Generalization of FCV, and Its Connection to Shell Clustering *

Katsuhiro Honda, Hidetomo Ichihashi
Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuen-cho, Sakai, Osaka, 599-8531, Japan
honda@ie.osakafu-u.ac.jp

Abstract

This paper proposes a fuzzy clustering technique that captures local non-linear structures of data sets by generalizing Fuzzy c-Varieties (FCV) proposed by Bezdek *et al.*. While FCV is regarded as the simultaneous application of fuzzy clustering and Principal Component Analysis (PCA), the proposed method is a hybrid technique of fuzzy clustering and generalized PCA that is useful to non-linear dimension reduction. The clustering result is closely related to Shell clustering that partitions data sets into several shell-shape clusters by using quadric shells as the prototypes.

Keywords: Fuzzy clustering, generalized principal component analysis, shell clustering.

1 Introduction

Fuzzy c-Varieties (FCV) proposed by Bezdek *et al.* [1, 2] is the fuzzy clustering method that partitions a data set into several linear clusters by using linear varieties as the prototypes of the clusters. Because the FCV algorithm estimates the vectors spanning the prototypical linear varieties by solving the eigenvalue problems of the fuzzy scatter matrices, it can be said that the vectors are equivalent to local principal component vectors derived in each cluster and the algorithm performs a simultaneous application of fuzzy clustering and Principal Component Analysis (PCA). In the algorithm, the eigenvectors that correspond to the largest eigenvalues of the fuzzy scatter matrix span the prototypical linear varieties. Using the linear combination of the objective function of FCV and that of Fuzzy c-Means (FCM) [1] that is useful to derive spherical clusters, FCV can be expanded to Fuzzy c-Elliptotypes (FCE) for the detection of ellipsoidal clusters.

*Proc. of Joint 1st International Conference on Soft Computing and Intelligent Systems and 3rd International Symposium on Advanced Intelligent Systems, #24B4-4 (2002)

On the other hand, Generalized Principal Component Analysis [3, 4] contrasts with PCA and tries to estimate a suitable non-linear coordinate system when the sample data points have non-linear distribution. The solution of an enlarged eigenvalue problem for any class of coordinates consisting of arbitrary functions of the original coordinates gives the non-linear coordinate system that characterizes the data configuration, just as in the linear case. If the goal of the analysis is non-linear reduction of dimensionality or function fitting, the function defined by the smallest eigenvalue is most meaningful and the problem can be regarded as Minor Component Analysis [5, 6, 7] in the enlarged non-linear coordinate system.

In this paper, we propose a technique for capturing local non-linear data structures by generalizing the FCV algorithm to an arbitrary coordinate system consisting of the linear combination of arbitrary functions. The goal of the analysis is to estimate the non-linear coordinate system that is most in agreement with the data distribution in each cluster and the problem corresponds to the function fitting considering the local structures of the data set.

The data partitioning derived by our method is closely related to Shell clustering in which the prototypes of clusters are defined by non-linear equations. Fuzzy c-Shells (FCS) [8] and Adaptive Fuzzy c-Shells (AFCS) [9] are the useful methods for detecting circular and elliptical shapes. However they are computationally demanding because the system of non-linear equations must be solved using a certain numerical technique in each iteration. Fuzzy c-Spherical Shells (FCSS) [10] captures circular shapes using the algebraic distance and provides the clustering result faster than the FCS algorithm because the prototypes are given explicitly. To capture not only circles and ellipses but also hyperbolas, parabolas or linear clusters, Krishnapuram *et al.* proposed Fuzzy c-Quadric Shells (FCQS) [11, 12, 13] that uses a quadratic polynomial as the distance function. The explicit equations for the prototypes are provided by solving eigenvalue problems in the FCQS algorithm. Our proposed method provides a data partitioning similar to the FCQS algorithm and the relation between the two methods is discussed in section III.

Finally, the characteristic features of our method are shown in numerical examples.

2 Generalization of Fuzzy c-Varieties

Let $X = (x_{ij})$ denotes an $(n \times m)$ data matrix consisting of m dimensional observations of n samples. FCV is a clustering method that partitions a data set into C linear clusters. The goal of the linear fuzzy clustering is to partition the data set into several clusters using local principal component vectors to express local linear structures. The objective function of FCV consists of squared distances from data points to p dimensional prototypical linear varieties spanned by linearly independent unit vectors $\boldsymbol{a}_{cj}$'s. FCV can be extended to FCE that extracts ellipsoidal clusters using the linear combination of the objective function of FCV and that of FCM. The objective function is defined as follows:

$$
\begin{aligned}
L_{fce} &= \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}^{\ \theta}\Big\{(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top(\boldsymbol{x}_i - \boldsymbol{b}_c) \\
&\qquad - \alpha \sum_{j=1}^{p} \boldsymbol{a}_{cj}^{\ \top} R_{ci}\boldsymbol{a}_{cj}\Big\}, \qquad (1)\\
R_{ci} &= (\boldsymbol{x}_i - \boldsymbol{b}_c)(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top, \qquad (2)
\end{aligned}
$$

where $\boldsymbol{b}_c = (b_{c1}, \cdots, b_{cm})^\top$ is the center of c-th cluster and u_{ci} denotes the membership degree of the data point $\boldsymbol{x}_i$ to c-th cluster with the constraint

$$
\sum_{c=1}^{C} u_{ci} = 1 \quad ; i = 1, \cdots, n. \qquad (3)
$$

The weighting exponent θ is added for fuzzification. The larger θ is, the fuzzier the membership assignments are. α is a constant which defines the trade-off between FCM and local principal component analysis. When α is 0, Eq.(1) is equivalent to the objective function of FCM. When α is 1, Eq.(1) is the sum of the squared distances between data points and prototypical linear varieties. The second term in the bracket tries to maximize the sample variance in the linear coordinate system spanned by $\boldsymbol{a}_{cj}$'s. Because we can derive $\boldsymbol{a}_{cj}$'s as the eigenvectors corresponding to large eigenvalues of the fuzzy scatter matrix, $\boldsymbol{a}_{cj}$'s are local principal component vectors. Therefore, the FCV algorithm can be regarded as the simultaneous application of PCA and fuzzy clustering and the objective function is the linear combination of two objective functions (PCA and FCM).

If the data set has non-linear singularity, we wish to reveal the data structure using non-linear functions. Gnanadesikan [3, 4] proposed the generalization of PCA to determine the non-linear coordinate system that is most nearly in agreement with the data configuration. Suppose that the bivariate response (x_1, x_2) is given and one is seeking a quadratic coordinate system. Generalized PCA searches for the feature value z

$$
z = a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4{x_1}^2 + a_5{x_2}^2 \qquad (4)
$$

such that the variance of z is maximum among all such quadratic functions of x_1 and x_2.

Considering a five dimensional vector $\boldsymbol{x}^*$

$$
\begin{aligned}
\boldsymbol{x}^* &= (x_1^*, x_2^*, x_3^*, x_4^*, x_5^*)^\top \\
&= (x_1, x_2, x_1x_2, {x_1}^2, {x_2}^2)^\top, \qquad (5)
\end{aligned}
$$

the quadratic PCA is formulated as the problem of finding the vector of coefficients $\boldsymbol{a}$ so that the variance of $z = \boldsymbol{a}^\top\boldsymbol{x}^*$ is maximum under such a normalization constraint as $\boldsymbol{a}^\top\boldsymbol{a} = 1$. The procedure is quite similar to that of linear

PCA and the optimal coefficients are derived by searching for the eigenvector corresponding to the largest eigenvalue of covariance matrix of the generalized response $\boldsymbol{x}^*$. As the method for the reduction of dimensionality, we are interested in the coordinate system in which z has minimum variance. Such coordinate system is defined by using the eigenvector corresponding to the smallest eigenvalue. For example, if all observations lie on a certain quadratic curve, the minimum variance will be 0 and z will be a constant for all observations.

In the following, we consider a technique for capturing local non-linear structures by combining the objective function of FCM and that of Generalized PCA. For the illustrative purposes, we consider the bivariate ($m = 2$) case.

Let $\boldsymbol{x}_i = (x_{i1}, x_{i2})^\top, i = 1, \cdots, n$ be a set of bivariate feature vectors and the goal is to partition the data set into C clusters estimating the quadratic curve in each cluster. The objective function for the simultaneous application of FCM and quadratic PCA is defined as follows:

$$
\begin{aligned}
L_{qfce} &= \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}^{\ \theta}\Big\{\alpha(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top(\boldsymbol{x}_i - \boldsymbol{b}_c) \\
&\qquad + (1-\alpha)\boldsymbol{a}_c^\top R_{ci}^* \boldsymbol{a}_c\Big\}, && (6) \\
R_{ci}^* &= (\boldsymbol{x}_i^* - \boldsymbol{b}_c^*)(\boldsymbol{x}_i^* - \boldsymbol{b}_c^*)^\top. && (7)
\end{aligned}
$$

The first term in the bracket $(\boldsymbol{x}_i - \boldsymbol{b}_c)^\top(\boldsymbol{x}_i - \boldsymbol{b}_c)$ corresponds to the clustering criterion of FCM in the original 2-D space and $\boldsymbol{b}_c$ is the center of c-th cluster. $\boldsymbol{x}_i^*$ is the five dimensional vector of response where the first two elements are the original variables and the remaining three elements are functions of the first two elements,

$$
\begin{aligned}
\boldsymbol{x}_i^* &= (x_{i1}^*, x_{i2}^*, x_{i3}^*, x_{i4}^*, x_{i5}^*)^\top \\
&= (x_{i1}, x_{i2}, x_{i1}x_{i2}, x_{i1}^{\ 2}, x_{i2}^{\ 2})^\top, && (8)
\end{aligned}
$$

and $\boldsymbol{b}_c^*$ is the cluster center whose first and second elements are equivalent to b_{c1} and b_{c2} respectively. The second term $\boldsymbol{a}_c^\top R_{ci}^* \boldsymbol{a}_c$ tries to minimize the variance of the linear combination

$$
z_c = a_{c1}x_1 + a_{c2}x_2 + a_{c3}x_1x_2 + a_{c4}{x_1}^2 + a_{c5}{x_2}^2 \qquad (9)
$$

in c-th cluster. α is the trade-off coefficient. Because the goal of the analysis is to estimate the local non-linear curves, we consider the FCM clustering only in the earlier stage of the analysis for fear that the derived non-linear curves should be distorted. The trade-off coefficient α decreases according to the annealing schedule,

$$
\alpha = \alpha_0 \beta^{t-1}, \qquad (10)
$$

where α_0, β and t are the initial value of α, the cooling rate and the iteration index respectively.

The solution algorithm is a fixed-point iteration scheme as in the FCV (or FCE) algorithm. Using Lagrangean multiplier method, new cluster center and membership are derived as follows:

$$\boldsymbol{b}_c^* = \frac{\sum_{i=1}^{n} u_{ci}^{\theta} \boldsymbol{x}_i^*}{\sum_{i=1}^{n} u_{ci}^{\theta}}, \tag{11}$$

$$u_{ci} = \left\{ \sum_{l=1}^{C} \left(\frac{E_{ci}}{E_{li}} \right)^{\frac{1}{m-1}} \right\}^{-1}, \tag{12}$$

where

$$\begin{aligned} E_{li} &= \alpha(\boldsymbol{x}_i - \boldsymbol{b}_l)^\top (\boldsymbol{x}_i - \boldsymbol{b}_l) \\ &\quad + (1-\alpha)\boldsymbol{a}_l^\top R_{li}^* \boldsymbol{a}_l. \end{aligned}$$

The optimal coefficient vector $\boldsymbol{a}_c$ that minimizes the variance of z_c in c-th cluster is derived by solving the following eigenvalue problem.

$$S_{fc}\boldsymbol{a}_c = \lambda_c \boldsymbol{a}_c, \tag{13}$$

where λ_c is the Lagrangean multiplier and S_{fc} denotes the generalized fuzzy scatter matrix in c-th cluster,

$$S_{fc} = \sum_{i=1}^{n} u_{ci}^{\theta} R_{ci}. \tag{14}$$

Because the variance of z_c is equivalent to λ_c, the optimal $\boldsymbol{a}_c$ is derived by using the eigenvector corresponding to the smallest eigenvalue of S_{fc}. Therefore, $\boldsymbol{a}_c$ can be regarded as the local minor component vector derived in c-th cluster [14].

The following three-step algorithm is used to obtain the optimal solution.

The Generalized FCV Algorithm

Step 1, Initialize the memberships u_{ci}'s randomly in each cluster and normalize them so that they satisfy the constraints Eq.(3).

Step 2, Calculate $\boldsymbol{b}_c^*$ using Eq.(11) in each cluster.

Step 3, Solve the eigenvalue problem of Eq.(13) to derive the optimal $\boldsymbol{a}_c$ in each cluster.

Step 4, Update the memberships u_{ci}'s using Eq.(12).

Step 5, If

$$\max_{i,c} \mid u_{ci}^{NEW} - u_{ci}^{OLD} \mid < \epsilon,$$

then stop. Otherwise, update α using Eq.(10) and return to **Step 2**.

If the goal of the generalized PCA is function fitting, we wish to obtain the non-linear equations that represent the characteristic feature of each cluster. The non-linear equation defines the hyper-plane of closest fit in the enlarged 5-D space and the closeness is measured by the weighted sum of the squared distances. In c-th cluster, the equation of a hyper-plane passing through a point $\boldsymbol{y}$ is defined as

$$\boldsymbol{x}^{*\top}\boldsymbol{a}_c = \boldsymbol{y}^\top \boldsymbol{a}_c, \tag{15}$$

and the weighted sum of the distances between data points and the hyper-plane is

$$d = \sum_{i=1}^{n} u_{ci}^{\theta} |(\boldsymbol{x}_i^* - \boldsymbol{y})^\top \boldsymbol{a}_c|^2. \tag{16}$$

Then, an equation of the hyper-plane of the closest fit to the data in c-th cluster is

$$\boldsymbol{x}^{*\top}\boldsymbol{a}_c = \boldsymbol{b}_c^{*\top}\boldsymbol{a}_c, \tag{17}$$

where $\boldsymbol{b}_c^*$ is the cluster center derived from Eq.(11).

Here, the validity of the derived local non-linear equations is measured by using the eigenvalues of the fuzzy scatter matrices because the weighted sum of the squared distance in c-th cluster is equivalent to the smallest eigenvalue λ_c. Then, the mean squared distance d_m can be written as follows:

$$\begin{aligned} d_m &= \sum_{c=1}^{C} \frac{\sum_{i=1}^{n} u_{ci}^{\theta} |(\boldsymbol{x}_i^* - \boldsymbol{b}_c^*)^\top \boldsymbol{a}_c|^2}{\sum_{i=1}^{n} u_{ci}^{\theta}} \\ &= \sum_{c=1}^{C} \frac{\lambda_c}{\sum_{i=1}^{n} u_{ci}^{\theta}}. \end{aligned} \tag{18}$$

The above observations are easily expanded into m-dimensional cases with arbitrary non-linear coordinate systems. For example, we can estimate arbitrary cubic equations from 2-dimensional observations by using enlarged nine dimensional responses.

3 Relation with Shell Clustering

Although we proposed a technique for capturing local non-linear structures by using an algorithm similar to FCV with generalized observations, the data partitioning is closely related to Shell clustering with non-linear curve prototypes. In this section, we discuss the relation between our technique and FCQS. FCQS is the shell clustering method whose prototypes are derived analytically and can capture not only circles and ellipses but also hyperbolas, parabolas or linear clusters. The equation of the prototypical curve in c-th cluster is defined as

$$\boldsymbol{x}^\top A_c \boldsymbol{x} + \boldsymbol{x}^\top \boldsymbol{v}_c + v_{c0} = 0, \tag{19}$$

where A_c is a symmetric matrix and $(A_c, \boldsymbol{v}_c, v_{c0})$ are the parameters that determine the shapes of the cluster. The (algebraic) distance between a data point $\boldsymbol{x}_i$ and the prototype of c-th cluster is defined as

$$d_{ic}^2 = (\boldsymbol{x}_i^\top A_c \boldsymbol{x}_i + \boldsymbol{x}_i^\top \boldsymbol{v}_c + v_{c0})^2, \tag{20}$$

and the weighted sum of the distances is minimized. Let rewrite the equation of the prototype as follows:

$$\boldsymbol{x}^\top A_c \boldsymbol{x} + \boldsymbol{x}^\top \boldsymbol{v}_c + v_{c0} = \boldsymbol{p}_c^\top \boldsymbol{q} = 0, \tag{21}$$

where

$$\begin{aligned} \boldsymbol{p}_c^\top &= (p_{c1}, \cdots, p_{cm}, p_{c(m+1)}, \cdots, p_{cr}, \\ &\quad p_{c(r+1)}, \cdots, p_{c(r+m)}, p_{cs}), & (22) \\ \boldsymbol{q}^\top &= (x_1^2, \cdots, x_m^2, x_1 x_2, \cdots, x_{n-1} x_n, \\ &\quad x_1, \cdots, x_m, 1), & (23) \\ s &= \frac{n(n+1)}{2} + n + 1 = r + n + 1. & (24) \end{aligned}$$

Krishnapuram *et al.* proposed to use the constraint [12]

$$||p_{c1}^2 + \cdots + p_{cm}^2 + \frac{1}{2} p_{c(m+1)}^2 + \cdots + \frac{1}{2} p_{cr}^2||^2 = 1 \tag{25}$$

to optimize the objective function. Under the constraint, the problem is reduced to the eigenvalue problem of a $\left(\frac{n(n-1)}{2} \times \frac{n(n-1)}{2}\right)$ matrix and the optimal $\boldsymbol{p}_c$ is derived by calculating the eigenvector corresponding to the smallest eigenvalue.

On the other hand, the objective function of our method is written as

$$\begin{aligned} L_{qfce} &= \sum_{c=1}^{C} \sum_{i=1}^{n} u_{ci}^{\theta} \left\{ \boldsymbol{a}_c^\top (\boldsymbol{x}_i^* - \boldsymbol{b}_c^*)(\boldsymbol{x}_i^* - \boldsymbol{b}_c^*)^\top \boldsymbol{a}_c \right\} \\ &= \sum_{c=1}^{C} \sum_{i=1}^{n} u_{ci}^{\theta} (a_{c1} x_1 + a_{c2} x_2 + a_{c3} x_1 x_2 \\ &\quad + a_{c4} {x_1}^2 + a_{c5} {x_2}^2 - {\boldsymbol{a}_c}^\top \boldsymbol{b}_c^*)^2, \end{aligned} \tag{26}$$

when $\alpha = 0$. Then the objective function is equivalent to that of FCQS replacing $-{\boldsymbol{a}_c}^\top \boldsymbol{b}_c^*$ with a constant term. That is, our proposed method makes it possible to solve the problem of finding non-linear clusters using the FCV-like algorithm considering the cluster centers instead of the constant term (or explicit equations of prototypes) in FCQS.

Generally, the results of Shell clustering algorithms are quite sensitive to the initial partitioning. Krishnapuram *et al.* [12] reported that a few iterations of the FCM algorithm followed by a few iterations of the G-K algorithm and a few iterations of the FCSS algorithm provided a good initialization for the FCQS algorithm. Then, it can be said that we should consider the spherical clustering in some degree in the application of Shell clustering. In our technique, we consider the spherical clustering using the objective function composed of the linear combination of two objective functions and try to solve the initialization problem using the simulated annealing technique. In our experiments, we used the validity measure of the mean squared distance in the enlarged space to evaluate the clustering results.

4 Numerical Experiments

We present two numerical examples to illustrate the characteristics of the proposed method with $\theta = 2$. In each example, we performed 100 trials changing the initial partitioning and used the following function as the annealing schedule.

$$\alpha = \begin{cases} 0.5 \times 0.9^{t-1} & ; \alpha > 0.001 \\ 0 & ; \alpha \leq 0.001 \end{cases}$$

Fig.1 shows the trajectory of the coefficient.

4.1 Extraction of Ellipses

The first artificial data set composed of 200 bivariate observations is made by using the following equations and shown in Fig.2(a).

- left ellipse (100 observations)
  $-0.42x_1 + 0.20x_2 - 0.49x_1x_2 + 0.52x_1^2 + 0.52x_2^2 = 0.19$
- right ellipse (100 observations)
  $0.42x_1 + 0.20x_2 + 0.49x_1x_2 + 0.52x_1^2 + 0.52x_2^2 = 0.19$

The data set includes two incomplete ellipses and the goal of the analysis is to capture the two ellipses partitioning the data points into two clusters ($C = 2$). Using five dimensional enlarged response Eq.(8), we applied the generalized FCV algorithm to the data set. The best result is shown in Fig.2(b) where the two elliptic curves represent the prototypes whose equations are

- prototype of first cluster
  $-0.41x_1 + 0.20x_2 - 0.50x_1x_2 + 0.52x_1^2 + 0.53x_2^2 = 0.20$
- prototype of second cluster
  $0.41x_1 + 0.20x_2 + 0.50x_1x_2 + 0.52x_1^2 + 0.53x_2^2 = 0.18$

Table 1: The Simulation Results

Method	Freq.	d_m
Generalized FCV	97	3.9×10^{-15}
FCQS	55	—

The proposed method captured the two ellipses properly although the two ellipses are incomplete. In this way, the proposed method is useful to capture the local non-linear structures.

For the comparison, we also applied the FCQS algorithm and the frequency of the best result in 100 trials are shown in Table 1. The table indicates that the generalized FCV algorithm tends to converge to the best solution in most cases and surpasses the FCQS algorithm. Thus the annealing technique is useful to overcome the initialization problem.

4.2 Estimation of Cubic Equations

The second example considers the more general case in which the data set includes subsets represented by cubic equations. The two subsets are made from

- left (lower) curve (100 observations)
  $-0.26x_1 - 0.13x_2 + 0.79x_2^2 - 0.53x_2^3 = 0.20$
- right (upper) curve (100 observations)
  $0.62x_1 - 0.12x_2 - 0.74x_1^2 + 0.25x_1^3 = 0.12$

and the data set is shown in Fig.3(a). Because the sample data points lie on the cubic equations, the FCQS algorithm cannot capture the local structures.

Using seven dimensional enlarged response (x_1, x_2, x_1x_2, x_1^2, x_2^2, x_1^3, x_2^3), we partitioned the data set into two clusters ($C = 2$). In the 100 trials, the minimum value of the validity measure is 7.1×10^{-11} and we could derive the partitioning 46 times. The result is shown in Fig.3(b) where the prototypes are represented by equations

- prototype of first cluster
  $-0.26x_1 - 0.13x_2 + 0.00x_1x_2 - 0.00x_1^2 + 0.79x_2^2$
  $+0.00x_1^3 - 0.53x_2^3 = 0.20$
- prototype of second cluster
  $0.62x_1 - 0.12x_2 - 0.00x_1x_2 - 0.74x_1^2 + 0.00x_2^2$
  $+0.25x_1^3 - 0.00x_2^3 = 0.12$

By using the suitable enlarged response, the algorithm recognized the cubic curves. In this way, the proposed method can be easily extended to the case where the shapes of clusters have highly non-linear singularity.

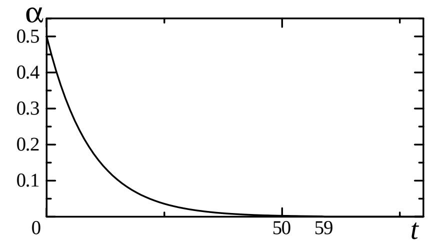


Figure 1: The trajectory of the trade-off coefficient

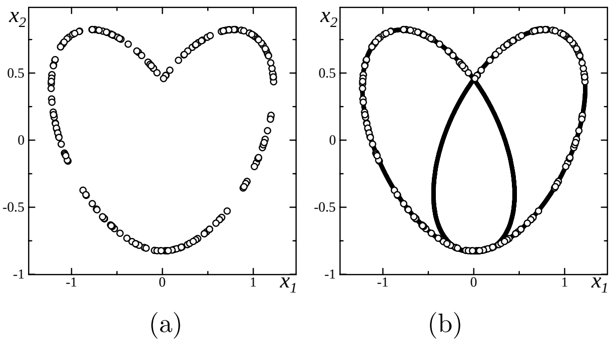
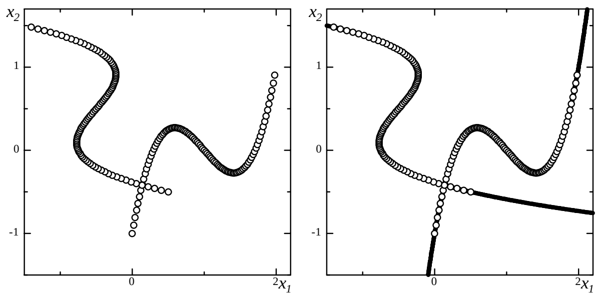


Figure 2: An example capturing two incomplete ellipses by the generalized FCV algorithm with enlarged 5-D response (a) 2-D plot of the sample data set (b) the clustering result

5 Conclusion

In this paper, we proposed to generalize the FCV algorithm, which can be regarded as the simultaneous application of PCA and fuzzy clustering, to estimate the non-linear coordinate systems that indicate the local non-linear singularity of data sets. While the proposed algorithm is similar to that of linear fuzzy clustering, the local models are closely related to that of FCQS in which the prototypes are represented by non-linear equations. Furthermore, the proposed method can be applied to more general cases by using suitable non-linear functions.

Krishnapuram *et al.* proposed the Modified FCQS algorithm in which the minimum distance between a data point and the prototype is used as the clustering criterion because the algebraic distance used in FCQS is biased, e.g. the distance is biased toward smaller curved when there are curves of highly-varying sizes. And they also proposed the possibilistic algorithms by eliminating the constraint Eq.(3) to remove the effects of noise. An alternative in robust clustering is to use absolute deviations instead of squared distances and we have proposed the algorithm for linear fuzzy clustering based on least absolute deviations [14].



(a) (b)

Figure 3: An example estimating two cubic curves by the generalized FCV algorithm with enlarged 7-D response (a) 2-D plot of the sample data set (b) the clustering result

Such modifications in the proposed method are included in the future works.

Acknowledgments

This work was supported through the scientific research fund from the Ministry of Education, Culture, Sports, Science and Technology of Japan (#13680375)

References

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

[2] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson, "Detection and Characterization of Cluster Substructure 2. Fuzzy c-Varieties and Convex Combinations Thereof," *SIAM J. Appl. Math.*, Vol.40, No.2, pp 358-372, 1981.

[3] R. Gnanadesikan and M. B. Wilk, "Data analytic methods in multivariate statistical analysis," *MultiVariate Analysis II (P. R. Krishnaiah, ed).*, Academic Press, pp 593-638, 1969.

[4] R. Gnanadesikan, *Methods for Statistical Data Analysis of Multivariate Observations*, John Wiley & Sons, 1977.

[5] E. Oja, "Principal components, minor components and liner neural networks," *Neural Networks*, Vol.5, pp.927-935, 1992.

[6] F.-L. Luo, R. Unbehauen and A. Cichocki, "A minor component analysis algorithm," *Neural Networks*, Vol.10, pp.291-297, 1997.

[7] H. Sakai and K. Shimizu, "A new adaptive algorithm for minor component analysis," *Signal Processing*, Vol.71, pp.301-308, 1998.

[8] R. N. Dave, "Fuzzy shell-clustering and application to circle detection in digital images," *Int. J. Gen. syst.*, Vol.16, pp.343-355, 1990.

[9] R. N. Dave and K. Bhaswan, "Adaptive fuzzy C-shells clustering and detection of ellipses," *IEEE Trans. on Neural Networks*, Vol.3, No.5, pp.643-662, 1992.

[10] R. Krishnapuram, O. Nasraoui and H. Frigui, "The fuzzy C spherical shells algorithms: A new approach," *IEEE Trans. on Neural Networks*, Vol.3, No.5, pp.663-671, 1992.

[11] R. Krishnapuram, H. Frigui and O. Nasraoui, "Quadric shell clustering algorithms and the detection of second degree curves," *Pattern Recognition Lett*, Vol.14, No.7, pp.545-552, 1993.

[12] R. Krishnapuram, H. Frigui and O. Nasraoui, "Fuzzy and possiblistic shell clustering algorithms and their application to boundary detection and surface approximation - partI," *IEEE Trans. on Fuzzy systems*, Vol.3, No.1, pp.29-43, 1995.

[13] R. Krishnapuram, H. Frigui and O. Nasraoui, "Fuzzy and possiblistic shell clustering algorithms and their application to boundary detection and surface approximation - partII," *IEEE Trans. on Fuzzy systems*, Vol.3, No.1, pp.44-60, 1995.

[14] K. Honda, N. Togo, T. Fujii and H. Ichihashi, "Linear fuzzy clustering based on least absolute deviations," *proc. of 2002 IEEE International Conference on Fuzzy Systems*, pp.1444-1449, 2002.