Clustering of TS-Fuzzy System

Getúlio Igrejas[1] and Paulo Salgado[2]
[1] *ESTiG - Instituto Politécnico de Bragança, Bragança, Portugal*
[2] *CETAV – Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal*
igrejas@ipb.pt, psal@utad.pt

Abstract

This paper presents a fuzzy c-means clustering method for partitioning symbolic interval data, namely the T-S fuzzy rules. The proposed method furnish a fuzzy partition and prototype for each cluster by optimizing an adequacy criterion based on suitable squared Euclidean distances between vectors of intervals. This methodology leads to a fuzzy partition of the TS-fuzzy rules, one for each cluster, which corresponds to a new set of fuzzy sub-systems. When applied to the clustering of TS-fuzzy system the result is a set of additive decomposed TS-fuzzy sub-systems. In this work a generalized Probabilistic Fuzzy C-Means algorithm is proposed and applied to TS-Fuzzy System clustering.

1. Introduction

Clustering methods seeks to organize a set of items into clusters such that items within a given cluster have a high degree of similarity, whereas items belonging to different clusters have a high degree of dissimilarity. These methods have been widely applied in various areas such as taxonomy, image processing, information retrieval, data mining, etc. Clustering techniques may be divided into hierarchical and partitioning methods: hierarchical methods yield complete hierarchy, whereas partitioning methods seek to obtain a single partition of the input data in a fixed number of clusters, usually by optimizing an objective function. However, the generalization of these techniques to clustering imprecisely or uncertainly data or objects is not yet explored. Recently, fuzzy set theory is more and more frequently used in intelligent systems, because of its simplicity and similarity to human reasoning.

For the past several years, fuzzy models have become an active research area because of their excellent capability of describing complex systems in a human intuitive way and have many successful applications in industry. One of the most outstanding fuzzy models in the literature is the structure introduced by Takagi and Sugeno (T-S fuzzy model) [1]. The T-S fuzzy models have functional-type consequences in fuzzy rules and thus have good performance in various applications. A simple parameterization of the consequent function is represented by a linear polynomial function. These models, for a n^{th} order polynomial, may approximate not only smooth functions but also their *n*-order derivatives with an arbitrary degree of accuracy [2].

In this work, a new fuzzy relational clustering algorithm, based on the fuzzy c-means algorithm is proposed to cluster T-S fuzzy models. This clustering process divides the fuzzy rules of a Fuzzy System into a set of classes or clusters of fuzzy rules based on similarity. From this new strategy, a T-S fuzzy system $f(x)$ can be decomposed into a set of additive T-S fuzzy systems. The proposed algorithm allows grouping a set of rules into *c* subgroups (clusters) of similar rules.

2. The Probabilistic Clustering Algorithm of T-S Fuzzy System

A fuzzy rule-based model suitable for describing a large class of nonlinear systems was introduced by Takagi and Sugeno [1] as follows:

$$R_l : IF\ x_1 \text{ is } A_1^l \text{ and } \cdots \text{ and } x_n \text{ is } A_n^l\ THEN\ y^l = f_l(\boldsymbol{x},\boldsymbol{a}) \quad (1)$$

where l=1, ... , M, R^l denotes the l^{th} IF-THEN rule and M is the numbers of rules in the rule base. x_i, i = 1, ... , n, are individual input variables, and A_i^l are the associated individual antecedent fuzzy sets of each input variable. $y^l \in R$ is the output of each rule and *a* the vector of parameters of the nonlinear function *f*. As a special case of *f*, we have a polynomial function, where *a*'s are the polynomials coefficients.

For any input vector, $\boldsymbol{x}=[x_1,\cdots,x_n]^T$, if the singleton fuzzifier, the product fuzzy inference and the centre average defuzzifier are applied, the output of the fuzzy model $\hat{y}$ is inferred as follows: [3]

$$\hat{y}=\sum_{l=1}^{M}\mu^l(x)y^l \Big/ \sum_{l=1}^{M}\mu^l(x), \text{ where } \mu^l(x)=\prod_{i=1}^{n}A_i^l(x_i) \quad (2)$$

The objective of fuzzy clustering partition is to separate a set of fuzzy rules $\Im=\{R_1, ..., R_M\}$ in c clusters in the antecedent space and e clusters in the consequent space, according to a "similarity" criterion. This process allows finding the optimal clusters centers V in the input space, the polynomial prototype Z at output space, the partition matrix, U, of combined input-output partition and the matrix W of scalars values. Each value u_{ijk} represents the membership degree of the k^{th} rule, R_k, belonging to the i^{th} cluster of the input space and j^{th} cluster of the output space. $w_{jk}\in\mathbb{R}$ is a value that express the translation of the consequent of the k^{th} rule fuzzy sets in direction of the center of j^{th} the output center of cluster. So, the projection of y^l in the cluster j is the function y_j^l, with $y_j^l = w_{jl}y^l$ and is expectable that:

$$\sum_{j=1}^{e} w_{jl} = 1, \quad l = 1,\cdots,M \quad (3)$$

Let $x_k \in S$ be a point covered by one or more fuzzy rules. Naturally, the membership degree of point x_k belonging to $(ij)^{th}$ cluster is the sum of products between the relevance of the rules l in x_k point and the membership degree of the rule l belonging to cluster ij, u_{ijl}, for all rules, i.e.:

$$\sum_{i=1}^{c}\sum_{j=1}^{e}\sum_{l=1}^{M} u_{ijl}\cdot\Re_l\left(\vec{x}_k\right) = 1, \quad \forall x_k \in S \quad (4)$$

where $\Re_l(\boldsymbol{x})$ represents the relevance function of the l^{th} fuzzy subsystem covering the point $\boldsymbol{x}$ of the Universe of Discourse.

The rule decomposition into $c \times e$ sub-relations will lead to an output fuzzy set decomposition as well.

For the Fuzzy Clustering of Fuzzy Rules Algorithm (FCFRA) the objective is to find $U=[u_{ijl}]$, $\boldsymbol{V}=[v_1,\cdots,v_c]\in R^{n\times c}$ and $\boldsymbol{Z}=[z_1,\cdots,z_e]\in R^e$ where:

$$J=\sum_{k=1}^{n}\sum_{i=1}^{c}\sum_{j=1}^{e}\sum_{l=1}^{M}\left(u_{ijl}\Re_l(\mathbf{x}_k)\right)^m\left(\left\|\mathbf{x}_k-\mathbf{v}_i\right\|^2+\left\|f_l(\mathbf{x}_k)\cdot w_{jl}-\mathbf{z}_j\right\|^2\right) \quad (5)$$

is minimized, with a weighting constant $m > 1$, with equation (3) and (4) as a constraint. z_j is the prototype function of the j^{th} cluster in output space, here considered to be of polynomial type of order one:

$$z_j(\boldsymbol{x}_k)=[1\ \boldsymbol{x}_k]^T\boldsymbol{z}_j=\boldsymbol{x}_k^T\boldsymbol{z}_j \quad (6)$$

It can be shown that the following algorithm may lead the triplet (U^*,V^*,W^*) to a minimum. The results can be expressed by the Probabilistic FCFRA - Algorithm:

Step 1– For a set of points X={x_1,..., x_n}, with $x_i\in S$, and a set of rules $\Im=\{R_1, ..., R_M\}$, with relevance $\Re_l(\boldsymbol{x_k})$, keep c, $2\le c < np$, and initialize $U(0)\in M_{fcm}$.

Step 2– On the r^{th} iteration, with r = 0, 1, 2, ... , compute the c mean vectors.

$$v_i^{(r)}=\sum_{k=1}^{n}\left(\sum_{l=1}^{M}U_{il}^m\cdot\Re_l^m(x_k)\cdot x_k\right)\Big/\sum_{k=1}^{n}\left(\sum_{l=1}^{M}U_{il}^m\cdot\Re_l^m(x_k)\right) \quad (7)$$

where $U_{il}^m = \sum_{j=1}^{e}u_{ijl}^m$, i =1, 2, ... , e.

Step 3– Compute the new partition matrix $U(r+1)$ using the expression:

$$u_{ijl}^{(r+1)}=1\Big/\sum_{r=1}^{c}\sum_{s=1}^{e}\left(\sum_{k=1}^{n}\Re_l^m(\mathbf{x}_k)\cdot D_{ijlk}\Big/\sum_{k=1}^{n}\Re_l^m(\mathbf{x}_k)\cdot D_{rslk}\right)^{\frac{1}{m-1}} \quad (8)$$

where $D_{ijlk}=\left\|\mathbf{x}_k-\mathbf{v}_i\right\|^2+\left\|f_l(\mathbf{x}_k)\cdot w_{jl}-\mathbf{z}_j\right\|^2$.

Step 4 – Compute the new partition matrix $W(r+1)$ with the expression:

$$w_{jl}^{(r+1)}=\left(1-\hat{\theta}_l^T\sum_{r=1}^{e}V_r\right)\Big/\sum_{r=1}^{e}\left(\bar{U}_{jl}/\bar{U}_{rl}\right)^m+\hat{\theta}_lV_j \quad (9)$$

with $\bar{U}_{jl}^m=\sum_{i=1}^{c}u_{ijl}^m$, $\hat{\theta}_l=\theta_l/\left(\theta_l^T\theta_l\right)$ and $\theta_l=\sum_{k=1}^{n}\Re_l^m(\mathbf{x}_k)f_l(\mathbf{x}_k)$

Step 5 – Compute z_j with:

$$\mathbf{z}_j^{(r+1)}=\frac{\sum_{l=1}^{M}\left[U_{jl}^m\cdot w_{jl}\left(\sum_{k=1}^{n}\Re_l^m(\mathbf{x}_k)\cdot f_l(\mathbf{x}_k)\cdot\boldsymbol{x}_k\right)\right]}{\sum_{l=1}^{M}\left[U_{jl}^m\cdot\left(\sum_{k=1}^{n}\Re_l^m(\mathbf{x}_k)\cdot\boldsymbol{x}_k^T\cdot\boldsymbol{x}_k\right)\right]} \quad (10)$$

Step 6– If $\|U(r+1)\text{-U}(r)\| < \varepsilon$ then the process ends. Otherwise let $r = r + 1$ and go to step 2.

More details about this method can be found in [4].

3. Conclusions

In this work, the mathematical fundaments for Possibilistic fuzzy clustering of the T-S Fuzzy System were presented. In the FCFR the relevance concept has a significant importance.

4. References

[1] Sugeno, M., and Kang, G.T.: 'Structure identification of fuzzy model', Fuzzy Sets Syst., 1988, 28, (1), pp. 15–33

[2] Rovatti, R.: 'Takagi-sugeno models as approximators in Sobolev norms: the SISO case'. Proc. 5th IEEE Conf. on Fuzzy Systems, New Orleans, 1996, pp. 1060–1066

[3] Wang, L.X.: 'Adaptive fuzzy systems and control: design and stability analysis' (Prentice-Hall, NJ, 1994)

[4] Salgado, P. (2007), Hierarchical decomposition of the fuzzy systems by clustering process, accepted to publish.

Acknowledgements: This work was supported by Fundação para a Ciência e Tecnologia (FCT) under grant POSI/SRI/41975/2001 and CITAB-UTAD.