal complementary space of s in the pattern space. Remark also that there is the quite similarity between Eq. (58) in Algorithm 3 and Eq. (9.29) in [3] for dealing with the missing data in linear fuzzy clustering followed by the optimal completion strategy [4], because Eq. (58) means that the optimal solution of the tolerance in a certain condition is not affected by the constraint of the tolerance $\|\varepsilon_i\| \leq \kappa_i^2$, while Eq. (9.29) in [3] does not consider the constraint for missing component of data from the beginning.

4 Numerical Example

In this section, we show some examples of classification by fuzzy c-lines for data with tolerance (Algorithm 3). In each example, 100 trials for Algorithm 3 with different initial cluster centers are tested and the solution with the minimal objective function value is selected as the final result. The classified dataset is Gustaffson's cross [5] shown in Fig. 2. This dataset consists of 20 points ($N = 20$) in $\mathbf{R}^2$ and forms two visually apparent linear clusters with the crossing shape.

We fix $m = 2$ and test four different values of $\kappa_i \in \{0, 0.016, 0.032, 0.064\}$. All cases of κ_i produce the correctly classified results shown in Fig. 3, 4 and 5, respectively. Especially, the cases of $\kappa_i \in \{0.032, 0.064\}$ produce the same results shown in Fig. 5. From Fig. 4, we can see that each data moves to the direction of its own prototype. From Fig. 5, we can see that each data moves just on its own prototype.

From these results, we consider the following hypothesis: if Gustaffson's cross is generated by adding error followed by uniform distribution with a width after sampling from the two crossing exactly linear shaped population, FCL-T (Algorithm 3) produces the shape of the original population by $v_j + ts_j$ and also produces the width of the uniform distribution by κ_i, though we have not proved it yet.

5 Conclusion

In this paper, we proposed fuzzy c-lines for data with tolerance. The proposed method is obtained by introducing the idea of tolerance, which means that each datum moves within a region, into fuzzy c-lines. This is the analogy with that fuzzy c-means for data with tolerance is obtained by introducing the idea of tolerance into fuzzy c-means. In the numerical examples, we can see that each data moves to the direction of its own prototype in setting the positive maximal tolerance, and that each data moves on its own prototype in setting larger maximal tolerance.

ISBN: 978-989-95079-6-8

From this feature of the proposed method, we consider the following hypothesis: if dataset is generated by adding error followed by uniform distribution with a width after sampling from the exactly linear shaped population, FCL-T (Algorithm 3) produces the shape of the original population and also produces the width of the uniform distribution. But we must investigate the proposed method theoretically and through many numerical experiments, which is the first future work.

As another future works, (1) entropy regularlized fuzzy c-means [6] for data with tolerance will be proposed by the similar discussion, (2) the cases of error followed by other probabilistic distribution will be considered, e.g., the case of error followed by normal distribution can be considered by adding the penalty term of squared tolerance, and Laplace one by the penalty term of absolute value of tolerance.

References

[1] Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms, Plenun, New York, (1981).

[2] Endo, Y., Miyamoto, S.: "Fuzzy c-Means for Data with Tolerance", Proc. 2005 International Symposium on Nonlinear Theory and Its Applications, pp.123–127, 2005.

[3] Miyamoto, S., Ichihashi, H., Honda, K.: "Algorithms for Fuzzy Clustering", Springer, Berlin (2008).

[4] Hathaway, R.J., Bezdek, J.C.: "Fuzzy c-means Clustering of Incomplete Data", IEEE Trans. on Systems, Man and Cybernetics Part B 31(5), pp.735–744 (2001).

[5] Gustaffson, D.E., Kessel, W.: "Fuzzy Clustering with a Fuzzy Covariance Matrix", Proc. IEEE–CDC, Vol.2, pp.761–766 (1979).

[6] Miyamoto, S. and Mukaidono, M.: "Fuzzy c-Means as a Regularization and Maximum Entropy Approach", Proc. of the 7th International Fuzzy Systems Association World Congress (IFES'97), Vol. II, pp. 86–92, Prague, 1997.

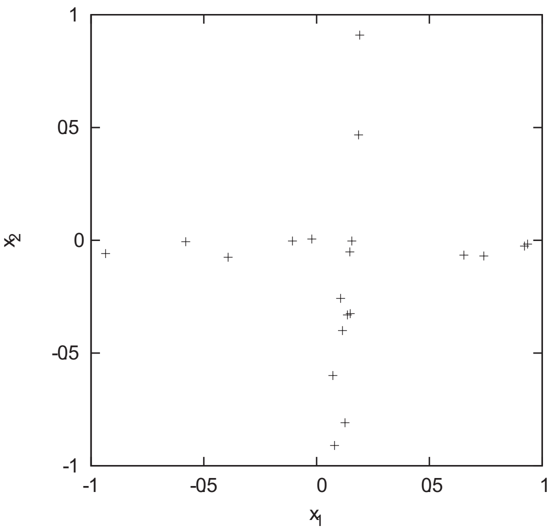

Figure 2: Gustaffson's Cross

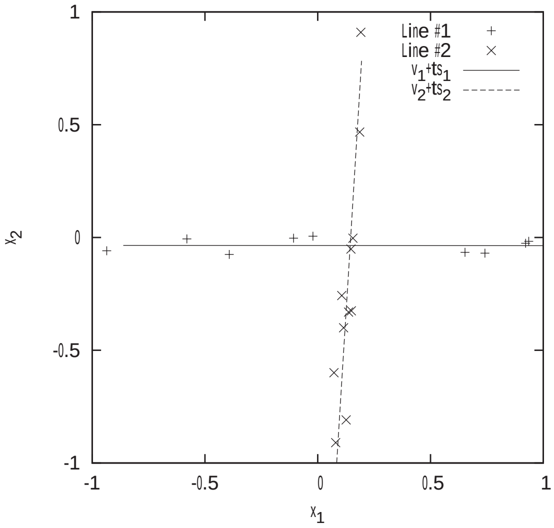

Figure 3: Successful Classification Result of Fig. 2 by FCL-T 3 with $m = 2$ and $\kappa_i = 0$

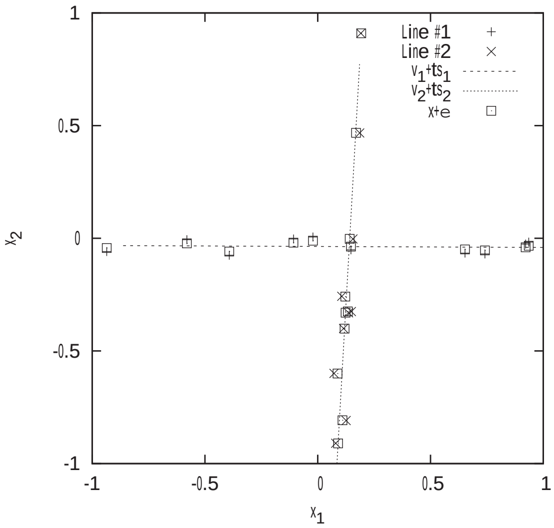

Figure 4: Successful Classification Result of Fig. 2 by FCL-T 3 with $m = 2$ and $\kappa_i = 0.016$

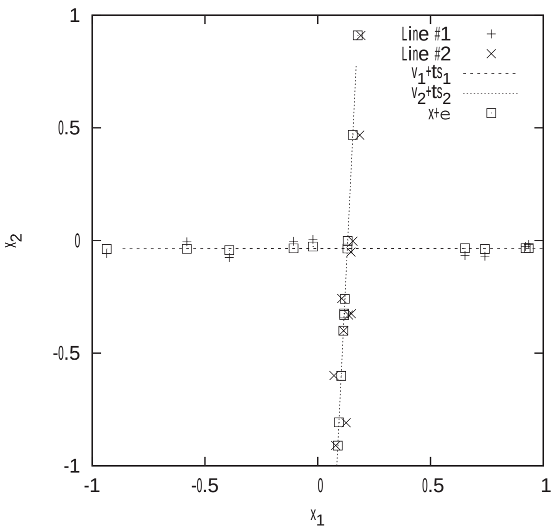

Figure 5: Successful Classification Result of Fig. 2 by FCL-T 3 with $m = 2$ and $\kappa_i = 0.032$

ISBN: 978-989-95079-6-8