Clustering Algorithms for Fuzzy Rules Decomposition

Paulo Salgado
CETAV – Universidade de Trás-os-Montes e Alto Douro
psal@utad.pt

Getúlio Igrejas
Departamento de Electrotecnia
ESTiG – Instituto Politécnico de Bragança
igrejas@ipb.pt

Abstract

This paper presents the development, testing and evaluation of generalized Possibilistic fuzzy c-means (FCM) algorithms applied to fuzzy sets. Clustering is formulated as a constrained minimization problem, whose solution depends on the constraints imposed on the membership function of the cluster and on the relevance measure of the fuzzy rules. This fuzzy clustering of fuzzy rules leads to a fuzzy partition of the fuzzy rules, one for each cluster, which corresponds to a new set of fuzzy sub-systems. When applied to the clustering of a flat fuzzy system results a set of decomposed sub-systems that will be conveniently linked into a Hierarchical Prioritized Structures.

1 Introduction

In recent years, fuzzy modelling techniques became an active research area due to its successful application to complex system models, where classical methods are difficult to apply due to the lack of sufficient knowledge. The idea of fuzzy modelling consists on establish qualitative relations expressed via linguistic rules based on expert knowledge [1]. A fuzzy model is defined as a set of IF-THEN rules, used to describe these input-output relations of a complex system.

Fuzzy rules induce in this way a fuzzy partition of the product space of the input-output variables. Generally, fuzzy clustering algorithms are very suitable techniques to detect this fuzzy partition. Different authors have proposed the use of the fuzzy clustering techniques in this process, see for example [2, 3, 4]. Fuzzy clustering in the Cartesian product-space of the inputs and outputs is another tool that has been quite extensively used to obtain the antecedent membership functions [3], [5, 6]. Attractive features of this approach are the simultaneous identification of the antecedent membership functions along with the consequent local linear models and the implicit regularization [19]. However, all these techniques are used as unsupervised identification methods and never to organize the fuzzy rules of a fuzzy system.

The automated modelling techniques will be used to obtain accurate and transparent rule-based models from system measurements. This last objective is generally relinquished by a posterior reduction of the model complexity by reduction [20], merging [21] or selection [22] of the most significant rules of the fuzzy system, which can produce some lost of information. Other strategy [7], more efficient, consists on organizing the flat fuzzy system f(x) into a set of n fuzzy sub-systems $f_1(x)$, $f_2(x)$, ..., $f_n(x)$, organized in a particularly structure, by transferring the information from the original fuzzy system to other sub-systems. Each of these systems may contain information related with particular aspects of the system $f(x)$. This process has been called Separation of Linguistic Information Methodology, SLIM [11, 17]. This objective can be reached if the fuzzy system is represented in different hierarchical structures, as the HPS (Hierarchical Prioritized Structure), which allows organize the information in the prioritized fashion, [13, 14, 15].

This work addresses this fundamental aim of fuzzy modelling by using an algorithm that implements Possibilistic Fuzzy Clustering of Fuzzy Rules (P-FCFR). The proposed algorithm permits the group of a set of rules into c subgroups (clusters) of similar rules. It is a generalization of the Possibilistic Clustering Algorithm, here applied to rules instead of points in [11, 12]. With this algorithm, the system obtained from the data is transformed into a new system, organized into several subsystems, in HPS structures. The application of the P-FCFR in the fuzzy system is here used to organize the information, by distributing it among various layers of the HPS structure.

The paper is organized as follows. Firstly, a brief introduction to fuzzy systems is presented. The concept of relevance of a set of rules and of fuzzy system is reviewed. The HPS structure is described in section 3. In section 4 the FCFR strategy is proposed. An example is presented in section 5. Finally, the main conclusions are outlined in section 6.

2 The Relevance

A generic fuzzy model is presented as a collection of fuzzy rules in the following form:

R^i: IF x_1 is A_{il} and x_2 is A_{il} ... and x_i is A_{in} THEN $y=z_i(\vec{x})$

where $\vec{x} = (x_1, x_2, \cdots, x_n)^T \in U$ and $y \in V$ are linguistic variables, A_{ij} are fuzzy sets of the universes of discourse $U_i \in \boldsymbol{R}$, and $z_i(\vec{x})$ is a function of the input variables. Typically, z can take one of the following three forms: fuzzy set (Mamdani type fuzzy systems), singleton (Takagi-Sugeno) or polynomial function (Takagi-Sugeno-Kang, TSK) type fuzzy systems. Takagi-Sugeno fuzzy systems with centre average defuzzification, product-inference rule and singleton fuzzification are represented by:

$$f(\vec{x}_k) = \sum_{l=1}^{M} p^l(\vec{x}_k) \cdot \theta^l \qquad (1)$$

where $p^l(\vec{x}) = \mu^l(\vec{x}) \Big/ \sum_{l=1}^{M} \mu^l(\vec{x})$ is the *fuzzy basis functions* (FBF), M represent the number of rules, θ^l is the point at which the output fuzzy set l achieves its maximum value, and μ^l is the membership of the antecedent of rule l. The defuzzified output y of the fuzzy model is calculated as a *weighed average* [8, 9] of the outputs of all fuzzy rules.

Fuzzy Logic Systems, FLS, are based on a set of rules that map regions in an input space, U, into regions in an output space, V, describing a region in a product space $S = U \times V$. The fuzzy rules are fuzzy relations in the product space S described by a set of rules $\Im$, which create a power set of fuzzy rules $\tilde{P}(\Im)$. In the traditional systems, as equation (1), all the rules are considered as having the same contribution in the characterization of the fuzzy system. However, they will have different importance in different regions of space or in modeling fundamental relationships.

For the characterization of relative importance of sets of rules, in the modelling process, it is essential to define a relevance function.

The relevance is a measure of the relative importance of the rules that describe the region S. The relevance is a special fuzzy measure that involves the relativity of a support region, which we see as a fuzzy measure only if the support of rules agrees with region S.

Depending on the context where the relevance is to be measured, different metrics may be defined.

Definition 1: The relevance of the rule $R \in \tilde{P}(\Im)$ on a region S can be characterized by a real positive value. The normalized relevance function maps the power set of fuzzy rules $\tilde{P}(\Im)$ on the real interval $[0\ ,\ 1]$, i.e.: $\Re_S(R) \in [0\ ,\ 1]$.

In the context of fuzzy systems there are many definitions of relevance of fuzzy rules. Next, we propose one of them for the fuzzy system (1).

Definition 2: Let $\Im$ be a set of rules that map U into V, describing completely the region S. The *relevance* of a rule $R_l \in \Im$, of fuzzy system (1) in S space is defined as:

$$\Re_l(\vec{x}_k) = \frac{\mu^l(\vec{x}_k) \cdot \delta^l}{\sum_{l=1}^{M} \mu^l(\vec{x}_k) \cdot \delta^l} \qquad (2)$$

i.e., the relevance in ($\vec{x}$,y) is the maximum of the ratio between the value of the output membership function of rule l in ($\vec{x}$,y), and the value of the membership of the union (sum) of all the functions in ($\vec{x}$, y).

Let consider the Fuzzy Systems that obey to definition 3.

Definition 3: The relevance of fuzzy system in the point $\vec{x}_k \in S$ is the sum of the relevance of all rules point $\vec{x}_k \in S$ and equal to one:

$$\Re_\Im(\vec{x}_k) = \sum_{l=1}^{M} \Re_l(\vec{x}_k) = 1 \qquad (3)$$

3 The Hierarchical Prioritized Structure

A clustering algorithm is used in this work to implement the separation of information among the various subsystems. These subsystems are organized into a HPS structure, as illustrated in figure 1. This structure allows the prioritization of the rules by using a hierarchical representation, as defined by Yager [13]. If $i < j$ the rules in level i will have a higher priority than those in level j.

Consider a system with i levels, i=1,…, n-1, each level with M_i rules:

I) If U is A_{ij} and $\hat{V}_{i-1}$ is low, then V_i is B_{ij} and rule II is used;

II) V_1 is $\hat{V}_{i-1}$;

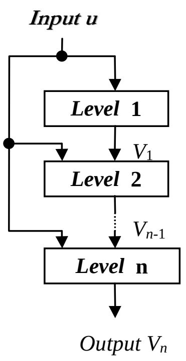

Figure 1 - Hierarchical Prioritised Structure (HPS)

Rule I is activated if two conditions are satisfied: U is A_{ij} and $\hat{V}_{i-1}$ is low. $\hat{V}_{i-1}$, which is the maximum value of the output membership function of V_{i-1}, may be interpreted as a measure of satisfaction of the rules in the previous levels. If these rules are relevant, i.e. $\hat{V}_{i-1}$ is not low, the information conveyed by these rules will not be used. On the other hand, if the rules in the previous levels are not relevant, i.e. $\hat{V}_{i-1}$ is low, this information is used. Rule II states that the output of level i is the union of the output of the previous level with the output of level i.

The output of a generic level i is given by the expression:

$$G_i = \left((1-\alpha_{i-1}) \wedge \bigcup_{l=1}^{M_i} F_i^l \right) \cup G_{i-1} \qquad (4)$$

where $F_i^l = A_{il}\left(x^*\right) \wedge B_{il}$, $l = 1, \cdots, M_i$ is the output membership function of rule l in level i.

Equation (4) can be interpreted as an aggregation operation, for the hierarchical structure. The coefficient α_i translates the *relevance* of the set of rules in level i. Level 1 gets from level 0: α_0=0, $G_0 = \varnothing$.

Paulo Salgado [16, 17] proposed a general definition of relevance for the HPS structure:

Definition 4: (*Relevance of fuzzy system just i level*) Let S_i be the input-output region covered by the set of rules of level i. The relevance of the set of rules in level i is defined as:

$$\alpha_i = S\left(\alpha_{i-1} \; , \; T\left[(1-\alpha_{i-1}), \Re_{S_i}\right]\right) \qquad (5)$$

where $\Re_{S_i} = S\left(\left\{\Re_{S_i}\left(F_i^l\right), \; l = 1, \cdots, M_i\right\}\right)$ where S and T are, respectively, *S-norm* e *T-norm* operations, and $\Re_{S_i}$ is restricted by the proposed axiomatic.

$\Re_{S_i}\left(F_i^l\right)$ represents the relevance of rule l in level i, and is defined in this work by (2). Using the product implication rule, and considering B_{il} a centroid of amplitude δ_{il} centered in $y=\ddot{y}$, then

$$\Re_{S_i}\left(F_i^l\right) = A_{il}\left(x^*\right) \cdot \delta_{il} \qquad (6)$$

When the relevance of a level i is 1, the relevance of all the levels below is null. If the *S-norm* and *T-norm* operations used in (5) are continuous and derivable, and a cost function is given, it is possible to develop an optimization process for tuning the parameters of the membership functions of the rules in a HPS system.

In the next section, an algorithm implementing the SLIM methodology for the HPS structure is presented. For simplicity, the algorithm is presented for a particular HPS structure (with three levels, n=3, two inputs, ni=2, and one output, no=1). All rules are of the type:

I) If U is A_{ij} and $\hat{V}_{i-1}$ is low, then V_i is θ_{ij} and rule II is used;

II) V_1 is $\hat{V}_{i-1}$;

The consequent rules are of the Sugeno type, with singleton membership functions, θ_{ij}. The antecedent sub-sets A_{ij} (product aggregation of the input membership functions) are of the Gaussian type. The system output is obtained by a center area defuzzifier:

$$y^* = \frac{\sum_{i=1}^{n}\sum_{j=1}^{M_i} G_n\left(b_{ij}\right) \cdot \theta_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{M_i} G_n\left(b_{ij}\right)} \qquad (7)$$

F_i may be written as

$$F_i = \bigcup_{j=1}^{M_i} \left\{ A_{ij} \big/ \theta_{ij} \right\} \qquad (8)$$

and G_n as (Yager, 1998):

$$G_n\left(b_{ij}\right) = (1-\alpha_{i-1}) \cdot A_{ij}(x) \qquad (9)$$

The relevance of the set of rules just i level, as definition 3, is here simplified by using the next expression:

$$\alpha_i = \begin{cases} \alpha_{i-1} + \Re_{S_i} & ; \text{ if } \alpha_{i-1} + \Re_{S_i} \leq 1 \\ 1 & ; \text{ otherwise} \end{cases} \qquad (10)$$

where $\Re_{S_i} = \sum_{l=1}^{M_i} \Re_{S_i}\left(F_i^l\right)$.

4 The Possibilistic Clustering Algorithm of Fuzzy Rules

4.1 The Probabilistic and Possibilistic c-means Algorithm

Clustering is well established as a way to separate a set $\boldsymbol{X} = \{x_1, x_2, \cdots, x_n\}$ into c subsets that represent (sub)structures of $\boldsymbol{X}$. A partition can be described by a $c \times n$ partition matrix U. Each element u_{ik}, $i = 1, \cdots, c$, $k = 1, \cdots, n$ of the partition matrix represents the membership of $x_k \in \boldsymbol{X}$ in the ith cluster. We distinguish two particular sets of partition matrices: the set of *fuzzy partitions*

$$M_{fcm} = \left\{ U \in [0,1]^{cn} \left| \sum_{i=1}^{c} u_{ik} = 1,\ k = 1, \cdots, n;, \cdots, c; \right. \right. \\ \left. \sum_{i=1}^{c} u_{ik} > 0,\ i = 1 \right\} \tag{11}$$

and the set of *possibilistic partitions*

$$M_{pcm} = \left\{ U \in [0,1]^{cn} \left| \sum_{i=1}^{c} u_{ik} > 0,\ k = 1, \cdots, n \right. \right\} \tag{12}$$

The most widely used fuzzy clustering [3] model is the *fuzzy c-means* (FCM) [10]. FCM is defined as the following problem: Given the data set $\boldsymbol{X}$, any norm $\|\cdot\|$ on $\mathbb{R}^p$ and a fuzziness parameter $m \in (1, \infty)$, minimize the objective function

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^{m} \cdot \|\boldsymbol{x}_k - \boldsymbol{v}_i\|^2, \quad 1 < m \leq \infty \tag{13}$$

where $U \in M_{fcm}$ and $V = \{v_1, \cdots, v_c\} \subset \mathbb{R}^p$ is a set of prototype points (cluster centers). To assign low membership to noise points in each cluster the normalization condition:

$$\sum_{i=1}^{c} u_{ik} = 1,\ k = 1, \cdots, n$$

in (11) must be dropped, which leads to possibilistic instead of fuzzy partitions. To avoid the trivial solution $u_{ik} = 0$, $i = 1, \ldots, c$, $k = 1, \ldots, n$, Krishnapuram and Keller [18] added a punishment term for low memberships to the objective function in (13) and obtained the *possibilistic c-means* (PCM) model: Given $\boldsymbol{X}$, $\|\cdot\|$, and $m \in (1, \infty)$, minimize the objective function:

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^{m} \cdot \|\boldsymbol{x}_k - \boldsymbol{v}_i\|^2 + \eta_i (1 - u_{ik})^m \tag{14}$$

where $U \in M_{pcm}$, $V = \{v_1, \cdots, v_c\} \subset \mathbb{R}^p$ as in (5), and the distance parameters $\eta_1, \cdots, \eta_c \in \mathbb{R}^+ \setminus \{0\}$ are user specified. FCM and PCM both use point prototypes for V.

It can be shown that the following algorithm may lead the pair ($\boldsymbol{U}^*$, $\boldsymbol{V}^*$) to a minimum, using alternating optimization [3]:

Possibilistic Fuzzy C-Means Algorithm

Step 1– For a set of points $X=\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$, with $\boldsymbol{x}_i \in R^p$, keep c, $2 \leq c < n$, and initialize $U^{(0)} \in M_{cf}$.

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

$$v_i^{(r)} = \frac{\sum_{k=1}^{n} \left(u_{ik}^{(r)}\right)^m \cdot \boldsymbol{x}_k}{\sum_{k=1}^{n} \left(u_{ik}^{(r)}\right)^m}, \ i=1, 2, ..., c. \tag{15}$$

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

$$u_{ik}^{(r+1)} = \frac{1}{1 + \left(\frac{\left\|\boldsymbol{x}_k - \boldsymbol{v}_i^{(r)}\right\|^2}{\eta_k} \right)^{\frac{1}{m-1}}} \tag{16}$$

for, $1 \leq i \leq c$, $1 \leq k \leq n$, where $\eta_k \in \mathbb{R}$.

Step 4– Compare $U^{(r)}$ with $U^{(r+1)}$: If $\| U^{(r+1)} - U^{(r)} \| < \varepsilon$ then the process ends. Otherwise let $r=r+1$ and go to step 2. ε is a small real positive constant.

The equation (16) defines the possibility (P-FCM) membership function for cluster i in the universe of discourse of all data vectors $\boldsymbol{X}$.

4.2 Fuzzy Clustering of Fuzzy Rules

The objective of the fuzzy clustering partition is to separate a set of fuzzy rules $\Im=\{R_1, R_2, ..., R_M\}$ in c clusters, according to a "similarity" criterion. This process allows finding the optimal clusters centre, V, and the partition matrix, U. Each value u_{ik} represents the membership degree of the k^{th} rule, R_k, belonging to the i^{th} cluster i, A_i, and obeys to equations (11) or (12), respectively for the FCM or PCM method.

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 i^{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 i, u_{il}, for all rules, i.e.:

$$\sum_{i=1}^{c}\sum_{l=1}^{M} u_{il}\cdot\Re_l\left(\vec{x}_k\right)=1\ ,\quad \forall x_k\in S \tag{17}$$

for the FCM algorithm, and only

$$\sum_{i=1}^{c}\sum_{l=1}^{M} u_{il}\cdot\Re_l\left(\vec{x}_k\right)>0\ ,\quad \forall x_k\in S \tag{18}$$

for the PCM algorithm.

A rule decomposition into c sets (sub-relations) will lead to an output fuzzy set decomposition as well. For fuzzy system (1), for each rule l we will have new centers θ_i^l associated to the output prototypes z_i, for $i=1,\cdots,c$.

$$\theta_i^l = w_{il}\theta^l$$

and is expectable that:

$$\sum_{i=1}^{c} w_{il}=1\ ,\quad l=1,\cdots,M$$

where $w_{il}\in\mathbb{R}$ is now a parameter that characterizes the deviation from the original centroid.

For fuzzy clustering, each rule and x_k point, must obey simultaneously to equations (3) and (11) or (12) , respectively for FCM or PCM algorithms. This requirements and the relevance condition of equation (3) are completely satisfied in equation (17). So, for the Fuzzy Clustering of Fuzzy Rules Algorithm, FCFRA, the objective is to find a $\boldsymbol{U}=[u_{ik}]$ and $\boldsymbol{V}=[v_1,v_2,\cdots,v_C]$ with $v_i\in R^p$ where:

$$J=\sum_{k=1}^{n}\sum_{i=1}^{c}\sum_{l=1}^{M}\left(u_{il}\Re_l\left(\boldsymbol{x}_k\right)\right)^m\left(\left\|\boldsymbol{x}_k-\boldsymbol{v}_i\right\|^2+\left\|\boldsymbol{y}_k-w_{il}\boldsymbol{z}_i\right\|^2\right) \tag{19}$$

is minimized, with a weighting constant $m>1$, with equation (17) as a constraint.

It can be shown that the following algorithm may lead the pair $(\boldsymbol{U}^*,\boldsymbol{V}^*)$ to a minimum. The models specified by the objective function (19) were minimized using alternating optimization. The results can be expressed by the following algorithm:

Possibilistic Fuzzy Clustering algorithms of fuzzy rules – P-FCAFR

Step 1– For a set of points $X=\{\boldsymbol{x}_1, \boldsymbol{x}_2,..., \boldsymbol{x}_n\}$, with $\boldsymbol{x}_i\in S$, and a set of rules $\Im=\{R_1, R_2,..., R_M\}$, with relevance $\Re_l\left(\boldsymbol{x}_k\right)$, k= 1, … , M, keep c, $2\le c<n$, and initialize $U^{(0)}\in M_{cf}$.

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

$$v_i^{(r)}=\frac{\sum_{l=1}^{M}\left[\left(u_{il}^{(r)}\right)^m\cdot\sum_{k=1}^{np}\left(\Re_l\left(\boldsymbol{x}_k\right)\right)^m\cdot\boldsymbol{x}_k\right]}{\sum_{l=1}^{M}\left[\left(u_{il}^{(r)}\right)^m\cdot\sum_{k=1}^{np}\left(\Re_l\left(\boldsymbol{x}_k\right)\right)^m\right]} \tag{20}$$

where $\left[u_{il}^{(r)}\right]=U^{(r)}$, i=1, 2, ... , c.

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

$$u_{il}^{(r+1)}=\frac{1}{\sum_{j=1}^{c}\left(\sum_{k=1}^{np}\left(\Re_l\left(\boldsymbol{x}_k\right)\right)^m\cdot\left\|\boldsymbol{x}_k-\boldsymbol{v}_i^{(r)}\right\|\Big/\eta_l\right)^{\frac{2}{m-1}}} \tag{21}$$

with $1\le i\le c$, $1\le l\le M$.

Step 4– Compare $U^{(r)}$ with $U^{(r+1)}$: If $\| U^{(r+1)}-U^{(r)}\| < \varepsilon$ then the process ends. Otherwise let r=r+1 and go to step 2. ε is a small real positive constant.

The application of the *P-FCAFR* algorithm on fuzzy system (1) rules results in a fuzzy system with a HPS structure, i.e. a system with the form:

$$f\left(x\right)=\frac{\sum_{i=1}^{c}\left(\sum_{l=1}^{M}\left(\theta^l w_{il}\right)\cdot\mu^l\left(x\right)\cdot u_{il}\right)}{\sum_{i=1}^{c}\left(\sum_{l=1}^{M}\mu^l\left(x\right)\cdot u_{il}\right)} \tag{22}$$

In conclusion, this methodology allows expand the original fuzzy system (1) into the HPS structure, with the form of equation (7).

If the rules describe one region S, instead of a set of points, the equation (21) will be reformulate to:

$$u_{il}^{(r+1)}=\frac{1}{\sum_{j=1}^{c}\left(\int_S\left(\Re_l\left(\boldsymbol{x}\right)\right)^m\cdot\left\|\boldsymbol{x}-\boldsymbol{v}_i^{(r)}\right\|\cdot d\boldsymbol{x}\Big/\eta_k\right)^{\frac{2}{m-1}}} \tag{23}$$

and if the membership of relevance function of the rules is symmetrical, the last equation can be rewrited as:

$$u_{il}^{(r+1)}=\frac{1}{\sum_{j=1}^{c}\left(\frac{\left\|\overline{\boldsymbol{x}}_l-\boldsymbol{v}_i^{(r)}\right\|}{\eta_k^*}\right)^{\frac{2}{m-1}}} \tag{24}$$

where $\overline{\boldsymbol{x}}_l$ is the center of rule l.

Similarly conclusion can be obtained between *FCAFR* [12,17] and the P-FCAFR algorithms: If all rules have the same shape, the possibilistic fuzzy clustering of fuzzy rules can be determined by only considering the center of the rules.

5 Experimental Results

In this section, an example is given to illustrate the proposed strategy for possibilistic clustering in “fuzzy rules domain”. Figure 2 shows a volcano’s surface, generated with 40×40 data points. The exercise is to capture in an HPS system the description of the function, trough the clustering decomposition of a flat fuzzy system (FS). The original structure of FS is identified from the data points using the Nearest Neighborhood Identification method, with a radius of 1.2 and a negligible error. A set of 380 fuzzy rules was generated. It is general perception that the volcano function, *W*=*F*(*U*,*V*), can be generated by the following three level hierarchical HPS structure, with one rule in each level:

Level 1: IF (*U*,*V*) is very close to (5,5) THEN *W* is quasi null (Rule 1);

Level 2: IF (*U*,*V*) is close to (5,5) THEN *W* is high (Rule 2);

Level 3: IF *U* and *V* are anything THEN *W* is low (Rule 3);

The fuzzy sets membership functions very close, close, anything, quasi null, high and low can be defined as two dimensional Gaussian functions:

$\mu_{I(i)}(\vec{x}) = e^{-\left(\frac{x_1-\bar{x}_{i1}}{\sigma_{i1}}\right)^2} \times e^{-\left(\frac{x_1-\bar{x}_{i2}}{\sigma_{j2}}\right)^2}$ with $\vec{x} = (x_1, x_2)$ and $I = \{$"**very close**", "**close**", "**anything**"$\}$;

$\mu_{J(i)}(w) = \left\{ 1/\theta_i \right\} = \left\{ 1/0 \right\}$;

with $J = \{$"**quasi null**", "**high**", "**low**"$\}$;

where $\bar{x}_{ij}$ and σ_{ij} are the central value and the width of the input membership functions, respectively, and θ_i is the central value of the singleton output membership function for the *i* level. Because there is only one rule at each level its subscripts were omitted.

Now, we begin building the HPS structure in line with the SLIM-HPS Algorithm. As mentioned, in the first step, the system is modeled by a set of rules, which is an accuracy modeling of the identified system. The output of the system at this stage is practically identical of the one shown in Figure 2.

The second step consists in the decomposition of the fuzzy rules of the FS into 3 clusters (η= 4, 10, 200, respectively for cluster 1, 2 and 3; *m*=1.2). Each one of these clusters represents a fuzzy system in a HPS structure, using the FCAFR algorithm presented. Figure 3 to Figure 5 shows the individual output response of each hierarchical fuzzy model. The original image can be described as the aggregation (equation (7)) of these three clusters surfaces.

So, the use of the P-FCAFR clustering algorithm makes the stratification of the early flat fuzzy system into a HPS structure. The results of this algorithm shown the same basic shape as the one represented in Figure 2, with just minor discrepancies.

The membership values of the fuzzy rules for each cluster are shown in Figure 6 to Figure 8 (note that the membership functions for each cluster are represented by a surface instead of its discrete values). From these figures we can observe where each cluster is “relevant” in description of the various regions of the surface. It must be noted that the 1st cluster indentifies the mountain of the volcano without the interior cavity and this last one is modeled by the 2nd cluster. The 3rd cluster identifies the foot of the mountain.

6 Conclusions

In this work, the mathematical fundaments for Possibilistic fuzzy clustering of fuzzy rules were presented. In the P-FCFR the relevance concept has a significant importance. Based on this concept, it is possible to make a possibilistic fuzzy clustering algorithm of fuzzy rules, which is naturally a generalization of possibilistic clustering algorithms.

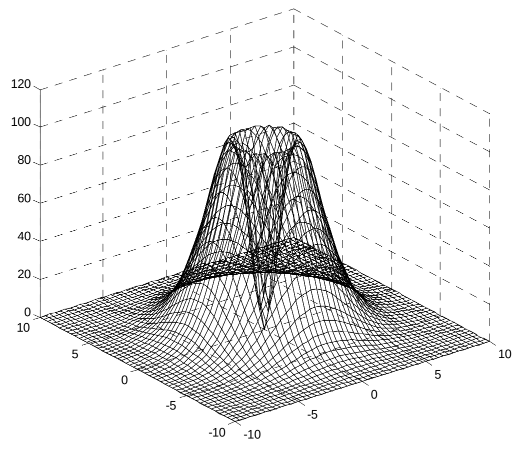

Figure 2 – Volcano surface – original system.

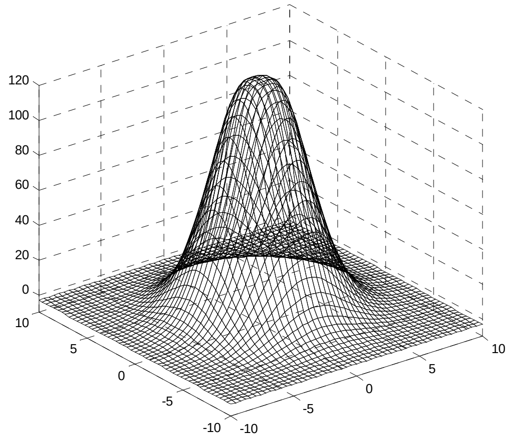

Figure 3 – Surface generated by first fuzzy system cluster.

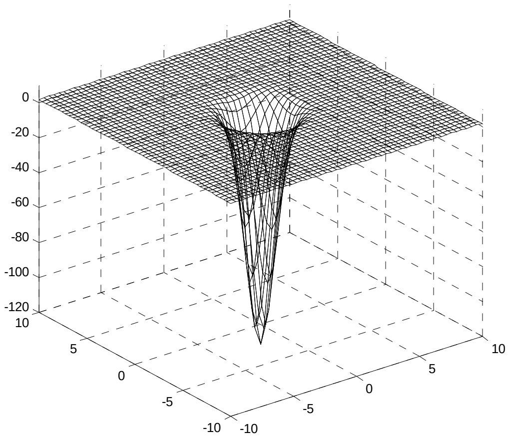

Figure 4 - Surface generated by second cluster fuzzy system - the hall.

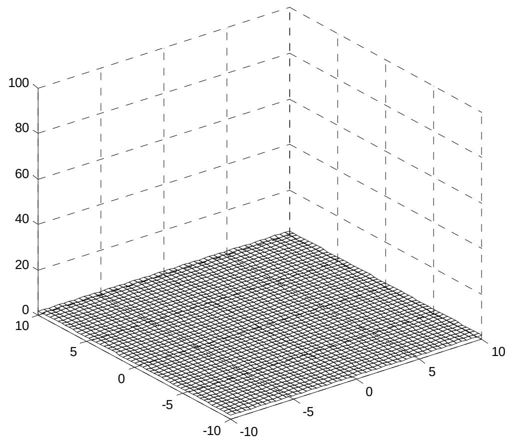

Figure 5- Surface generated by third fuzzy system cluster – the background of surface.

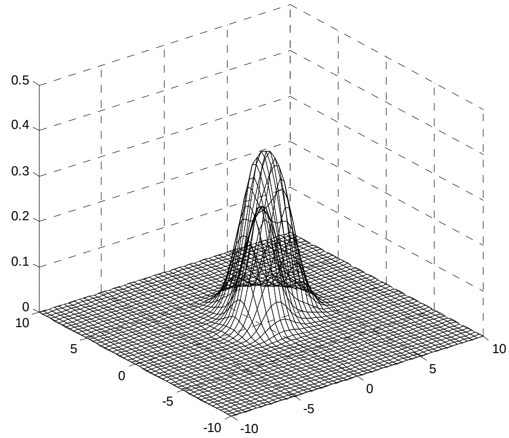

Figure 6 - Membership function u_{il}, for cluster 1.

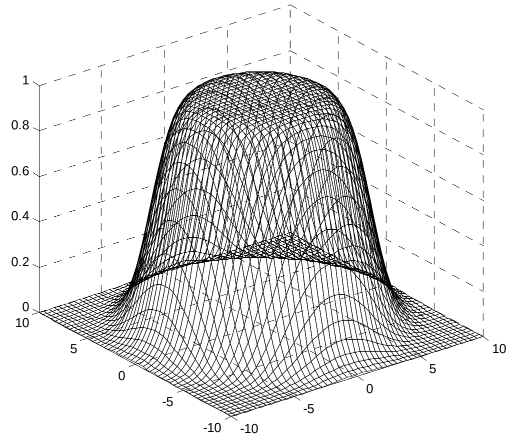

Figure 7 - Membership function u_{il}, for cluster2.

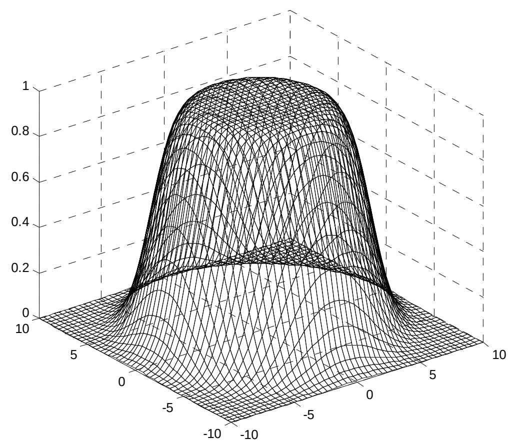

Figure 8: Membership function u_{il}, for cluster3.

Acknowledgements

This work was supported by Fundação para a Ciência e Tecnologia (FCT) under grant POSI/SRI/41975/2001 and by CETAV-Centro de Estudos Tecnológicos do Ambiente e da Vida.

References

[1] Zadeh, L. (1996). Fuzy Sets, Fuzzy Logic and Fuzzy Systems, World Scientific.

[2] Gath, I. & Geva, A. B. (1989). Unsupervised optimal fuzzy clustering, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 773–781.

[3] Höppner, F., Klawonn, F., Kruse, R. & Runkler, T. (1999). Fuzzy Cluster Analysis—Methods for Image Recognition, Classification, and Data Analysis. New York: Wiley.

[4] Abonyi, J., Babuˇska, R. & Szeifert, F. (2002). Modified Gath-Geva Fuzzy Clustering for Identification of Takagi-Sugeno Fuzzy Models, *IEEE Transactions on Systems, Man, and Cybernetics—part b: Cybernetics*, vol. 32, nº. 5, pp. 612-621.

[5] Babuska, R. (1998). Fuzzy Modeling for Control. Norwell, MA: Kluwer.

[6] Sugeno, M. & Yasukawa, T., (1993). A fuzzy-logic-based approach to qualitative modeling, *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 7–31, Jan.

[7] Salgado, P. & Boaventura, J. (2005). Greenhouse climate hierarchical fuzzy modelling, *Control Engineering Practice*, 13 pp. 613-628.

[8] Pedcrycz, W., Reformat, M. (1997). Rule-Based Modelling of Nonlinear Relationships, *IEEE Trans. On Fuzzy Systems*, 5 (2), pp. 256-269

[9] Roventa, E., Spircu, T. (2003). Averaging Procedures in Defuzzification Processes, *Fuzzy Sets and Systems* 136, pp. 375-385.

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

[11] Salgado, Paulo, (2005). Clustering and hierarchization of fuzzy systems, *Soft Computing Journal*, Vol. 9, nº 10, pp. 715-731, October, 2005, Springer Verlag.

[12] Salgado, P. (2001). Fuzzy Rule Clustering, *IEEE Conf. on Syst., Man and Cybernetics 2001*, Arizona, Tucson, AZ, USA, pp. 2421-2426.

[13] Yager, R. (1993 a). On a Hierarchical Structure for Fuzzy Modeling and Control, *IEEE Trans. On Syst., Man, and Cybernetics*, 23, pp. 1189-1197.

[14] Yager, R. (1993 b). Hierarchical representation of fuzzy if-then rules. In *Advanced Methods in Artificial Intelligence*, B. Bouchon-Meunier, L. Valverde, and R. R. Yager, Eds. Berlin, Germany: Springer-Verlag, pp. 239-247.

[15] Yager, R. (1998). On the Construction of Hierarchical Fuzzy Systems Models, *IEEE Trans. On Syst., Man, and Cyber. –Part C*, 28, pp. 55-66.

[16] Salgado, P. (1999). *Novos métodos de Identificação Difusa*, Ph.D. Thesis, UTAD, Portugal.

[17] Salgado, P (2002). Relevance of the fuzzy sets and fuzzy systems. In: *Systematic Organization of Information in Fuzzy Logic*, NATO Advanced Studies Series, IOS Press.

[18] Krishnapuram, R. & Keller, (1993). A possibilistic approach to clustering, *IEEE Trans. Fuzzy Syst.*, 1, pp. 85-110.

[19] Johansen, T. A. & Babuˇska, R. (2002). On multi-objective identification of Takagi-Sugeno fuzzy model parameters, in *Preprints 15th IFAC World Congress*, Barcelona, Spain, July.

[20] Magne Setnes, Member, IEEE, & Robert Babuˇska, (2001). Rule Base Reduction: Some Comments on the Use of Orthogonal Transforms, *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, vol. 31, no. 2, May, pp. 189-206

[21] Magne Setnes, Robert Babuˇska, Uzay Kaymak, & Hans R. van Nauta Lemke (1998). Similarity Measures in Fuzzy Rule Base Simplification, *Ieee Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics*, vol. 28, no. 3, June, pp. 376-286.

[22] Demirli, K. & Muthukumaran, P. (2000). Higher order fuzzy system identification using subtractive clustering, *Journal of Intelligent and Fuzzy Systems*, 9, pp. 129–158.