Fuzzy Regression Analysis using Fuzzy Clustering

Mika Sato-Ilic

Abstract— **This paper proposes an estimation method for fuzzy cluster loading [7] using the kernel method [3]. Fuzzy cluster loading was proposed in order to interpret the result of fuzzy clustering by obtaining the relationship between the obtained fuzzy clusters and the variables of the given data. From the structure of the model for fuzzy cluster loading, it is known that the estimate is obtained using the estimate of the weighted regression analysis [2]. In this paper, we propose a method to obtain the estimate in a higher space then the space in the given data using the idea of the kernel method. The significant properties of this technique are 1) we use high dimension space to estimate the fuzzy cluster loading, due to this, we can get a better result to extract the data structure, 2) through the cluster structure of given data, we can extract a clearer structure of the given data. Several numerical examples show the validity of the proposed technique and the efficiency of the use of the cluster structure in the given data.**

Keywords— **Kernel method. Clustering validity. Regression analysis.**

I. Introduction

Regression analysis is one widely used and well known data analysis method. If the data performs irregularly in the spatial variables, then the conventional regression analysis can not extract the data structure. So, a geographically weighted regression analysis [2] was proposed for spatial data which are not stationary situated according to geographical area. The main difference between conventional regression analysis and weighted regression analysis is the consideration of the difference among the area by the weight that shows the degree of the relationship of objects to each area.

Fuzzy cluster loading was proposed [7] in order to find the interpretation of fuzzy clustering result by the degree of the relation between a fuzzy cluster and a variable. The degree of the belongingness of objects to the fuzzy clusters also represents the state of the irregularity of the data structure, so the estimate of the fuzzy cluster loading and the estimate of the regression coefficients of the weighted regression analysis is closely related to each other. It has been shown that the fuzzy cluster loading is obtained by the same method used to estimate the regression coefficient of the weighted regression analysis. [7]

In this paper, we propose a method to obtain the fuzzy cluster loading in a higher dimension space than the data space and show that we can extract the data structure more efficiently. In order to extend the data space to higher dimension space which are nonlinearly related with the data space, we use the kernel method.[3] This model is for the nonlinear case of the fuzzy cluster loading model. We show how we can extend the model for the nonlinear case.

M. Sato-Ilic is with the Institute of Policy and Planning Sciences, Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Japan. E-mail: mika@sk.tsukuba.ac.jp .

Several numerical examples show the validity of the method and better performance when compared with the estimate of data space. Also, the results show the advantage of the use of the fuzzy cluster structure for the method.

II. Weighted Regression Analysis

The geographically weighted regression was proposed by C. Brunsdon et al. in 1998 [2], and the model is represented as follows:

$$y = V_h X \beta_h + e_h, \tag{2.1}$$

$$V_h = \begin{pmatrix} v_{h1} & 0 & \cdots & \cdots & 0 \\ 0 & v_{h2} & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & v_{hn} \end{pmatrix},$$

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{pmatrix},$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \beta_h = \begin{pmatrix} \beta_{0h} \\ \beta_{1h} \\ \vdots \\ \beta_{ph} \end{pmatrix}, \quad e_h = \begin{pmatrix} e_{1h} \\ e_{2h} \\ \vdots \\ e_{nh} \end{pmatrix},$$

where y is a vector of dependent variables and β_h a vector of regression coefficients at h-th area. V_h shows a matrix whose diagonal element v_{hi} shows the weight of i-th object to h-th area and the weights are estimated explicitly. e_h is an error vector. By roughly speaking, the main difference between a conventional regression model and the weighed model is to consider the difference among the geographical areas. The estimate of β_h is obtained as

$$\beta_h = (X^t V_h^2 X)^{-1} X^t V_h y_h. \tag{2.2}$$

III. Fuzzy Regression based on Fuzzy Clustering

In order to obtain interpretation of fuzzy clustering result, we have proposed the following model [7]:

$$u_{ik} = \sum_{a=1}^{p} x_{ia} z_{ak} + \varepsilon_{ik}, \; i = 1, \cdots, n, \; k = 1, \cdots, K, \tag{3.1}$$

where, ε_{ik} is an error.

An observed 2-way data which is composed of n objects, p variables is denoted as $X = (x_{ia}), \quad i = 1, \cdots, n, \; a = 1, \cdots, p$. z_{ak} shows the fuzzy degree which represents the amount of loading of cluster k to variable a and we call this fuzzy cluster loading. This parameter will show how each cluster can be explained by each variable. z_{ak} is assumed to be satisfied by the following conditions:

$$\sum_{k=1}^{K} z_{ak} = 1, \quad z_{ak} \in [0,1], \quad a = 1,\cdots,p, \quad k = 1,\cdots,K. \tag{3.2}$$

u_{ik} shows the obtained fuzzy clustering result as the degree of belongingness of an object i to a cluster k. The essence of fuzzy clustering is to consider not only the belonging status to the assumed clusters, but also to consider how much the objects belong to the clusters. So, there is a merit to representing the complex data situations which real data usually have.

The state of fuzzy clustering is represented by a partition matrix whose elements show the grade of belongingness of the objects to the clusters, u_{ik}, $i = 1,\cdots,n$, $k = 1,\cdots,K$, where n is number of objects and K is number of clusters. In general, u_{ik} satisfies the following conditions:

$$u_{ik} \in [0,1], \quad \sum_{k=1}^{K} u_{ik} = 1.$$

The fuzzy clustering is to obtain the adaptable partition u_{ik} from the data X.

Then the purpose of the model (3.1) is to estimate the z_{ak}, which minimize the following normalized sum of squared errors β^2 under condition (3.2).

$$\beta^2 = \frac{\sum_{i=1}^{n}\sum_{k=1}^{K}(u_{ik} - \sum_{a=1}^{p} x_{ia} z_{ak})^2}{\sum_{i=1}^{n}\sum_{k=1}^{K}(u_{ik} - \bar{u})^2},$$

$$\text{where,} \quad \bar{u} = \frac{1}{nK}\sum_{i=1}^{n}\sum_{k=1}^{K} u_{ik}.$$

We call the model (3.1) fuzzy regression based on fuzzy clustering.

The model (3.1) is rewritten as

$$\mathbf{1} = U_k X \mathbf{z}_k + \mathbf{e}_k, \tag{3.3}$$

using

$$U_k = \begin{pmatrix} u_{1k}^{-1} & 0 & \cdots & \cdots & 0 \\ 0 & u_{2k}^{-1} & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & u_{nk}^{-1} \end{pmatrix},$$

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix},$$

$$\mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{z}_k = \begin{pmatrix} z_{1k} \\ z_{2k} \\ \vdots \\ z_{pk} \end{pmatrix}, \quad \mathbf{e}_k = \begin{pmatrix} e_{1k} \\ e_{2k} \\ \vdots \\ e_{nk} \end{pmatrix}.$$

From (2.1), (2.2), and (3.3), we obtain the estimate of $\mathbf{z}_h$ as

$$\mathbf{z}_k = (X^t U_k^2 X)^{-1} X^t U_k \mathbf{1}. \tag{3.4}$$

So, if we do not assume the condition (3.2), then it is enough to use the estimate (3.4).

IV. Kernel Method

Kernel method originally developed in the context of support vector machines [3], the efficient advantage of which has been is widely reconized in many areas. The essence of the kernel method is arbitrary mapping from lower dimension space to higher dimension space. Note that the mapping is an arbitrary mapping, so we do not need to find the mapping, this is called the kernel trick.

Suppose an arbitrary mapping Φ:

$$\Phi: R^p \rightarrow F,$$

where F is a higher dimension space than R^p.

We assume

$$k(\boldsymbol{x}, \boldsymbol{y}) = \Phi(\boldsymbol{x})^t \Phi(\boldsymbol{y}),$$

where k is the kernel function which is defined in R^p and $\boldsymbol{x}$, $\boldsymbol{y} \in R^p$.

The typical examples of the kernel function are as follows:

$$k(\boldsymbol{x}, \boldsymbol{y}) = \exp(-\frac{\| \boldsymbol{x} - \boldsymbol{y} \|}{2\sigma^2}). \quad \text{(Gaussian kernel)} \tag{4.1}$$

$$k(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x} \cdot \boldsymbol{y})^d. \quad \text{(Polynomial kernel of degree d)} \tag{4.2}$$

$$k(\boldsymbol{x}, \boldsymbol{y}) = \tanh(\alpha(\boldsymbol{x} \cdot \boldsymbol{y}) + \beta). \quad \text{(Sigmoid kernel)} \tag{4.3}$$

By the introduction of this kernel function, we can analyze the data in F without finding the mapping Φ explicitly.

V. Kernel Fuzzy Regression based on Fuzzy Clustering

From (3.4), we can obtain the following:

$$\begin{aligned} \mathbf{z}_k &= (X^t U_k^2 X)^{-1} X^t U_k \mathbf{1} \\ &= ((U_k X)^t (U_k X))^{-1} (U_k X)^t \mathbf{1} \\ &\equiv (C_k^t C_k)^{-1} C_k^t \mathbf{1}, \end{aligned} \tag{5.1}$$

where $C_k = (c_{ia(k)})$, $c_{ia(k)} \equiv u_{ik}^{-1} x_{ia}$, $i = 1,\cdots,n$, $a = 1,\cdots,p$. Using $\boldsymbol{c}_{a(k)}^t = (c_{ia(k)},\cdots,c_{na(k)})$, we can represent (5.1) as follows:

$$\mathbf{z}_k = (\boldsymbol{c}_{a(k)}^t \boldsymbol{c}_{b(k)})^{-1} (\boldsymbol{c}_{a(k)}^t \mathbf{1}), \; a, b = 1,\cdots,p, \tag{5.2}$$

where $C_k^t C_k = (c_{a(k)}^t c_{b(k)})$, $C_k^t \mathbf{1} = (c_{a(k)}^t \mathbf{1})$, $a, b = 1, \cdots, p$.

Then we consider the following mapping Φ:

$$\Phi: R^p \to F, \ c_{a(k)} \in R^p. \tag{5.3}$$

From (5.2) and (5.3), the fuzzy cluster loading in F is as follows:

$$\tilde{z}_k = (\Phi(c_{a(k)})^t \Phi(c_{b(k)}))^{-1} (\Phi(c_{a(k)})^t \Phi(\mathbf{1})), \ a, b = 1, \cdots, p, \tag{5.4}$$

where $\tilde{z}_k$ shows the fuzzy cluster loading in F.

Using the kernel representation $k(x, y) = \Phi(x)^t \Phi(y)$ which is mentioned in the above section, (5.4) is rewritten as follows:

$$\tilde{z}_k = (k(c_{a(k)}, c_{b(k)}))^{-1} (k(c_{a(k)}, \mathbf{1})), \ a, b = 1, \cdots, p. \tag{5.5}$$

From this, using the kernel method, we can estimate the fuzzy cluster loading in F. We call this method the kernel fuzzy regression based on fuzzy clustering.

VI. Numerical examples

At first, we use artificially created data in order to get the validity of (5.5). The data is shown in table 6.1 and figure 6.1. Table 6.1 shows values of 8 objects, $o_1, \cdots, o_8$ with respect to 7 variables $v_1, \cdots, v_7$. In figure 6.1, the abscissa shows each variable and ordinate shows the values for each variable. The solid lines show the objects o_1, o_2, o_3, and o_4 and the dotted lines show the objects o_5, o_6, o_7, and o_8. From this, we can see the similarity among objects o_1, o_2, o_3 and o_4 as well as the similarity among objects o_5, o_6, o_7 and o_8. The feature of this data is that the first group which consists of objects o_1, o_2, o_3, and o_4 has the property that values of variables v_1 and v_7 are large and values of variables v_2 to v_6 are small. While, the other group has an opposite property.

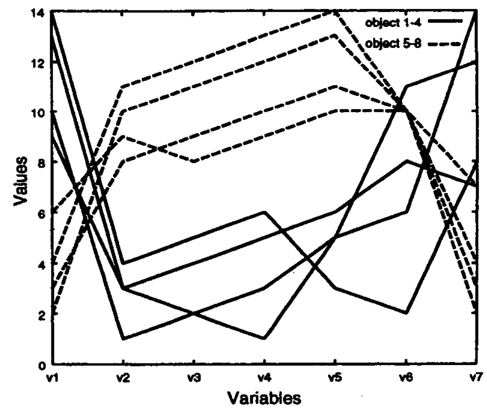

Figure 6.1 Artificial Data

The result of fuzzy k-means using $m = 2.0$ [1] is shown in table 6.2. In this table, each value shows degree of belongingness to each cluster, cluster C_1 and cluster C_2. From this result we can see that objects o_1 to o_4 for the most part belong to cluster C_1 and objects o_5 to o_8 belong to cluster C_2. According to the feature of this data, which was created with the two clusters, this result is quiet adaptable.

However, usually we do not know the property of the data structure shown in figure 6.1, so after the clustering we get the result shown in table 6.2. The problem in this case is how to find the property of the clusters C_1 and C_2.

In order to solve this problem, we use the model of fuzzy cluster loading shown in (3.1) and the kernel fuzzy regression method shown in (5.5). Figure 6.2 shows the result. In figure 6.2, the abscissa shows each variable and ordinate shows the values of fuzzy cluster loading which was obtained by using (5.5). We used the gaussian kernel which is shown in (4.1) and $\sigma = 2.0$. The solid line shows the values of fuzzy cluster loading for cluster C_1 and the dotted line shows the values for cluster C_2. From this result, we can see that cluster C_1 is related with variables v_2 to v_6, because the values of degree of proportion are large. On the other hand, cluster C_2 is explained by variables v_1 and v_7, because the values of fuzzy cluster loading for these variable are large. This is a property which we can see in the data (shown in figure 6.1).

Table 6.1 Artificial data

Variables	v_1	v_2	v_3	v_4	v_5	v_6	v_7
o_1	10	1	2	3	5	11	12
o_2	9	3	4	5	6	8	7
o_3	13	3	2	1	5	6	14
o_4	14	4	5	6	3	2	8
o_5	4	11	12	13	14	10	7
o_6	6	9	8	9	10	10	2
o_7	2	10	11	12	13	10	3
o_8	3	8	9	10	11	10	4

Table 6.2 Fuzzy Clustering Result ($m = 2.0$)

Clusters	C_1	C_2
o_1	0.75	0.25
o_2	0.64	0.36
o_3	0.88	0.12
o_4	0.83	0.17
o_5	0.12	0.88
o_6	0.06	0.94
o_7	0.04	0.96
o_8	0.03	0.97

However, the conventional estimation method used the weighted regression analysis which is shown in (3.4), but could not extract the structure of the data. In this case, the solution of the fuzzy cluster loading is the solution which is obtained in R^p. Compared to this result, note that the result shown in figure 6.2 is the result in F (mapped higher dimension space). The result of (3.4) is shown in figure 6.3. From this figure, we could not find any adjusted result to the cluster structure of the data shown in figure 6.1.

Moreover, figure 6.4 shows the result of (5.5) when U_k, $k = 1, 2$ are $n \times n$ unit matrixes in (5.1), that is the case that we do not consider the fuzzy clustering result. In (5.5), we used the gaussian kernel when $\sigma = 2.0$. In figure 6.4, the abscissa shows each variable and ordinate shows the values of loading which show the relationship with each variable. From this result, we can see that this result also does not show any clear interpretation of the data. We can also see the merit of the use of fuzzy clustering for the method to obtain the significant latent structure of the given data.

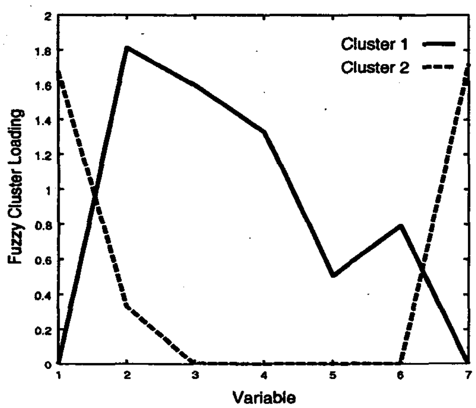

Figure 6.2 Result of Fuzzy Cluster Loading in F using Kernel Fuzzy Regression based on Fuzzy Clustering using Gaussian Kernel

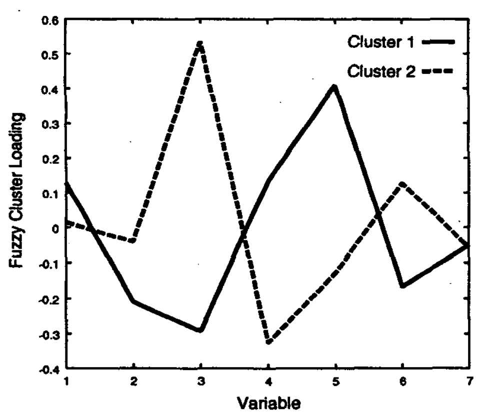

Figure 6.3 Result of Fuzzy Cluster Loading in R^p using Fuzzy Regression based on Fuzzy Clustering

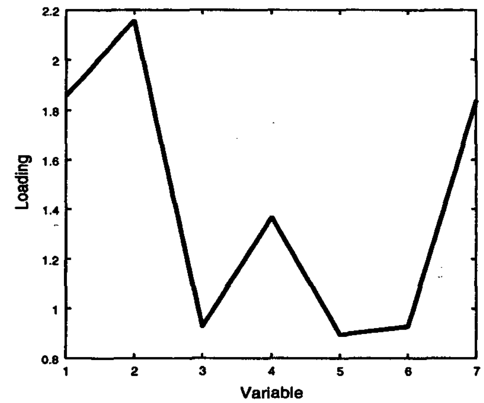

Figure 6.4 Result of Loading in F using Kernel Fuzzy Regression using Gaussian Kernel

Next, we will show an example which uses (5.5) when the kernel function is polynomial kernel shown in (4.2). The data is made up of the measurements of rain fall from 328 locations around Japan over a 12 months period. [8] Degree of belongingness of a location to each cluster, u_{ik}, is obtained by using the fuzzy k-means method when $m = 2.0$. The data was classified into 5 clusters, sapporo/sendai, tokyo, osaka, fukuoka, and okinawa areas.

Figure 6.5 shows the result of fuzzy cluster loadings in (5.5) when k is polynomial kernel, $d = 1$. Note when $d = 1$ in (4.2), then (5.5) is reduced to be (5.2). So, this is the same as finding the solution of fuzzy cluster loading in R^p. In figure 6.5, the abscissa shows each month (variable) and ordinate is the values of fuzzy cluster loadings. Each line shows each cluster. From this figure, we can see that the sapporo/sendai area has an opposite situation from the other areas. Especially, in the month of February, the sapporo/sendai area does not have as much rain fall,they receive snow due to lower temperatures.

Figure 6.6 shows the result of (5.5) used polynomial kernel when $d = 2$. In this case, the estimated fuzzy cluster loading is a solution which is obtained in F (mapped higher dimension space). From this figure, we can see that the same feature of February, that is, Sapporo/Sendai have remarkable difference compared with the other four area. Also, we can see clearer properties in figure 6.6, comparing with the result shown in figure 6.5. For example, in May, Sapporo/Sendai has clearly different feature from other four area in figure 6.5, but in figure 6.6, we see the difference is small and Sapporo/Sendai, Tokyo, and Osaka are similar to each other and next smaller value is Fukuoka and the smallest value is Okinawa. Since those five area are located from north to south according to the order, Sapporo/Sendai, Tokyo, Osaka, Fukuoka, Okinawa, the values for fuzzy cluster loading in May are arranged in order from north to south location. So, the result seems to be reasonable.

Moreover, in November, we can see that similarity between Sapporo/Sendai and Okinawa in figure 6.5, this is difficult to explain because the location of these two areas are completely different, that is, northern part and southern part. In figure 6.6, we can not find any remarkable similarity of those two area. From the comparison of these two results in figures 6.5 and 6.6, it seems clearer to use result in figure 6.6 to explain the data.

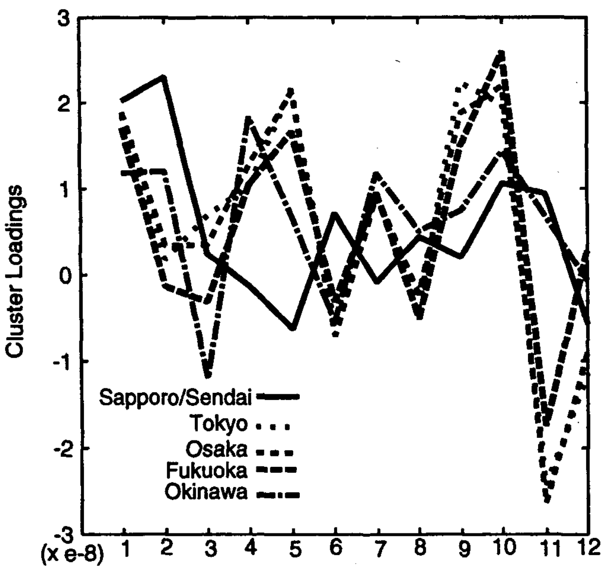

Figure 6.5 Fuzzy Cluster Loadings for Rain Fall Data using Polynomial Kernel ($d = 1$)

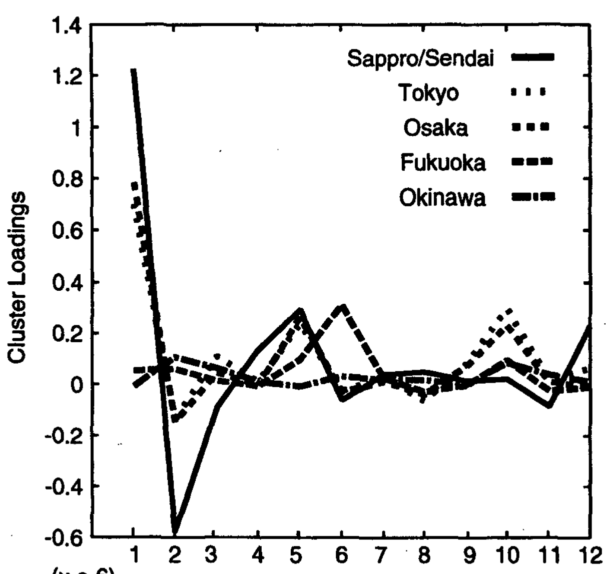

Figure 6.6 Fuzzy Cluster Loadings for Rain Fall Data using Polynomial Kernel ($d = 2$)

VII. Conclusion

In this paper, a method of using the kernel method to obtain fuzzy cluster loading was proposed.

Conventional clustering means classifying the given observation into exclusive subsets (clusters). So, we can discriminate clearly if an object belongs to a cluster or not. However, such a partition is not sufficient to represent many real situations. So, a fuzzy clustering method is offered to contract clusters with vague boundaries, this method allows one object to belong to some overlapping clusters with some grades. From this, it is known that fuzzy clustering is an efficient technique for real complex data.

However, replaced by the representativeness of fuzzy clustering to real complex data, the interpretation of such a fuzzy clustering causes us some confusion, because we sometimes think that objects which have a similar degree of belongingness can together form one more cluster. In order to have the interpretation of the obtained fuzzy clusters, we have proposed a fuzzy cluster loading which can show the relationship between the clusters and the variables.

Related with the interpretation of fuzzy clustering result, validity of the fuzzy clustering result have been discussed. In the conventional measures of validity of fuzzy clustering, partition coefficient or entropy coefficient are well known. [1] However, these measures are essentially based on the idea that clear classification is a better result. Using the idea of within class dispersion and between class dispersion, separation coefficients are introduced. [5] Moreover, according to the fuzzy hypervolume, partition density was discussed. [4] Recently, a method of evaluation of fuzzy clustering result which used the idea of the homogeneity of homogeneity analysis has been proposed. [6] These measures show the extraction of relations of the fuzzy clustering result, which is the degree of belongingness of objects to clusters, and observations indirectly. On the other hand, fuzzy cluster loading is a direct measure extracted from the data and the clustering result.

It is known that the estimate of the fuzzy cluster loading is reduced to the estimate of the regression coefficient of the weighted regression analysis. So, we can obtain the solution of the fuzzy cluster loading using the same method of weighted regression analysis.

In this paper, we extended the method to higher dimension space using the idea of the kernel method and in the high dimension space we obtained the solution of the fuzzy cluster loading.

Several numerical examples show the validity of this method and the efficiency of the idea to estimate the fuzzy cluster loading in the mapped higher space using the kernel trick of the kernel method.

For the further discussion, from the extension of fuzzy cluster loadings for 3-way data, we can consider the temporal situation in nonlinearly mapped higher dimension space.

Suppose an observed 3-way data which is composed of n objects, p attributes and T situations (or times) which is denoted as

$$X = (x_{ia}^{(t)}), \ i = 1, \cdots, n, \ a = 1, \cdots, p, \ t = 1, \cdots, T.$$

For the classification of such a 3-way data, there are two ways: first, to get a clustering result for each time we can obtain T clustering results as $U^{(t)} = (u_{ik}^{(t)}), \ t = 1, \cdots, T$. Second, to get a clustering result through times as $U = (u_{ik})$. In this case, there is merit in capturing the changing patterns of clusters through times.

Models for the interpretation of this fuzzy clustering result for 3-way data are proposed as follows:

$$u_{ik} = \sum_{a=1}^{p} x_{ia}^{(t)} z_{ak}^{(t)} + \varepsilon_{ik}, \qquad (7.1)$$

$$i = 1, \cdots, n, \ \ k = 1, \cdots, K, \ \ t = 1, \cdots T.$$

$$u_{ik} = \sum_{t=1}^{T} \omega_t \sum_{a=1}^{p} x_{ia}^{(t)} z_{ak} + \varepsilon_{ik}, \qquad (7.2)$$

$$i = 1, \cdots, n, \ \ k = 1, \cdots, K.$$

where, ε_{ik} is an error and u_{ik} shows the obtained fuzzy clustering result as the degree of belongingness of an object i to a cluster k. $z_{ak}^{(t)}$ shows the fuzzy degree which represents the amount of loading of cluster k to variable a at t-th time. This parameter will show how each cluster can be explained by each variable at each time. z_{ak} shows the fuzzy degree which represents the amount of loading of cluster k to variable a through all the times. So, the difference of purpose between these two models is the capturing of properties of the clusters that exist in each time, or properties of the clusters over all of the times. $\omega_t > 0$ shows weights which show degree of salience for each time. If $t = 1$, then the models for 3-way data are reduced to a model for 2-way data.

So, for (7.1), we can formulate the kernel fuzzy regression for 3-way data to obtain fuzzy cluster loading for 3-way data in a higher dimension space as follows:

$$\tilde{z}_k^{(t)} = (k(\mathbf{c}_{a(k)}^{(t)}, \mathbf{c}_{b(k)}^{(t)}))^{-1}(k(\mathbf{c}_{a(k)}^{(t)}, \mathbf{1})),$$

$$a, b = 1, \cdots, p,$$

where

$$(\mathbf{c}_{a(k)}^{(t)})^t = (c_{1a(k)}^{(t)}, \cdots, c_{na(k)}^{(t)}), \ \ c_{ia(k)}^{(t)} \equiv u_{ik}^{-1} x_{ia}^{(t)},$$

$$i = 1, \cdots, n, \ a = 1, \cdots, p, \ k = 1, \cdots, K, \ t = 1, \cdots T.$$

References

[1] J.C. Bezdek, J. Keller, R. Krisnapuram, N.R. Pal, *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Kluwer Academic Publishers, 1999.

[2] C. Brunsdon, S. Fotheringham and M. Charlton, Geographically weighted regression-modelling spatial non-stationarity, *Journal of the Royal Statistical Society*, Vol. 47, Part 3, pp. 431-443, 1998.

[3] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, 2000.

[4] I. Gath and A.B. Geva, Unsupervised optimal fuzzy clustering, *IEEE Trans. Patt. Anal. and Machine Intell.*, Vol. 11, pp. 773–781, 1989.

[5] R. Gunderson, Applications of fuzzy ISODATA algorithms to star-tracker printing systems, *Proc. Triannual World IFAC Congress*, pp. 1319–1323, 1978.

[6] M. Sato-Ilic, On Evaluation of Clustering using Homogeneity Analysis, *IEEE International Conference on Systems, Man and Cybernetics*, 3588–3593, 2000.

[7] M. Sato-Ilic, Fuzzy Cluster Loadings for Weighted Regression Analysis, *Japanese Classification Society*, pp. 71-78, 2001.

[8] Annual Report of Automated Meteorological Data Acquisition System (AMeDAS), Meteorological Business Support Center, 1999.