Paper:

Fuzzy *c*-Means Clustering for Uncertain Data Using Quadratic Penalty-Vector Regularization

Yasunori Endo*, **Yasushi Hasegawa****, **Yukihiro, Hamasuna***, **and Yuchi Kanzawa*****

*Department of Risk Engineering, Faculty of Systems and Information Engineering, University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan
Email: endo@risk.tsukuba.ac.jp
**DENSO Co., Ltd.
1-1 Showa-cho, Kariya, Aichi 448-8661, Japan
Email: yasushi_hasegawa@denso.co.jp
***Faculty of Engineering, Shibaura Institute of Technology
3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan
Email: kanzawa@sic.shibaura-it.ac.jp
[Received February 13, 2010; accepted April 22, 2010]

Clustering – defined as an unsupervised data-analysis classification transforming real-space information into data in pattern space and analyzing it – may require that data be represented by a set, rather than points, due to data uncertainty, e.g., measurement error margin, data regarded as one point, or missing values. These data uncertainties have been represented as interval ranges for which many clustering algorithms are constructed, but the lack of guidelines in selecting available distances in individual cases has made selection difficult and raised the need for ways to calculate dissimilarity between uncertain data without introducing a nearest-neighbor or other distance. The tolerance concept we propose represents uncertain data as a point with a tolerance vector, not as an interval, while this is convenient for handling uncertain data, tolerance-vector constraints make mathematical development difficult. We attempt to remove the tolerance-vector constraints using quadratic penalty-vector regularization similar to the tolerance vector. We also propose clustering algorithms for uncertain data considering optimization and obtaining an optimal solution to handle uncertainty appropriately.

Keywords: clustering, fuzzy *c*-means, uncertain data, optimization, penalty vector

1. Introduction

The data from numerous natural and social phenomena accumulated into humongous computer databases is clearly too massive and complex to analyze manually. In roughly supervised and unsupervised data-analysis classification, data clustering is the most widely used unsupervised classification, dividing sets of objects into categories called clusters. Clustering is classified into hierarchical and non-hierarchical methods.

Hard c-Means clustering [1] (HCM), the most widely used non-hierarchical approach, classifies a set of objects simply into c clusters using deterministic assignments, e.g., an object x is assigned exclusively to one cluster. There are many HCM extensions including Fuzzy c-Means clustering [2] (FCM), which extends HCM using fuzzy sets, although not requiring the detailed fuzzy set theory.

Data in real space generally transformed into a point in pattern space and analyzed in clustering often must be represented by a set instead of a point due to data uncertainty. Take the following three examples:

- **Errors introduced when the object is mapped from real space to pattern space**

  In this general case, rounding error is introduced when data is mapped from real space to pattern space.

- **Uncertainty introduced in real space**

  In measuring the scale of an astronomical object, the scale estimated has many errors due to the uncertainty such as measurement-gauge fluctuation or a lack of knowledge about the object.

- **Original uncertainty with data**

  In mapping an object from real space to three-dimensional (3D) RGB color space, the object represented as one point in pattern space may actually have colors making it more logical to represent it as an interval, not a single point, in pattern space.

These data uncertainties have been represented as interval ranges for which numerous clustering algorithms have been constructed [3–5] using distances between intervals, e.g., nearest neighbor, furthest neighbor, or Hausdorff distance, to calculate dissimilarity in target data. A lack of guidelines for selecting the distance between intervals has complicated selection.

The tolerance concept we introduce here includes data uncertainties for which we propose clustering algorithms [6, 7]. The basic tolerance concept assumes that data $x \in \Re^p$ with uncertainty is represented as interval $[\underline{\kappa}, \overline{\kappa}] = [(\underline{\kappa}_1, \ldots, \underline{\kappa}_p)^T, (\overline{\kappa}_1, \ldots, \overline{\kappa}_p)^T] \subset \Re^p$. This is represented in one of two ways as $[x + \underline{\kappa}, x + \overline{\kappa}]$ and as $x + (\overline{\kappa} + \underline{\kappa})/2 + \varepsilon$ $(\varepsilon = (\varepsilon_1, \ldots, \varepsilon_p)^T)$ under constraints $|\varepsilon_j| \leq (\overline{\kappa}_j - \underline{\kappa}_j)/2$ $(j = 1, \ldots, p)$. This second way corresponds to the basic tolerance concept, and ε is called the tolerance vector. Although convenient in handling uncertain data, tolerance vector constraints make mathematical development difficult.

To remove constraint, we use quadratic regularization in which the tolerance vector is called a penalty vector because the penalty vector role differs from tolerance vectors in its absence of constraint. After considering optimization problems applying a quadratic regularization term of penalty vectors instead of norms of tolerance constraints, we construct clustering algorithms based on standard Fuzzy c-Means (sFCM) and entropy-regularized Fuzzy c-Means (eFCM) on a Euclidean (L_2) norm for data with tolerance while solving the optimization problems.

2. Penalty Vector Concept

As stated regarding point vs. set representation, we have handled the uncertainty range as tolerance and define the tolerance vector within tolerance [6, 7].

To define terms, $X = \{x_1, \ldots, x_n\}$ is a subset on p-dimensional vector space $\Re^p$ written $x_k = (x_{k1}, \ldots, x_{kp})^T \in \Re^p$. Dataset X is classified into clusters C_i $(i = 1, \ldots, c)$. Let $v_i = (v_{i1}, \ldots, v_{ip})^T \in \Re^p$ and $v_i \in V$. v_i is cluster center C_i. $u_{ki} \in [0, 1]$ is a membership grade of x_k belonging to C_i. A partition matrix is $U = [u_{ki}]$. $\varepsilon_k \in E$ is tolerance vector under constraint:

$$|\varepsilon_{kj}| \leq \kappa_{kj}, \quad (\kappa_{kj} \geq 0) \quad \ldots\ldots\ldots \quad (1)$$

κ_{kj} is the maximum tolerance range, meaning the width of data uncertainty. Uncertain data is represented as $x_k + \varepsilon_k$ with constraint Eq. (1). The basic tolerance concept although well known, is new in that it is introduced into objective functions to construct clustering algorithms. We have handled data uncertainty as tolerance and proposed new clustering algorithms elsewhere [6, 7], but those algorithms could not handle uncertain data with an unknown range, and uncertainty is not given in many cases. Algorithms may also fail to consider data distribution, so we introduce the penalty vector, which is denoted as $\delta_k = (\delta_{k1}, \ldots, \delta_{kp})^T \in \Re^p$ and a set of penalty vectors as $\Delta = \{\delta_1, \ldots, \delta_n\}$. In conventional tolerance work, uncertain data is represented as $x_k + \varepsilon_k$ with constraint Eq. (1). Here we represent data as $x_k + \delta_k$ with no constraint – the most difficult point from tolerance. **Fig. 1** is an example of the penalty vector in $\Re^2$.

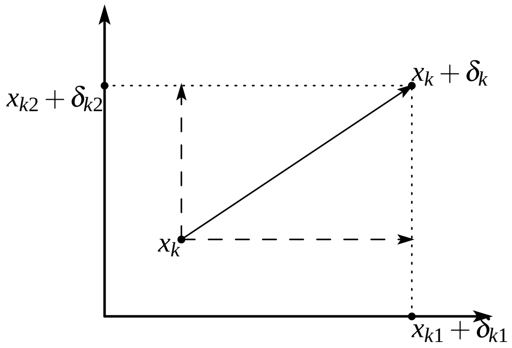


Fig. 1. An example of penalty vector in $\Re^2$.

3. Objective Functions and Optimal Solutions

In discussing a basic theory to construct new fuzzy c-means clustering algorithms for uncertain data using quadratic regularization, we define two objective functions and derive optimal solutions minimizing the functions using a Lagrange multiplier.

3.1. sFCM for Uncertain Data Using Quadratic Regularization

The objective function and constraints of sFCM are as follows:

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ki})^m \|x_k - v_i\|^2, \quad \ldots\ldots \quad (2)$$

$$\sum_{i=1}^{c} u_{ki} = 1. \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

We introduce the following quadratic regularization term:

$$\sum_{k=1}^{n} \delta_k{}^T W_k \delta_k = \sum_{j=1}^{p} \sum_{l=1}^{p} w_{klj} \delta_{kl} \delta_{kj}. \quad \ldots\ldots \quad (4)$$

where

$$W_k = \begin{pmatrix} w_{k11} & \cdots & w_{k1p} \\ \vdots & \ddots & \vdots \\ w_{kp1} & \cdots & w_{kpp} \end{pmatrix}$$

is a penalty matrix and $w_{kj} (w_{kj} \geq 0)$ is a penalty coefficient. W_k is assumed to be a symmetrical matrix, i.e., $w_{kj} = w_{jk}$. When W_k is diagonal, the quadratic regularization term is as follows:

$$\sum_{k=1}^{n} \delta_k{}^T W_k \delta_k = \sum_{k=1}^{n} \sum_{j=1}^{p} w_{kjj} \left(\delta_{kj}\right)^2.$$

We add term Eq. (4) to the objective function of sFCM Eq. (2) to obtain the following objective function:

$$J_s(U,V,\Delta) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ki})^m \|x_k + \delta_k - v_i\|^2 + \sum_{k=1}^{n} \delta_k{}^T W_k \delta_k. \quad \ldots\ldots\ldots \quad (5)$$

The nearer it approaches $(x_k + \delta_k)$ by v_i, the smaller the first term of Eq. (5) becomes. The penalty term of Eq. (5), however, grows proportionally to squared δ_{kj}, so the bigger the w_{kj}, the smaller the optimal solution δ_{kj}, which minimizes Eq. (5). The smaller the w_{kj}, the larger the optimal solution δ_{kj}, which minimizes Eq. (5). W_k thus means the uncertainty of data x_k.

Our goal here is to derive optimal solutions U, V, and Δ, which minimize objective function Eq. (5) under the following constraint:

$$\sum_{i=1}^{c} u_{ki} = 1. \quad \ldots \quad (6)$$

We first derive optimal solutions of U, introducing the following Lagrange function:

$$L_s = J_s(U,V,\Delta) + \sum_{k=1}^{n} \gamma_k (\sum_{i=1}^{c} u_{ki} - 1).$$

From

$$\frac{\partial L_s}{\partial u_{ki}} = m(u_{ki})^{m-1} \|x_k + \delta_k - v_i\|^2 + \gamma_k = 0, \quad \ldots \quad (7)$$

we have

$$u_{ki} = \left(\frac{-\gamma_k}{m\|x_k + \delta_k - v_i\|^2} \right)^{\frac{1}{m-1}}.$$

From constraint Eq. (6), we have

$$\sum_{l=1}^{c} \left(\frac{-\gamma_k}{m\|x_k + \delta_k - v_l\|^2} \right)^{\frac{1}{m-1}} = 1. \quad \ldots \quad (8)$$

From Eqs. (7) and (8), we have

$$u_{ki} = \frac{s_{ki}}{\sum_{l=1}^{c} s_{kl}},$$

where

$$s_{ki} = \|x_k + \delta_k - v_i\|^{-\frac{2}{m-1}}.$$

We next consider how to derive optimal solutions of V. From

$$\frac{\partial L_s}{\partial v_{ij}} = -\sum_{k=1}^{n} 2(u_{ki})^m (x_{kj} + \delta_{kj} - v_{ij}) = 0,$$

we have

$$v_{ij} = \frac{\sum_{k=1}^{n} (u_{ki})^m (x_{kj} + \delta_{kj})}{\sum_{k=1}^{n} (u_{ki})^m}.$$

We last consider how to derive Δ. From

$$\begin{aligned}\frac{\partial L_s}{\partial \delta_{kj}} &= \frac{\partial}{\partial \delta_{kj}} \left(\sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{p} (u_{ki})^m (x_{kj} + \delta_{kj} - v_{ij})^2 \right. \\ &\qquad \left. + \sum_{l=1}^{p} \sum_{j=1}^{p} w_{klj} \delta_{kl} \delta_{kj} \right) \\ &= 2\sum_{i=1}^{c} (u_{ki})^m (x_{kj} + \delta_{kj} - v_{ij}) + 2\sum_{l=1}^{p} w_{klj} \delta_{kl} \\ &= 0,\end{aligned}$$

we have the following:

$$\left(\sum_{i=1}^{c} (u_{ki})^m \right) \delta_{kj} + \sum_{l=1}^{p} w_{klj} \delta_{kl} + \sum_{i=1}^{c} (u_{ki})^m (x_{kj} - v_{ij}) = 0.$$

The above equation holds for any j $(1 \le j \le p)$. Hence,

$$\begin{pmatrix} \sum_{i=1}^{c} (u_{ki})^m & & 0 \\ & \ddots & \\ 0 & & \sum_{i=1}^{c} (u_{ki})^m \end{pmatrix} \begin{pmatrix} \delta_{k1} \\ \vdots \\ \delta_{kp} \end{pmatrix} + \begin{pmatrix} w_{k11} & \cdots & w_{kp1} \\ \vdots & \ddots & \vdots \\ w_{k1p} & \cdots & w_{kpp} \end{pmatrix} \begin{pmatrix} \delta_{k1} \\ \vdots \\ \delta_{kp} \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^{c} (u_{ki})^m (x_{k1} - v_{i1}) \\ \vdots \\ \sum_{i=1}^{c} (u_{ki})^m (x_{kp} - v_{ip}) \end{pmatrix} = 0. \quad \ldots \quad (9)$$

We use the following symbols:

$$\begin{aligned} A_k &= \begin{pmatrix} \sum_{i=1}^{c} (u_{ki})^m & & 0 \\ & \ddots & \\ 0 & & \sum_{i=1}^{c} (u_{ki})^m \end{pmatrix} + \begin{pmatrix} w_{k11} & \cdots & w_{kp1} \\ \vdots & \ddots & \vdots \\ w_{k1p} & \cdots & w_{kpp} \end{pmatrix} \\ &= \left(\sum_{i=1}^{c} (u_{ki})^m \right) I + W_k^T, \\ B_k &= \begin{pmatrix} \sum_{i=1}^{c} (u_{ki})^m (x_{k1} - v_{i1}) \\ \vdots \\ \sum_{i=1}^{c} (u_{ki})^m (x_{kp} - v_{ip}) \end{pmatrix}. \end{aligned}$$

I is a unit matrix.

Using symbols A_k and B_k, Eq. (9) is rewritten as follows:

$$A_k \delta_k + B_k = 0.$$

We then get the following solution:

$$\delta_k = -(A_k)^{-1} B_k.$$

We require that A_k be regularized for optimal solutions Δ.

3.2. eFCM for Uncertain Data Using Quadratic Regularization

The objective function and constraints of eFCM are as follows:

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ki} \|x_k - v_i\|^2 + \lambda^{-1} \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ki} \log u_{ki}. \quad \ldots \quad (10)$$

The constraint is the same as sFCM Eq. (3).

We add the quadratic regularization term Eq. (4) to the objective function of eFCM Eq. (10) to obtain the following objective function:

$$J_e(U,V,E) = \sum_{k=1}^{n}\sum_{i=1}^{c} u_{ki}\|x_k+\delta_k-v_i\|^2 + \lambda^{-1}\sum_{k=1}^{n}\sum_{i=1}^{c} u_{ki}\log u_{ki} + \sum_{k=1}^{n} \delta_k{}^T W_k \delta_k, \quad (11)$$

Our goal here is again to derive optimal solutions U, V, and Δ which minimize objective function Eq. (11) under constraints Eq. (6). As in Section 3.1, we introduce the following Lagrange function to solve the optimization problem:

$$L_e = J_e(U,V,E) + \sum_{k=1}^{n} \gamma_k (\sum_{i=1}^{c} u_{ki} - 1).$$

After considering how to derive optimal solutions of U, from

$$\frac{\partial L_e}{\partial u_{ki}} = \|x_k+\delta_k-v_i\|^2 + \lambda^{-1}(\log u_{ki}+1) + \gamma_k = 0,$$

we have

$$u_{ki} = \frac{s_{ki}}{\sum_{l=1}^{c} s_{kl}},$$

where

$$s_{ki} = e^{-\lambda\|x_k+\delta_k-v_i\|^2}.$$

Optimal solutions of V are obtained as in Section 3.1:

$$v_{ij} = \frac{\sum_{k=1}^{n} u_{ki}(x_{kj}+\delta_{kj})}{\sum_{k=1}^{n} u_{ki}}.$$

To optimally solve Δ, we do again as done in Section 3.1:

$$\delta_k = -(A'_k)^{-1}B'_k.$$

Here,

$$A'_k = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} + \begin{pmatrix} w_{k11} & \cdots & w_{kp1} \\ \vdots & \ddots & \vdots \\ w_{k1p} & \cdots & w_{kpp} \end{pmatrix} = I + W_k^T,$$

$$B'_k = \begin{pmatrix} \sum_{i=1}^{c} u_{ki}(x_{k1}-v_{i1}) \\ \vdots \\ \sum_{i=1}^{c} u_{ki}(x_{kp}-v_{ip}) \end{pmatrix}.$$

We again require that A'_k be regularized optimal solutions Δ.

4. Proposed Algorithms

We propose new fuzzy c-means clustering algorithms for uncertain data using quadratic regularization, i.e., standard fuzzy c-means clustering algorithms for uncertain data with quadratic penalty-vector regularization (sFCMQ) and entropy-regularized c-means clustering for uncertain data with quadratic penalty-vector regularization (eFCMQ). Numerical examples are shown below.

Algorithm 1 sFCMQ

sFCMQ1 Give the values of m and W_k, and set initial values of V and Δ.

sFCMQ2 Update U using the following optimal solution:

$$u_{ki} = \frac{s_{ki}}{\sum_{l=1}^{c} s_{kl}},$$

where

$$s_{ki} = \|x_k+\delta_k-v_i\|^{-\frac{2}{m-1}}.$$

sFCMQ3 Update V using the following optimal solution:

$$v_{ij} = \frac{\sum_{k=1}^{n} (u_{ki})^m (x_{kj}+\delta_{kj})}{\sum_{k=1}^{n} (u_{ki})^m}.$$

sFCMQ4 Update Δ using the following optimal solution:

$$\delta_k = -(A_k)^{-1}B_k.$$

sFCMQ5 If (U,V,Δ) is convergent, stop. Otherwise, return to **sFCMQ2**.

Algorithm 2 eFCMQ

eFCMQ1 Give the values of λ and W_k, and set initial values of V and Δ.

eFCMQ2 Update U using the following optimal solution:

$$u_{ki} = \frac{s_{ki}}{\sum_{l=1}^{c} s_{kl}},$$

where

$$s_{ki} = e^{-\lambda\|x_k+\delta_k-v_i\|^2}.$$

eFCMQ3 Update V using the following optimal solution:

$$v_{ij} = \frac{\sum_{k=1}^{n} u_{ki}(x_{kj}+\delta_{kj})}{\sum_{k=1}^{n} u_{ki}}.$$

eFCMQ4 Update Δ using the following optimal solution:

$$\delta_k = -(A'_k)^{-1}B'_k.$$

eFCMQ5 If (U,V,Δ) is convergent, stop. Otherwise, return to **eFCMQ2**.

4.1. sFCMQ

In proposing an sFCMQ algorithm using the above optimal solutions, we iteratively optimize the algorithm (**Algorithm 1**).

4.2. eFCMQ

In proposing an eFCMQ algorithm using the above optimal solutions, we iteratively optimize the algorithm (**Algorithm 2**).

Table 1. The table shows the artificial dataset which consists of two clusters and one point regarded as noise. Each cluster includes five data.

(x_{k1}, x_{k2})	(x_{k1}, x_{k2})
$(10, 10)$	$(9, 10)$
$(11, 10)$	$(10, 9)$
$(10, 11)$	$(20, 10)$
$(19, 10)$	$(21, 10)$
$(20, 9)$	$(20, 11)$
$(17, 80)$	

5. Numerical Examples

In examples of classification using the above algorithms, we use an artificial dataset and Polaris data.

In all algorithms, the convergence condition is

$$\max_{i,j} |v_{ij} - \bar{v}_{ij}| < 10^{-6},$$

where $\bar{v}_{ij}$ is the previous optimal solution.

5.1. Artificial Dataset

We use a simple artificial dataset in two-dimensional (2D) Euclidean space. The dataset consists of two clusters and one point regarded as noise. Each cluster includes five data (**Table 1**).

We classify the dataset into two clusters using sFCM, sFCMQ, eFCM and eFCMQ. We set $v_1 = (0, 0)$ and $v_2 = (3, 0)$ on initial cluster centers and $w_{kjl} = 1.0\delta^{jl}$. δ^{jl} is the Kronecker delta.

$\bullet$ and $\times$ is data in the first and second clusters. $\boxed{\cdot}$ is an individual cluster center. The line from each data is the penalty vector.

Figures 2 and **3** show the classification results for sFCM and eFCM. Noise data interrupts appropriate clustering results in both cases. Both sFCMQ and eFCMQ, however, give appropriate results from pulling noise data by penalty vectors (**Figs. 4** and **5**).

Both cluster centers are close to each cluster in sFCMQ, while one cluster center is closer to a cluster but the other is further in eFCMQ than in sFCMQ, so we consider eFCMQ to be more sensitive to initial values than sFCMQ.

5.2. Polaris Data

A classified dataset is the Polaris star chart and its neighboring stars are Polaris data. We classify the dataset into three clusters and set $v_1 = (0, 0)$, $v_2 = (0.5, 1.0)$ and $v_3 = (1.0, 0.0)$ on the initial cluster center. We set penalty coefficient W_k of two patterns for each algorithm – $w_{kjl} = 1.0\delta^{jl}$ and $w_{kjl} = 4.0\delta^{jl}$ as shown in **Figs. 6** and **7**.

The penalty vector becomes larger as data becomes more distant from the cluster center in these examples. The greater the value of w_{kj}, the smaller the penalty vector norm, reducing data uncertainty.

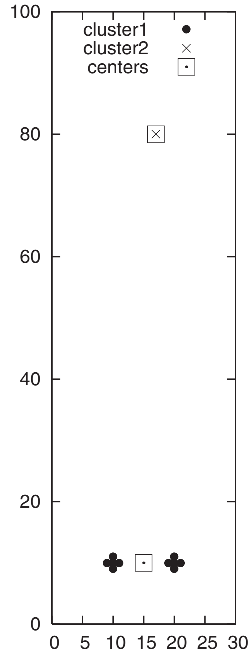


Fig. 2. This figure shows the classification result for the artificial dataset by sFCM with $m = 2$.

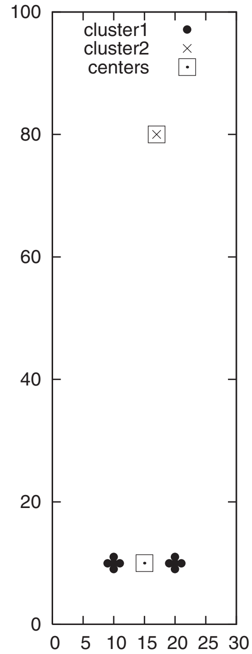


Fig. 3. This figure shows the classification result for the artificial dataset by eFCM with $\lambda = 0.1$.

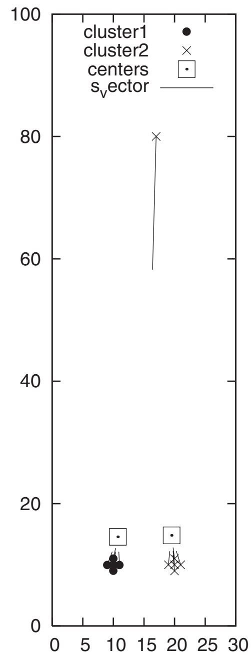


Fig. 4. This figure shows the classification result by sFCMQ with $m = 2$ and $w_{kjl} = 1.0\delta^{jl}$. δ^{jl} is the Kronecker delta.

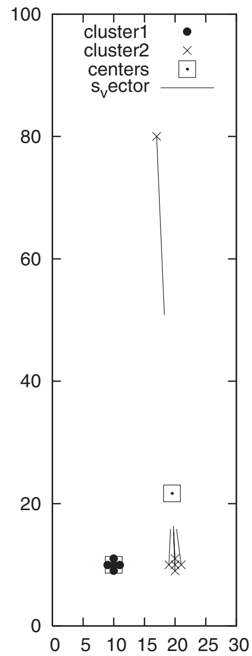


Fig. 5. This figure shows the classification result for the artificial dataset by eFCMQ with $\lambda = 0.1$ and $w_{kjl} = 1.0\delta^{jl}$. δ^{jl} is the Kronecker delta.

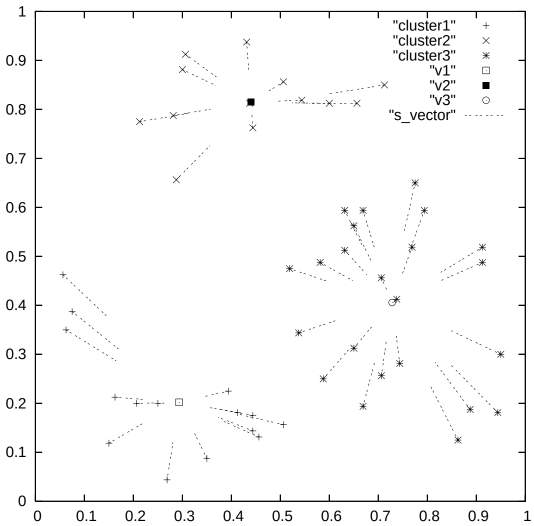


Fig. 6. This figure shows the classification result for Polaris data by sFCMQ with $w_{kjl} = 1.0\delta^{jl}$. δ^{jl} is Kronecker delta.

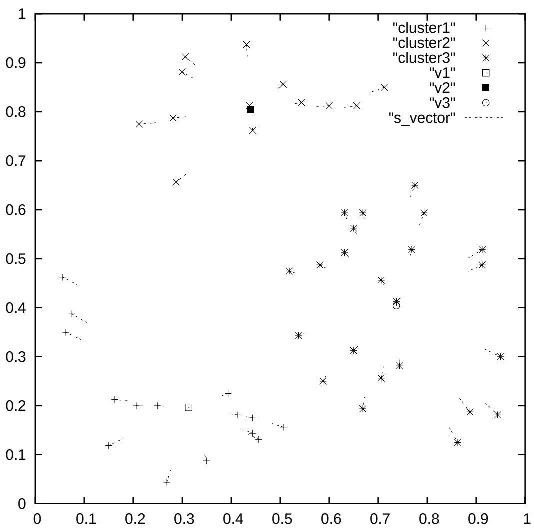


Fig. 7. This figure shows the classification result for Polaris data by sFCMQ with $w_{kjl} = 4.0\delta^{jl}$. δ^{jl} is Kronecker delta.

6. Conclusion

The clustering algorithms we have proposed for uncertain data use quadratic penalty-vector regularization. Their effectiveness was verified in numerical examples.

We now plan to derive appropriate penalty coefficient W_k corresponding to data distribution for individual algorithms. We expect this to make proposed algorithms more effective for handling uncertain data.

Acknowledgements

This study was supported in part by Grant-in-Aid for Scientific Research (C) No. 21500212 from the Japan Society for the Promotion of Science.

References:

[1] J. MacQueen, "Some methods for classification and analysis of multivariate observations," Proc. of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol.1, pp.281-297, 1967.

[2] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," Plenum, New York, 1981.

[3] O. Takata and S. Miyamoto, "Fuzzy clustering of Data with Interval Uncertainties," J. of Japan Society for Fuzzy Theory and Systems, Vol.12, No.5, pp. 686-695, 2000. (in Japanese)

[4] Y. Endo and K. Horiuchi, "On Clustering Algorithm for Fuzzy Data," In Proc. 1997 Int. Symposium on Nonlinear Theory and Its Applications, pp. 381-384, 1997.

[5] Y. Endo, "Clustering Algorithm Using Covariance for Fuzzy Data," In Proc. 1998 Int. Symposium on Nonlinear Theory and Its Applications, pp. 511-514, 1998.

[6] Y. Endo, R. Murata, H. Haruyama, and S. Miyamoto, "Fuzzy *c*-Means for Data with Tolerance," Proc. 2005 Int. Symposium on Nonlinear Theory and Its Applications, pp. 345-348, 2005.

[7] R. Murata, Y. Endo, H. Haruyama, and S. Miyamoto, "On Fuzzy *c*-Means for Data with Tolerance," J. of Advanced Computational Intelligence and Intelligent Informatics Vol.10, No.5, pp. 673-681, 2006.

Name:
Yasunori Endo

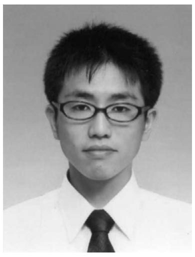
Affiliation:
Associate Professor, Department of Risk Engineering, Faculty of Systems and Information Engineering, University of Tsukuba

Address:
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

Brief Biographical History:
1994-1997 Research Assistant, Waseda University
1997-2001 Assistant Professor, Tokai University
2001-2004 Assistant Professor, University of Tsukuba
2004- Associate Professor, University of Tsukuba

Main Works:
• Y. Endo, R. Murata, H. Haruyama, and S. Miyamoto, "Fuzzy c-Means for Data with Tolerance," Proc. 2005 Int. Symposium on Nonlinear Theory and Its Applications, pp. 345-348, 2005.
• Y. Endo and K. Horiuchi, "Risk Analysis of Fuzzy Control Systems with $(n+1)$-inputs and 1-output FLC," Fuzzy Sets and Systems, Vol.147, No.3, pp. 341-361, 2004.

Membership in Academic Societies:
• Japan Society for Fuzzy Theory and Systems (SOFT)
• The Institute of Electrical and Electronics Engineers (IEEE)
• The Institute of Electronics, Information and Communication Engineers (IEICE)



Name:
Yasushi Hasegawa

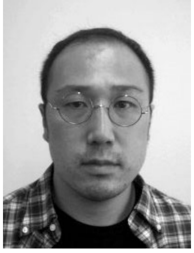
Affiliation:
DENSO Co., Ltd.

Address:
1-1 Showa-cho, Kariya, Aichi 448-8661, Japan

Brief Biographical History:
2007-2009 Masters Program, Graduate School of Systems and Information Engineering, University of Tsukuba

Main Works:
• Y. Hasegawa, Y. Endo, Y. Hamasuna, and S. Miyamoto, "Fuzzy *c*-Means for Data with Tolerance Defined as Hyper-rectangles," Modeling Decisions for Artificial Intelligence (MDAI2007), 2007.

Membership in Academic Societies:
• Japan Society for Fuzzy Theory and Systems (SOFT)



Name:
Yuchi Kanzawa

Affiliation:
Associate Professor, Faculty of Engineering, Shibaura Institute of Technology

Address:
3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan

Brief Biographical History:
2004- Associate Professor, Shibaura Institute of Technology

Main Works:
• Y. Kanzawa, Y. Endo, and S. Miyamoto, "Fuzzy *c*-Means Algorithms for Data with Tolerance Using Kernel Functions," IEICE Trans. Fundamentals, Vol.E91-A, No.9, pp. 2520-2534, 2008.

Membership in Academic Societies:
• The Institute of Electronics, Information and Communication Engineers (IEICE)

Name:
Yukihiro Hamasuna

Affiliation:
Research Fellow, the Japan Society for the Promotion of Science
Department of Risk Engineering, Faculty of Systems and Information Engineering, University of Tsukuba

Address:
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

Brief Biographical History:
2007- Doctoral Program, Graduate School of Systems and Information Engineering, University of Tsukuba
2009- Research Fellow, the Japan Society for the Promotion of Science

Main Works:
• Y. Hamasuna, Y. Endo, S. Miyamoto, and Y. Hasegawa, "On Hard Clustering for Data with Tolerance," J. of Japan Society for Fuzzy Theory and Intelligent Informatics, Vol.20, No.3, pp. 388-398, 2008. (in Japanese)
• Y. Hamasuna, Y. Endo, and S. Miyamoto, "On Tolerant Fuzzy *c*-Means Clustering," J. of Advanced Computational Intelligence and Intelligent Informatics (JACIII), Vol.13, No.4, pp. 421-428, 2009.
• Y. Hamasuna, Y. Endo, and S. Miyamoto, "On Tolerant Fuzzy *c*-Means Clustering and Tolerant Possibilistic Clustering," Soft Computing, Vol.14, No.5, pp. 487-494, 2010.

Membership in Academic Societies:
• Japan Society for Fuzzy Theory and Systems (SOFT)
• The Institute of Electrical and Electronics Engineering (IEEE)