UNIVERSITÀ DEGLI STUDI DI MACERATA
Dipartimento di Istituzioni Economiche e Finanziarie

Fuzzy clustering of WTO members position in the Doha Agricultural Negotiation.

Dario Bruzzese

Temi di discussione n. 41

Febbraio 2007

Fuzzy clustering of WTO members position in the Doha Agricultural Negotiation.*

Dario Bruzzese

Abstract

In this paper a fuzzy clustering approach is exploited in order to identifies potential alliances among the members of the World Trade Organization within the round of agricultural negotiations following the November 2001 Ministerial Conference in Doha. Each country is described by a vector of ratings that summarizes its position with respect to several issues ranging from Tariffs to Labour Standard and Environmental concerns. The data have been taken from a study conducted in the Danish Research Institute of Food Economics, where the same objective has been pursued by a two-stage clustering strategy consisting of a hierarchical classification followed by a k-means method. The main objective of this research is to highlight how a fuzzy approach could better grasp the blurry contours that inevitably characterize clusters deriving from such complex issues and, at the same time, could more naturally point out the specific roles played by the different members in each cluster. The whole analysis has been carried out by means of the R language and the R-code which has been written to implement the fuzzy clustering strategy is provided.

Dario Bruzzese, University of Naples ‘Federico II’.
E-mail: dbruzzes@unina.it.

* *Financial support received by the Italian Ministry for University and Scientific and Technological Research is gratefully acknowledged (Research Program of National Scientific Relevance "Doha Development Agenda - Research Network"; Research Unit: Dipartimento di Istituzioni Economiche e Finanziarie, Universita’ di Macerata).*

1 Introduction

In the Doha ministerial conference, that took place on November 2001 in Qatar, 142 members of WTO agreed to start a new round of global trade negotiations across a broad range of topics. The scope of the negotiating mandates were specified in the Ministerial Declaration WT/MIN(01)/DEC/1 (20 November 2001) which also established the term of 1 January 2005 for the conclusion of the negotiations. Besides the agreement of most countries on several key issues, there were important topics standing out for the variance of positions among the major players.
In this framework Bjørnskov and Lind (BjLi02) proposed to analyze the position of WTO members on such debated issues in order to identify "natural" group of negotiation partners with a particular attention to the developing countries. The objective was to show how developing countries could obtain great benefits, in terms of bargaining power in the agricultural negotiations, by forming alliances with free trade-oriented WTO members.
Their approach provided a statistical clustering of WTO members in 9 different clusters. The classification relied on two standard statistical methods which are often used in a complementary manner: hierarchical classification and k-means. Each country was described by a vector of scores which summarize its position in removing both tariff and non-tariff restrictions on international trading system.
The aim of this paper is to show how Fuzzy set theory could provide an effective theoretical framework for facing such complex and hazy issues. In particular, by allowing each country to belong simultaneously, even though with different membership levels, to more than one clusters it is possible to have a deeper insight in the alliance forming process.
The paper is structured in five sections. Section 2 sketches the method that Bjørnskov and Lind adopted in their research and reports the major results in terms of the cluster structure they obtained. Further, the reasons that lead the proposed approach are outlined. Section 3 introduces the Fuzzy set theory and describes fuzzy c-means algorithm (*FCM*) from a formal point of view. A variant of the FCM algorithm allowing to face with missing values is also described. The comparison among the results of the two approaches is discussed in Section 4. Finally some concluding remarks about the necessity to adopt an external criteria to validate the classification will conclude the paper. The R-code written to implement the fuzzy strategy will be provided in the Appendix.

2 Crisp analysis of WTO members' positions

In the paper of Bjørnskof and Lind the position of 120 WTO countries (out of the 142 countries meeting in Doha) on 13 issues belonging to 4 different broad categories have been rated. In table 1 the list of specific topics that have been considered by the authors is provided. We remand to the original paper for the explanation of the economic implication of those areas of contention.

The rating system is based on a 5 point ordinal scale (ranging from 0 to 4) were higher values indicate an higher commitment for reducing real and potential barriers in international trading. Each rating has been expressed on a subjective basis by analyzing three primary sources: proposals submitted to the agricultural negotiations in 2000 and 2001, official statements during these negotiations, and official statements and declarations during and following the Doha meeting.
Actually it was impossible for the Authors to rate each member on every selected issues. About 34 % of the overall set of members positions have not been rated due to absence of clear official

Table 1: The 13 rated issues according to competence macro-area

Market Access	Tariffs; Tariffs Peak and Escalation; Tariffs Rate Quotas; Special Safeguard Clause
Export Support	Export Subsidies; Export Credit
Domestic Support	Green Box; Blue Box; Development Box; AMS; De Minimis Level
Non-trade concern	Broad vs. Narrow Round; Labour Standards and Environments

positions. Thus the data matrix contains a huge amount of missing data points that have been replaced, in the original paper, by the overall average position on any given issue.
The analysis performed by the authors is based on a two stage procedure: at first, a hierarchical classification (using an Euclidean distance and the Ward Method) has been performed in order to identify the optimal number of clusters. In particular, by cutting the tree, the authors found out 9 main clusters. These cluster has become the input for the second stage, were a non hierarchical clustering method (k-means) has been used in order to refine the previous partition. In table 2 the clusters structure obtained at the end of the whole procedure is shown.
According to our opinion the approach followed by the authors suffers of two main drawbacks. At first, the estimation of missing data by the overall mean inflates artificially the homogeneity of the data. On average, each country is characterize by 34 % of missing features (30% of the rows has at least 6 missing features). Each variable has a percentage of missing values ranging from 20% of TRQ Size to 67% of Labour Standards and Environment. It is thus mandatory to set up a classification procedure that includes such specificity in the clusters forming process.
Secondarily, the adopted methodology, like all the crisp clustering methods, lets each statistical unit to be a part just of one cluster without marking out the similarity that could exist with the other clusters. Moreover, the exclusive belonging to a specific group implies that all the units have the same importance in defining the property of their own cluster. Evidently, in a crisp clustering, it is always possible to measure the relative position of each unit with respect to all the clusters by computing, for instance, the Euclidean distance from the barycentres; however in such an approach the cluster barycentres are themselves obtained by giving the same weight to all the members and thus causing a vicious circle that could only be broken if each country is allowed to belong simultaneously, but with different degrees, to different clusters.
The latter objective can be easily reached by adopting a fuzzy approach for the identification of clusters. However, as it will be detailed in the following section, the most popular algorithm for fuzzy clustering does not allow missing features in the data set and thus an adaptation of this algorithm has been used in order to cope with the particular structure of the available data set.

3 Fuzzy Sets Theory and Fuzzy Clustering

Fuzzy sets theory was introduced by L. Zadeh (Za65) in order to encompass some of the main drawbacks of the classical approaches for modeling uncertainty in real problems. From a formal point of view, given a collection of objects $\Omega \subseteq R^n$ a fuzzy set A is defined as an ordered couple

Cluster label	Countries
Gradual liberalisation	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Côte dŠIvoire, Djibouti, Gabon, the Gambia, Ghana, Guinea, Guinea- Bissau, Lesotho, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Niger, Rwanda, Senegal, Sierra Leone, Swaziland, Tanzania, Togo, and Zambia
Market Access	Cuba, Haiti, Nicaragua, and Peru
Average	Albania, Antigua, Brunei, Bulgaria, Burma, Canada, Croatia, the Czech Republic, Dominica, Ecuador, Estonia, Grenada, Hungary, Jamaica, Japan, Jordan, the Kyrgyz Republic, Latvia, Lithuania, Mongolia, Poland, St. Kitts, St. Lucia, St. Vincent, Singapore, the Slovak Republic, Slovenia, South Korea, Suriname, Switzerland, Trinidad, Tunisia, Turkey and Venezuela
Narrow Round	Argentina, Australia, Bolivia, Brazil, Chile, Colombia, Costa Rica, Fiji, Guatemala, Honduras, Indonesia, Malaysia, Namibia, New Zealand, Paraguay, South Africa, Thailand, Uruguay and the US
Small changes	India, Mexico and the Philippines
-	The EU and Israel
Free Trade	The Dominican Republic, Egypt, El Salvador, Kenya, Nigeria, Pakistan, Sri Lanka, Uganda and Zimbabwe
-	Norway
Developing multifunctionality	Barbados and the Democratic Republic Congo

Table 2: The 9 clusters obtained by Bjørnskof and Lind

$\{\Omega, \mu_A(\omega)\}$ where

$$\mu_A(\omega) : \omega \rightarrow [0, 1]$$

is a *membership function* which measures the coherence of each element $\omega \in \Omega$ with the properties that characterize the fuzzy set A.
The major difference with respect to the *crisp* set theory is the fact that the membership function is not constrained in the set $\{0, 1\}$ but it can assume any values in the real interval $[0, 1]$ where values closer to 1 denote stronger membership of the element to the fuzzy set and vice versa.
Fuzzy set theory has been successfully used in several applicative and methodological contexts (from Credit scoring to Medicine, from Regression to Expert Systems - see (Zi91) for an in-depth survey) but certainly, the framework where fuzzy set theory has become more and more popular is Pattern Recognition and, more specially, Cluster Analysis. Here the objective is to classify *n* objects $x_1, x_2, ..., x_n \in R^p$ in to C homogeneous groups in a manner such that the relationship between groups is revealed. Usually this goal is reached through minimization of internal variation while maximizing variation between groups.
Several algorithm have been developed to efficiently perform pattern recognition in a data set; the description of such methods is outside the scope of this paper. In the work of Bjørnskov and Lind, as already pointed out, the clusters were obtained mainly with the k-means method while hierarchical classification had just played a preliminary role in order to find the correct number

of clusters.
In our research we have replaced the two-stage analysis with a unified approach that exploits fuzzy theory both to find the number of cluster and to perform the classification. The core of this strategy is represented by the Fuzzy c-means algorithm.
Fuzzy c-means, also known as Fuzzy ISODATA, represents the natural extension of the k-means in the framework of fuzzy set theory. It has been developed by Dunn in 1973 (Du73) and later improved by Bezdek ((Be73) (Be81)). It is based on the minimization of the following objective function which represents a weighted version of the within-class sum square errors:

$$J_m(U,v) = \sum_{k=1}^{c} \sum_{i=1}^{n} U_{i,k}^{m} \|x_i - v_k\|^2$$

where c is the number of clusters, n is the number of objects to classify, $m > 0$ is a fuzzification parameter, $U_{i,k}$ is the degree of membership of the i^{th} object to the k^{th} cluster, v_k is the p-dimensional prototype of the k^{th} cluster and $\|x\|^2$ is the Euclidean norm: $\|x\|^2 = \sum_{j=1}^{p} x_{ij}^2$ [1]. By tuning the value of m toward $+\infty$ the resulting partition become fuzzier while values of m near to 1 makes the minimizing values $U_{i,k}$ closer to the extreme limits of the membership support.
The minimization is subject to the following constraints set on the membership function values:

$$U_{i,k} \in [0,1] \forall i,k; \sum_{k=1}^{c} U_{i,k} = 1 \forall i; \sum_{i=1}^{n} U_{i,k} > 0 \forall k$$

in order to ensure that they are normalized to one and that no empty clusters can be output by the algorithm.
The optimization of the objective function is obtained, from a numerical point of view, through an iterative scheme which alternates the minimization of $J_m(U,v)$ over $U_{i,k}$ and v_k using the following formulas (i=1,...,n and k=1,...,C.):

$$U_{i,k} = \frac{1}{\sum_{l=1}^{c} \left(\frac{\|x_i - v_k\|}{\|x_i - v_l\|} \right)^{\frac{2}{m-1}}} \tag{3.1}$$

$$v_k = \frac{\sum_{i=1}^{n} U_{i,k}^{m} x_i}{\sum_{i=1}^{n} U_{i,k}^{m}} \tag{3.2}$$

These expressions are derived equaling to 0 the partial derivatives of $J_m(U,v)$ with respect to $U_{i,k}$ $(i = 1,..,n, k = 1,...c)$ and $v_k(k = 1,..,c)$.
The FCM algorithm can be detailed as follows:

Initialization Step

Fix m, c and ϵ with $m > 1$, $1 < c < n$ and $\epsilon > 0$. Initialize the values of prototypes $v^{(0)} \subset R^p$

[1] The Euclidean norm can be replaced by any A-norm: $\|x\|_A = x'Ax$

Minimization Step
For r=1,2,... do
Compute membership values $U_{i,k}^{(r)}$ using equation 2.1 with $v_k = v_k^{(r-1)}$
Compute the prototypes $v_k^{(r)}$ using equation 2.2 with $U_{i,k} = U_{i,k}^{(r)}$

Stopping Rule
If $\|v^{(r)} - v^{(r-1)}\| < \epsilon$ then stop. Otherwise put r=r+1 and return to the Minimization Step.

It should be noted that in the *Minimization Step*, the Fuzzy c-means algorithm updates the prototypes through a weighted average of all the units, using membership values as weights. This, in turn, implies that FCM cannot directly be applied to data sets containing missing feature values. Several generalizations of the FCM algorithm have been proposed in order to make it feasible in case of missing values. However, some of these ((HaBe02), (ZhCh03)) require the data to be already in the form of a relational matrix and not in the usual units $\times$ variables structure. In order to face with incomplete data set, Hathaway and Bezdek (HaBe01) proposed a modified version of the FCM were the missing values are iteratively estimated during the optimization process. In particular, denoting with X_M the subsets of data vectors with at least one missing feature, that is $X_m = \{x_i(i = 1, ..., n) : \exists \ s \ (s = 1, ...p) : x_{i,s} = NA\}$, the optimization of $J_m(U, v)$ is pursued by adding one more step in the Minimization Phase where missing values estimates are obtained by zeroing the gradient of $J_m(U, v)$ with respect to the missing features, thus obtaining the following equation:

$$x_{i,s}^{(r+1)} = \frac{\sum_{k=1}^{c} \left(U_{i,k}^{(r)}\right)^m v_{k,s}^{(r)}}{\sum_{k=1}^{c} \left(U_{i,k}^{(r)}\right)^m}$$

Obviously, in the Initialization the missing features have to be picked at random or, as we did, initialized with some kind of column statistic. In particular, due to the ordinal nature of the data, we adopted the column median to initialize missing values.
Actually in (HaBe01) four different strategies are exploited to perform a fuzzy c-means clustering of incomplete data sets. The one we adopted in this paper, called Optimal Completion Strategy (OPS), has been chosen because it produces the most accurate terminal prototypes in case of very high proportion of missing features, which is the case of the data gathered by Biørnskov and Lind. Moreover, this strategy produces an estimate of the missing values as a by product of computation.
Once the membership values are determined, is always possible to transform the fuzzy partition into a crisp one by assigning each unit to the cluster with the highest membership. However this process, called *defuzzification*, can lead to empty clusters especially in case of very strong overlapping among them.
The fuzzy c-means algorithm (as well as the k-means) requires the user to specify, among the other parameters, the number of clusters in the Initialization phase. In order to make fully automatic the whole procedure it is possible to reiterate the algorithm by varying each time the parameter *c* and computing for each classification a validity index which measures the quality of the resulting classification. Different cluster validity index have been proposed ((XiBe91), (GaGe89), (Ro82)). Some of them rely on geometric property of the clusters (such as compactness); others use the information of fuzzy membership grades to evaluate clustering results. In this paper we adopted the Partition Entropy Coefficient (which belongs to the latter category)

defined as:

$$PE = -\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{c} U_{i,k}\log(U_{i,k})$$

PE varies in the range $[0, \log(c)]$ where smaller values indicate a better (crisper) partition. In order to compare the validity of classifications based on a different number of clusters is thus necessary to normalize PE with respect to its maximum. The Normalized Partition Entropy is thus defined as:

$$PE* = \frac{PE}{\log(c)}$$

4 Results of Fuzzy c-Means

In this section the results of applying the Fuzzy c-means algorithm to the data reported in (BjLi02) will be discussed. By using the Normalized Partition Entropy, twelve clusters have been found, but three of them turned out to be empty after the *defuzzification* of the membership functions. Thus, the automatic procedure confirmed the cluster structure found in the paper of Biørnskov and Lind. In table 3 the composition of the nine clusters is reported; for each country the highest degree of membership of each country is also shown.

By comparing the two classifications several considerations can be made. At first, in each clusters obtained through the Fuzzy c-mans algorithm it is quite direct to identify leader members that are those with the highest values of the membership function. All these countries belong to the same cluster in both the classifications. The only cluster where does not seem to exist such pattern is cluster number two. This includes the cluster labeled *Developing multifunctionality* in the work of Bjørnskov and Lind (Barbados and the Democratic Republic Congo) and several other countries that stems almost from the whole set of the initial clusters. The fuzzy approach clearly identifies these countries as members who do not express a clear and specific position on the rated issues. It is plausible that by repeating the procedure, and thus by allowing small changes in the membership values, some of these could be incorporated in different clusters.
The isolated position of the Norway with respect to all other members is confirmed.
By looking at the clusters 5 it can be noted that it completely absorbs the *Market Access* group; however if we observe the membership values, it turns out that Cuba is characterized by a lower membership rate than Uganda (and basically it is equal to the membership value of Kenya) which does not belong to the *Market Access* cluster in the original classification. This is clearly a consequence of the different approach adopted to face with missing values.
Another consideration refers to the cluster number 9 which is almost equal to the *EU and Israel* group with the extra membership of Japan which originally belongs to the *Average* cluster. The belonging of Japan to the EU group reflects mainly their conjoint opposition to the elimination or drastic reduction of domestic and export subsidies
The counterpart of the fusion of the *Developing multifunctionality* cluster is represented by the splitting of the *Average* group mainly in to the cluster number 1 and cluster 6. In particular, among the members who belong to the latter, some of them are characterized by high membership values (Bulgaria (0.97), Kyrgyz Republic (0.97), Latvia (0.97)). By looking at the original values it turns out that those countries are among those with the highest number of missing features (about 85%) but have a very clear position on *Tariffs Peak and Escalation* and *De minimis* (they all have on such issues, respectively, a rating of 3 and 0). The substitution of missing features with the column mean in the original paper of Bjørnskov and Lind has clearly flattened this

Table 3: The 9 clusters obtained after the *defuzzification* of the membership values.

Cluster nr.	*Countries (cluster membership)*
1	Ecuador (0.95), Estonia (0.95), Mongolia (0.95), Slovenia (0.95), Jordan (0.24), Singapore (0.21), Slovak Republic (0.21), Albania (0.13), Canada (0.13), Croatia (0.12), Brunei (0.12), Myanmar (0.12), Morocco (0.1), El Salvador (0.09), India (0.09) and Poland (0.09)
2	Uruguay (0.11), Namibia (0.11), Thailand (0.11), Jamaica (0.11), Ghana (0.11), Indonesia (0.11), St Vincent (0.11), Antigua (0.1), South Africa (0.1), Honduras (0.1), Argentina (0.1), Egypt (0.1), Zimbabwe (0.1), Barbados (0.09), Turkey (0.09) and Congo DR (0.08).
3	Australia (0.78), Brazil (0.78), Guatemala (0.4), Fiji (0.35), Paraguay (0.23), CostaRica (0.17), Chile (0.17), Colombia (0.15), USA (0.14), New Zealand (0.13), Bolivia (0.13), Malaysia (0.12), Philipines (0.1) and Mexico (0.1).
4	Grenada (0.98), St Kitts (0.98), St Lucia (0.98), Surinam (0.98), Trinidad (0.98), Tunisia (0.98), Venezuela (0.98), Dominica (0.21), South Korea (0.16), Switzerland (0.15), Swaziland (0.14), Cote d'Ivoire (0.11), Lesotho (0.11) and Mauritius (0.1).
5	Haiti (0.98), Nicaragua (0.98), Peru (0.98), Uganda (0.56), Kenya (0.23), Cuba (0.22), Dominican Republic (0.21), Pakistan (0.14), Nigeria (0.13) and Sri Lanka (0.12).
6	Bulgaria (0.97), Kyrgyz Republic (0.97), Latvia (0.97), Hungary (0.45), Lithuania (0.4) and Czech Republic (0.19).
7	Angola (1), Botswana (1), Burundi (1), Central African Republic (1), Chad (1), Congo Republic (1), Djibouti (1), Gambia (1), Guinea-Bissau (1), Madagascar (1), Malawi (1), Mauritania (1), Niger (1), Rwanda (1), Tanzania (1), Togo (1), Zambia (1), Gabon (0.34), Mozambique (0.28), Senegal (0.22), Benin (0.21), Sierra Leone (0.17), Burkina Faso (0.15), Cameroon (0.13), Guinea (0.1) and Mali (0.1).
8	Norway (1)
9	Austria (1), Belgium (1), Denmark (1), Finland (1), France (1), Germany (1), Greece (1), Ireland (1), Italy (1), Luxemburg (1), Netherlands (1), Portugal (1), Spain (1), Sweden (1), United Kingdom (1), Israel (0.11) and Japan (0.1).

specificity and pushed such countries towards the *Average* cluster.

5 Concluding Remarks

In this paper a novel approach based on Fuzzy Set Theory is proposed in order to group WTO members into homogeneous clusters according to their official position on the agricultural issues raised in the Doha Development Agenda negotiations.
The preliminary results suggest that by using Fuzzy Clustering instead of a Crisp Clustering procedure the output partition could reflect more accurately the cross partnerships among the potential allies. Moreover, by implementing a method where missing values are esplicitly handled the risk of stamping out well-defined positions, though on a narrow set of issues, is avoided.
On the other hand greater instability in the membership functions is generated due to the iterative nature of the alghoritm and due to the presence of very uncertain positions. Despite this aspect can be viewed as one of the benefit of the fuzzy classification it makes the whole procedure too much sensitive to the initial random selection of the clusters' prototype. This drawback could be partly faced by estimating fuzzy confidence intervals (Mc83) on the punctual memebership values.
The next research questions that should be addressed concern the use of socio-economical variables at the end of the clustering process for caractherizing each cluster according to the main features of the belonging countries (GDP pro capite, political ideology, form of government etc.) and the implementation of a pattern recognition procedure allowing to rate the positions of WTO members by using the official statements and declarations during and following the WTO meetings.

6 Appendix: R-Code

The FCM algorithm has been already implemented in the R-Language and it is contained in the contributed package *e1071* which also includes functions for latent class analysis, short time Fourier transform and support vector machines. However, the core of the algorithm, namely the *Optimization Phase*, is C-compiled and thus the code cannot be easily modified neither to implement the Optimal Completion Strategy nor to automate the choice of the correct number of clusters. In this section the original R-code written to implement the proposed fuzzy clustering approach is provided.
We assume that the data set is loaded in to R as a `data.frame` named *x*, using the function `read.table()`. The input data set is the only mandatory input together with the minimum and the maximum number of clusters which define the range of solutions within which the optimal number (according to the Partition Entropy Coefficient) is searched for. The reason leading this device is that, in spite of the normalization, the partition Entropy, together with many other cluster validity indexes exhibits a monotonic behavior with respect to the number of clusters. By allowing the user to limit the search space it is possible to keep under control this tendency.

```
OCS_cmeans<-function (x, iter.max<-1000, m<-2, error<-0.001,
                    min.clusters, max.clusters) {
x <- as.matrix(x)
xrows <- nrow(x)
```

```
xcols <- ncol(x)

#Initialize missing features with the column median
if (any(is.na(x)))
 {
   na.index <- is.na(x)
   arrMedian <- apply(x,2,median, na.rm=TRUE)
   f<-array(rep(arrMedian,each=nrow(x)),dim=c(nrow(x),ncol(x)))
   x[is.na(x)]=f[is.na(x)]
 }

#inizialize number of clusters
ncenters <- min.clusters

#inizialize results list
Results=list()

#Outer loop for choosing the correct number of clusters
repeat
 {
   #Initialize prototypes
   centers <- x[sample(1:xrows, ncenters), , drop = FALSE]
   if (any(duplicated(centers)))
    {
      cn <- unique(x)
      mm <- nrow(cn)
      if (mm < ncenters) stop("More cluster centers
                                 than distinct data points.")
      centers <- cn[sample(1:mm, ncenters), , drop = FALSE]
   }
   cont=0

   #Inner loop for optimizing the objective function
   repeat
   {
    old.centers=centers
    cont=cont+1

    #Compute distance from prototypes
    Dis=matrix(1,nrow=xrows,ncol=ncenters)
    for (i in 1:ncenters)
    {
     Dis[1:xrows,i]=apply(sweep(x, 2, centers[i, 1:xcols], EuNorm),1,sum)
    }

    #Update membership matrix
    U= sweep(Dis^(1/(1-m)),1,as.matrix(apply(Dis^(1/(1-m)),1,sum)),"/")
```

```
      U[is.na(U)]=1

      #Update prototypes
      centers=sweep(t(U^m)%*%x,1,as.matrix(apply(U^m,2,sum)),"/")
      new.centers=centers

      #Update missing features
      if (any(is.na(x)))
      {
       Missing=sweep(U^m%*%centers,1,as.matrix(apply(U^m,2,sum)),"/")
      }

      #Inner Stopping rule
      if (max(abs(old.centers-new.centers))<error || cont>=iter.max)
      {

       #Compute the validation index
       PE=-((1/((xrows))*sum((log(U)*U))))/log(ncenters)

       #Obtain the closest hard partition
       Clusters=apply(U,1,Find.Unique.Max)

       #Break the inner loop
       break
      }
     }
    #Combine all the information in the output list
    Results=c(Results,list( Clusters=Clusters,Centers=centers,Distances=Dis,
                            Membership=U,Cont=cont, Index=PE))
    #Outer Stopping rule
    if (ncenters==max.clusters) break
    else ncenters=ncenters+1
}
#find the fuzzy partition with minimum value of Fuzzy Entropy
arr.Fuzzy.Entropy=unlist(Results[seq(6,length(Results),6)])
minimum=which(arr.Fuzzy.Entropy==min(arr.Fuzzy.Entropy))
#output the optimal fuzzy partition
OCS_cmeans<-list(Results[seq(((minimum-1)*6+1),((minimum-1)*6+1)+5)])
}
```

The main function calls two sub-routine, `EuNorm` and `Find.Unique.Max` which serve, respectively, to compute the Euclidean Norm between two vectors and to obtain the hard partition by assigning each unit to the cluster with higher membership (in case of equal membership the assignment is resolved in a random fashion). For reasons of completeness in the following the R-code of the two subroutines is provided.

```
EuNorm <- function (x,y)
{
 EuNorm=(x-y)^2
}

Find.Unique.Max <- function(x)
{
 index=which(x==max(x))
 if (length(index)>1)
 {
  index=index[ceiling(runif(1,1,length(index)))]
 }
 Find.Unique.Max=index
}
```

7 Bibliography

References

[Be81] J. C. Bezdek, 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York.

[Be73] J. C. Bezdek, 1973. *Fuzzy Mathemathics in Pattern Classification*, PhD Thesis, Applied Math. Center, Cornell University, Ithaca.

[BjLi02] c. Bjørnskov and K. M. Lind, 2002. *Where do Developing Countries Go After Doha? An Analysis of WTO Positions and Potential Alliances*, Danish Research Institute of Food Economics, Copenhagen.

[Du73] J.C. Dunn, 1973. *A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters*, Journal of Cybernetics, **3**, 32–57

[GaGe89] I. Gath and A. B. Geva, 1989. *Unsupervised Optimal Fuzzy Clustering*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **11** (7), 773–781.

[HaBe01] R.J. Hathaway and J.C. Bezdek, 2001. *Fuzzy c-Means Clustering of Incomplete Data*, Information and Control, **8** (3), 338–353.

[HaBe02] R.J. Hathaway and J.C. Bezdek, 2002. *Clustering incomplete relational data using the non-Euclidean relational fuzzy c-means algorithm*, Pattern Recognition Letters, **23** (1-3), 151–160.

[Mc83] R. A. McCain, 1983. *Fuzzy confidence intervals*, Fuzzy Sets and Systems, **10**, 281–290.

[Ro82] M. Roubens, 1982. *Fuzzy clustering algorithms and their cluster validity*, European Journal of Operational Research, **10**, 294–301.

[XiBe91] L. X. Xie and G. Beni, 1991. *Validity measure for fuzzy clustering*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **3** (8), 841–847.

[Za65] L.A. Zadeh, 1965. *Fuzzy sets*, Information and Control, **8** (3), 338–353.

[ZhCh03] D.Q. Zhang and S.C. Chen, 2003. *Clustering Incomplete Data Using Kernel-Based Fuzzy C-means Algorithm*, Neural Processing Letters, **18** (3), 155–162.

[Zi91] H. J. Zimmermann, 1991. *Fuzzy set theory and its applications*, Kluwer.