Chapter 1

CONSENSUS-BASED AGGLOMERATIVE HIERARCHICAL CLUSTERING

José Luis García-Lapresta[1] and David Pérez-Román[2]

[1] *PRESAD Research Group, BORDA Research Unit, IMUVA, Dept. de Economía Aplicada, Universidad de Valladolid, Spain*

[1] *PRESAD Research Group, BORDA Research Unit, Dept. de Organización de Empresas y C.I.M., Universidad de Valladolid, Spain*

Abstract In this contribution, we consider that a set of agents assess a set of alternatives through numbers in the unit interval. In this setting, we introduce a measure that assigns a degree of consensus to each subset of agents with respect to every subset of alternatives. This consensus measure is defined as 1 minus the outcome generated by a symmetric aggregation function to the distances between the corresponding individual assessments. We establish some properties of the consensus measure, some of them depending on the used aggregation function. We also introduce an agglomerative hierarchical clustering procedure that is generated by similarity functions based on the previous consensus measures.

Keywords: Consensus, clustering, aggregation functions, OWA operators.

1. Introduction

When a group on agents show their opinions about a set of alternatives, an important issue is to know the homogeneity of these opinions. In this chapter we consider that agents evaluate each alternative by means of a number in the unit interval. For measuring the consensus in a group of agents over a subset of alternatives, we propose to aggregate the distances between the corresponding individual assessments through an appropriate symmetric aggregation function. This outcome measures the dispersion of individual opinions in a similar way to the Gini index [19] measures the inequality of individual incomes (see

Yitzhaki [32]). The consensus measure we propose is just 1 minus the mentioned dispersion measure.

The most important is not to know the degree of consensus in a specific group of agents, but comparing the consensus of different group of agents with respect to an alternative or a subset of alternatives. This is the starting point of the agglomerative hierarchical clustering procedure we propose. We consider as linkage clustering criterion one generated by a consensus-based similarity function that merges clusters or individuals by maximizing the consensus.

The rest of the chapter is organized as follows. Section 2 contains some notation and basic notions. In Section 3, we introduce and analyze the proposed consensus measures. Section 4 contains our proposal of consensus-based agglomerative hierarchical clustering. In Section 5, we illustrate the introduced procedures with an example. Finally, in Section 6, we conclude with some remarks.

2. Preliminaries

Given $\boldsymbol{y}, \boldsymbol{z} \in [0,1]^k$, by $\boldsymbol{y} \geq \boldsymbol{z}$ we mean $y_i \geq z_i$ for every $i \in \{1, \ldots, k\}$. Given $\boldsymbol{y} \in [0,1]^k$, the decreasing reordering of the coordinates of $\boldsymbol{y}$ is indicated as $y_{[1]} \geq \cdots \geq y_{[k]}$. In particular, $y_{[1]} = \max\{y_1, \ldots, y_k\}$ and $y_{[k]} = \min\{y_1, \ldots, y_k\}$.

Given a real number y, by $\lfloor y \rfloor$ we denote the *integer part* of y, i.e., the greatest integer number smaller than or equal to y.

With $\#I$ we denote the cardinality of I. With $\mathscr{P}_2(A) = \{I \subseteq A \mid \#I \geq 2\}$ we denote the family of subsets of at least two elements.

We begin by defining standard properties of real functions on $[0,1]^k$. For further details the interested reader is referred to Fodor and Roubens [12], Calvo *et al.* [6], Beliakov *et al.* [4], Torra and Narukawa [27], Grabisch *et al.* [20] and Beliakov *et al.* [3].

DEFINITION 1.1 *Let $F : [0,1]^k \longrightarrow [0,1]$ be a function.*

1. *F is* idempotent *if for every $y \in [0,1]$ it holds $F(y \cdot \mathbf{1}) = y$.*
2. *F is* symmetric *if for every permutation π on $\{1, \ldots, k\}$ and every $\boldsymbol{y} \in [0,1]^k$ it holds $F(y_{\pi(1)}, \ldots, y_{\pi(k)}) = F(\boldsymbol{y})$.*
3. *F is* monotonic *if for all $\boldsymbol{y}, \boldsymbol{z} \in [0,1]^k$ it holds $\boldsymbol{y} \geq \boldsymbol{z} \Rightarrow F(\boldsymbol{y}) \geq F(\boldsymbol{z})$.*
4. *F is* compensative *if for every $\boldsymbol{y} \in [0,1]^k$ it holds $y_{[k]} \leq F(\boldsymbol{y}) \leq y_{[1]}$.*
5. *F is* self-dual *if for every $\boldsymbol{y} \in [0,1]^k$ it holds $F(\mathbf{1} - \boldsymbol{y}) = 1 - F(\boldsymbol{y})$.*
6. *F is* stable for translations *if for all $\boldsymbol{y} \in [0,1]^k$ and $t \in [0,1]$ such that $\boldsymbol{y} + t \cdot \mathbf{1} \in [0,1]^k$ it holds $F(\boldsymbol{y} + t \cdot \mathbf{1}) = F(\boldsymbol{y}) + t$.*

DEFINITION 1.2

1 *Given* $k \in \mathbb{N}$, *a function* $F^{(k)} : [0,1]^k \longrightarrow [0,1]$ *is called an* k-ary aggregation function *if it is monotonic and satisfies the boundary conditions* $F^{(k)}(\mathbf{0}) = 0$ *and* $F^{(k)}(\mathbf{1}) = 1$. *In the extreme case of* $k = 1$, *the convention* $F^{(1)}(y) = y$ *for every* $y \in [0,1]$ *is considered.*

2 *An* aggregation function *is a sequence* $F = \left(F^{(k)}\right)_{k\in\mathbb{N}}$ *of k-ary aggregation functions.*

3 *An aggregation function* $F = \left(F^{(k)}\right)_{k\in\mathbb{N}}$ *satisfies a property (in particular, those appearing in Definition 1.1) whenever* $F^{(k)}$ *satisfies the same property for every* $k \in \mathbb{N}$.

It is easy to see that for every k-ary aggregation function, idempotency and compensativeness are equivalent.

For the sake of simplicity, the k-arity is omitted whenever it is clear from the context.

An interesting class of aggregation functions is the family of OWA operators, introduced by Yager [29].

A *weighting vector* of dimension k is a vector $\boldsymbol{w} = (w_1, \ldots, w_k) \in [0,1]^k$ such that $\sum_{i=1}^{k} w_i = 1$.

DEFINITION 1.3 *Given a weighting vector* $\boldsymbol{w}$ *of dimension k, the* OWA operator associated with $\boldsymbol{w}$ *is the aggregation function* $F_{\boldsymbol{w}} : [0,1]^k \longrightarrow [0,1]$ *defined as*

$$F_{\boldsymbol{w}}(y_1, \ldots, y_k) = \sum_{i=1}^{k} w_i \cdot y_{[i]}.$$

Some well-known aggregation functions are specific cases of OWA operators. With appropriate weighting vectors $\boldsymbol{w} = (w_1, \ldots, w_k)$ we obtain

1 The *maximum*, for $\boldsymbol{w} = (1, 0, \ldots, 0)$.

2 The *minimum*, for $\boldsymbol{w} = (0, \ldots, 0, 1)$.

3 The *arithmetic mean*, for $\boldsymbol{w} = \left(\frac{1}{k}, \ldots, \frac{1}{k}\right)$.

4 The *t-trimmed means*:

- If $t = 1$, for $\boldsymbol{w} = \left(0, \frac{1}{k-2}, \ldots, \frac{1}{k-2}, 0\right)$.
- If $t = 2$, for $\boldsymbol{w} = \left(0, 0, \frac{1}{k-4}, \ldots, \frac{1}{k-4}, 0, 0\right)$.
- $\ldots$.

5 The *median*:

(a) If k is odd, for $w_i = \begin{cases} 1, & \text{if } i = \frac{k+1}{2}, \\ 0, & \text{otherwise.} \end{cases}$

(b) If k is even, for $w_i = \begin{cases} 0.5, & \text{if } i \in \left\{\frac{k}{2}, \frac{k}{2}+1\right\}, \\ 0, & \text{otherwise.} \end{cases}$

6 The *mid-range*, for $\boldsymbol{w} = (0.5, 0, \ldots, 0, 0.5)$.

OWA operators are continuous, idempotent (hence, compensative), symmetric, and stable for translations. They have been characterized by Fodor *et al.* [11].

Centered OWA operators have been introduced by Yager [31] in order to give "the most weight to the central scores in the argument tuples and less weighting to the extreme values". We now introduce a more general notion than that provided by Yager. It was introduced by García-Lapresta and Martínez-Panero [14].

DEFINITION 1.4 *Given a weighting vector* $\boldsymbol{w}$ *of dimension k, the OWA operator associated with* $\boldsymbol{w}$ *is* centered *if the following two conditions are satisfied:*

1 $w_{k+1-i} = w_i$ *for every* $i \in \{1, \ldots, k\}$.

2 $w_i \leq w_j$ *whenever* $i < j \leq \lfloor \frac{k+1}{2} \rfloor$ *or* $i > j \geq \lfloor \frac{k+1}{2} \rfloor$.

The first condition is equivalent to the property of self-duality (see García-Lapresta and Llamazares [13, Proposition 5]). The second condition is weaker than the original of Yager [31], called *strongly decaying*, that requires strict inequalities $w_i < w_j$.

Yager [31] requires a third condition in the definition of centered OWA operators, *inclusiveness*: $w_i > 0$ for every $i \in \{1, \ldots, k\}$. That condition is very restrictive for our purposes, since it eliminates some interesting OWA operators as median and trimmed means, among others.

DEFINITION 1.5 *An* extended OWA (EOWA) operator *is a sequence of OWA operators* $(F_{\boldsymbol{w}^k})_{k \in \mathbb{N}}$ *with associated weighting vectors* $\boldsymbol{w}^k = (w_1^k, \ldots, w_k^k)$*, one for each dimension* $k \in \mathbb{N}$.

Following Mayor and Calvo [24], Calvo and Mayor [7], Beliakov *et al.* [4, pp. 54-56] and Beliakov *et al.* [3, pp. 73-76]), we can show graphically an EOWA operator as a weighting triangle where the entries in each row add up to one:

$$
\begin{array}{ccccccccc}
 & & & & w_1^1 & & & & \\
 & & & w_1^2 & & w_2^2 & & & \\
 & & w_1^3 & & w_2^3 & & w_3^3 & & \\
 & w_1^4 & & w_2^4 & & w_3^4 & & w_4^4 & \\
w_1^5 & & w_2^5 & & w_3^5 & & w_4^5 & & w_5^5 \\
\cdots\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\cdots
\end{array}
$$

A very useful approach for obtaining the EOWA weights is the functional method introduced by Yager [30, 31]. Given a *BUM function*, i.e., a monotonic function $f : [0,1] \longrightarrow [0,1]$ such that $f(0) = 0$ and $f(1) = 1$, the associated EOWA weights are defined as

$$
w_i^k = f\left(\frac{i}{k}\right) - f\left(\frac{i-1}{k}\right), \; i = 1,\ldots,k. \tag{1}
$$

Yager [31] proposes to generate BUM functions by means of centering functions.

A *centering function* is a function $g : [0,1] \longrightarrow \mathbb{R}$ satisfying the following conditions:

1 $g(x) > 0$ for every $x \in [0,1]$.

2 $g(0.5+x) = g(0.5-x)$ for every $x \in [0,0.5]$.

3 $g(x) < g(y)$ for $x < y \leq 0.5$ and $g(x) < g(y)$ for $x > y \geq 0.5$.

Then, the function $f : [0,1] \longrightarrow [0,1]$ defined as

$$
f(x) = \frac{\displaystyle\int_0^x g(y)\,dy}{\displaystyle\int_0^1 g(y)\,dy} \tag{2}
$$

is a BUM function.

3. Consensus

For measuring the degree of consensus among a group of agents that provide their opinions on a set of alternatives, different proposals can be found in the literature (see Martínez-Panero [23] for an overview of different notions of consensus).

In the social choice framework, the notion of *consensus measure* was introduced by Bosch [5] in the context of linear orders. Additionally, Bosch [5] and Alcalde-Unzu and Vorsatz [1] provided axiomatic characterizations of several consensus measures in the context of linear orders. García-Lapresta and Pérez-Román [15] extended that notion to the context of weak orders and they analyzed a class of consensus measures generated by distances. Alcantud *et al.* [2] provided axiomatic characterizations of some consensus measures in the setting of approval voting. In turn, Erdamar *et al.* [8] extended the notion of consensus measure to the preference-approval setting through different kinds of distances, and García-Lapresta *et al.* [18] introduced another extension to the framework of hesitant linguistic assessments.

Let $A = \{1, \ldots, m\}$, with $m \geq 2$, be a set of agents and let $X = \{x_1, \ldots, x_n\}$, with $n \geq 2$, be the set of alternatives which have to be evaluated in the unit interval.

A *profile* is a matrix

$$V = \begin{pmatrix} v_1^1 & \cdots & v_i^1 & \cdots & v_n^1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^a & \cdots & v_i^a & \cdots & v_n^a \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^m & \cdots & v_i^m & \cdots & v_n^m \end{pmatrix} = (v_i^a)$$

consisting of m rows and n columns of numbers in $[0,1]$, where the element v_i^a represents the assessment given by the agent $a \in A$ to the alternative $x_i \in X$.

Let $V = (v_i^a)$ be a profile, π a permutation on A, σ a permutation on $\{1, \ldots, n\}$, $I \in \mathscr{P}_2(A)$ and $\emptyset \neq Y \subseteq X$. The profiles V^π, V_σ and V^{-1}, and the subsets I^π and Y_σ are defined as follows:

1 $V^\pi = (u_i^a)$ where $u_i^a = v_i^{\pi(a)}$.

2 $V_\sigma = (u_i^a)$ where $u_i^a = v_{\sigma(i)}^a$.

3 $V^{-1} = (u_i^a)$ where $u_i^a = 1 - v_i^a$.

4 $I^\pi = \left\{\pi^{-1}(a) \mid a \in A\right\}$, i.e., $a \in I^\pi \Leftrightarrow \pi(a) \in I$.

5 $Y_\sigma = \{x_{\sigma^{-1}(i)} \mid x_i \in Y\}$, i.e., $x_i \in Y_\sigma \Leftrightarrow x_{\sigma(i)} \in Y$.

We now introduce a consensus measure associated with a symmetric aggregation function. Given a profile, it assigns a degree of consensus in each subset of at least two agents with respect to a subset of alternatives.

DEFINITION 1.6 *Let $F = \left(F^{(k)}\right)_{k \in \mathbb{N}}$ be a symmetric aggregation function. Given a profile $V = (v_i^a)$, the* degree of consensus *in a subset of agents $I \in \mathscr{P}_2(A)$ over a subset of alternatives $\emptyset \neq Y \subseteq X$ is defined as*

$$C_F(V, I, Y) = 1 - F\left(\left| v_i^a - v_i^b \right|_{\substack{a,b \in I,\, a<b \\ x_i \in Y}} \right).$$

In Proposition 1.1 we establish some properties of the consensus notion introduced in Definition 1.6. Normalization means that the degree of consensus is always in the unit interval. Anonymity means that all agents are treated in the same way. Unanimity establishes necessary and sufficient conditions for reaching maximum consensus. Maximum dissension establishes necessary and sufficient conditions for reaching minimum consensus in two agents. Positiveness establishes that with more than two agents the degree of consensus is never minimum. Neutrality means that all alternatives are treated in the same way. And reciprocity means that if all the agents reverse their assessments, then the degree of consensus does not change.

PROPOSITION 1.1 *Let $F = \left(F^{(k)}\right)_{k\in\mathbb{N}}$ be a symmetric aggregation function. The following properties are satisfied:*

1 Normalization*: $C_F(V,I,Y) \in [0,1]$.*

2 Anonymity*: $C_F(V^\pi,I^\pi,Y) = C_F(V,I,Y)$ for every permutation π on A.*

3 Unanimity*: If for every $x_i \in Y$ there exists $t_i \in [0,1]$ such that $v_i^a = t_i$ for every $a \in I$, then $C_F(V,I,Y) = 1$.*

Additionally, if $F^{(k)}(\boldsymbol{y}) = 0 \Leftrightarrow \boldsymbol{y} = \boldsymbol{0}$, for all $k \in \mathbb{N}$ and $\boldsymbol{y} \in [0,1]^k$, and $C_F(V,I,Y) = 1$, then for every $x_i \in Y$ there exists $t_i \in [0,1]$ such that $v_i^a = t_i$ for every $a \in I$.

4 Maximum dissension*: If $\left(\left(v_i^a = 0 \text{ and } v_i^b = 1\right) \text{ or } \left(v_i^a = 1 \text{ and } v_i^b = 0\right)\right)$ for all $x_i \in Y$, then $C_F(V,\{a,b\},Y) = 0$.*

Additionally, if $F^{(k)}(\boldsymbol{y}) = 1 \Leftrightarrow \boldsymbol{y} = \boldsymbol{1}$, for all $k \in \mathbb{N}$ and $\boldsymbol{y} \in [0,1]^k$, and $C_F(V,\{a,b\},Y) = 0$, then $\left(\left(v_i^a = 0 \text{ and } v_i^b = 1\right) \text{ or } \left(v_i^a = 1 \text{ and } v_i^b = 0\right)\right)$ for all $x_i \in Y$.

5 Positiveness*: If $F^{(k)}(\boldsymbol{y}) = 1 \Leftrightarrow \boldsymbol{y} = \boldsymbol{1}$, for all $k \in \mathbb{N}$ and $\boldsymbol{y} \in [0,1]^k$, and $\#I > 2$, then $C_F(V,I,Y) > 0$.*

6 Neutrality*: $C_F(V_\sigma,I,Y_\sigma) = C_F(V,I,Y)$ for every permutation σ on $\{1,\ldots,n\}$.*

7 Reciprocity*: $C_F(V^{-1},I,Y) = C_F(V,I,Y)$.*

PROOF: It is straightforward.

REMARK 1.1 Let $(F_{\boldsymbol{w}^k})_{k\in\mathbb{N}}$ an EOWA operator with associated weighting vectors $\boldsymbol{w}^k = (w_1^k,\ldots,w_k^k)$, $k \in \mathbb{N}$. It is easy to check that $F_{\boldsymbol{w}^k}(\boldsymbol{y}) = 0 \Leftrightarrow \boldsymbol{y} = \boldsymbol{0}$, for every $\boldsymbol{y} \in [0,1]^k$, if and only if $w_1^k > 0$; and $F_{\boldsymbol{w}^k}(\boldsymbol{y}) = 1 \Leftrightarrow \boldsymbol{y} = \boldsymbol{1}$, for every $\boldsymbol{y} \in [0,1]^k$, if and only if $w_k^k > 0$.

Consequently, any EOWA operator satisfying $w_1^k > 0$ and $w_k^k > 0$ for every $k \in \mathbb{N}$ verifies all the properties included in Proposition 1.1. Therefore, when considering the EOWA operators generated by the maximum, the minimum, the trimmed means and the median, the corresponding consensus measures do not satisfy the strong versions of unanimity and maximum dissension.

For our purposes, an interesting class of EOWA operators is the one generated by centered OWA operators (in the sense of Definition 1.4) satisfying $w_1^k = w_k^k > 0$ for every $k \in \mathbb{N}$.

4. Clustering

There are many clustering algorithms (see Ward [28], Jain *et al.* [21] and Everitt *et al.* [9], among others). Most methods of hierarchical clustering use an appropriate metric (for measuring the distance between pairs of observations), and a linkage criterion which specifies the similarity/dissimilarity of sets as a function of the pairwise distances of observations in the corresponding sets.

Ward [28] proposed an agglomerative hierarchical clustering procedure, where the criterion for choosing the pair of clusters to merge at each step is based on the optimization of an objective function.

Usually, clusters are merged by minimizing a distance between clusters. The complete, single and average linkage clustering take into account the maximum, minimum and mean distance between elements of each cluster, respectively. In turn, centroid linkage clustering is based on the distances between the clusters centroids.

In all the mentioned linkage clustering criteria there is a loss of information. In our proposal, clusters are merged when maximizing the consensus and, consequently, all the information is used for merging clusters.

Definition 1.7 *Let* $F = \left(F^{(k)}\right)_{k\in\mathbb{N}}$ *be a symmetric aggregation function. Given a profile* $V = (v_i^a)$*, the* similarity function *relative to a subset of alternatives* $\emptyset \neq Y \subseteq X$

$$S_F^Y : \left(\mathscr{P}(A) \setminus \{\emptyset\}\right)^2 \longrightarrow [0,1]$$

is defined as

$$S_F^Y(I,J) = \begin{cases} C_F(V, I \cup J, Y), & \textit{if } \#(I \cup J) \geq 2, \\ 1, & \textit{if } \#(I \cup J) = 1. \end{cases}$$

Remark 1.2 In the extreme case of two agents and a single alternative, the similarity between these agents on that alternative is just 1 minus the distance between their assessments. More formally, given an alternative $x_i \in X$ and two different agents $a, b \in A$, we have

$$S_F^{\{x_i\}}\left(\{a\},\{b\}\right) = C_F\left(V,\{a,b\},\{x_i\}\right) = 1 - \left|v_i^a - v_i^b\right|.$$

The agglomerative hierarchical clustering procedure we propose has some similarities to the ones provided by García-Lapresta and Pérez-Román [16, 17], in different settings. Given an aggregation function $F = \left(F^{(k)}\right)_{k\in\mathbb{N}}$ and a profile $V = (v_i^a)$, our proposal consists of a sequential process addressed by the following stages:

1 The initial clustering is $\mathscr{A}_0^Y = \{\{1\},\ldots,\{m\}\}$.

2 Calculate the similarities between all the pairs of agents, $S_F^Y(\{a\},\{b\})$ for all $a,b \in A$.

3 Select the two agents $a,b \in A$ that maximize S_F^Y and construct the first cluster $A_1^Y = \{a,b\}$.

4 The new clustering is $\mathscr{A}_1^Y = \left(\mathscr{A}_0^Y \setminus \{\{a\},\{b\}\}\right) \cup \left\{A_1^Y\right\}$.

5 Calculate the similarities $S_F^Y(A_1^Y,\{c\})$ and take into account the previously computed similarities $S_Y(\{c\},\{d\})$, for all $\{c\},\{d\} \in \mathscr{A}_1^Y$.

6 Select the two elements of $\mathscr{A}_1^Y$ that maximize S_F^Y and construct the second cluster A_2^i.

7 Proceed as in previous items until obtaining the next clustering $\mathscr{A}_2^i$.

The process continues in the same way until obtaining the last cluster, $\mathscr{A}_{m-1}^Y = \{A\}$.

In the case of several pairs of agents or clusters are in a tie, then proceed in a lexicographic manner in $1,\ldots,m$.

5. An illustrative example

In order to illustrate the agglomerative hierarchical clustering procedure introduced in Section 4, consider a set of eight experts $A = \{1,2,3,4,5,6,7,8\}$ assessing a set of six alternatives $X = \{x_1,x_2,x_3,x_4,x_5,x_6\}$ through the following profile

$$V = \begin{pmatrix} 1.0 & 0.9 & 0.7 & 0.5 & 0.5 & 0.0 & 0.5 & 1.0 \\ 0.8 & 0.0 & 0.6 & 0.4 & 0.3 & 0.8 & 1.0 & 0.8 \\ 0.6 & 0.6 & 0.6 & 1.0 & 0.2 & 0.6 & 0.7 & 1.0 \\ 0.4 & 0.4 & 0.4 & 0.3 & 0.9 & 0.7 & 0.3 & 0.7 \\ 0.3 & 0.3 & 0.7 & 1.0 & 0.0 & 0.9 & 0.2 & 1.0 \\ 0.0 & 1.0 & 0.5 & 1.0 & 0.5 & 0.7 & 1.0 & 0.7 \end{pmatrix}.$$

In order to show the importance of the aggregation function for defining the consensus measure that generates the cluster formation, we have considered four centered EOWA operators (in the sense of Definition 1.4): the arithmetic mean, the 1-trimmed mean (or olympic EOWA operator) and two specific cases generated by the functional method introduced by Yager [30, 31].

The clustering processes have been carried out for the case of all the alternatives ($Y = X$). The outcomes are summarized in the corresponding dendrograms.

The dendrograms generated by the arithmetic mean and the 1-trimmed mean are shown in Figure 1.1 and Figure 1.2, respectively.

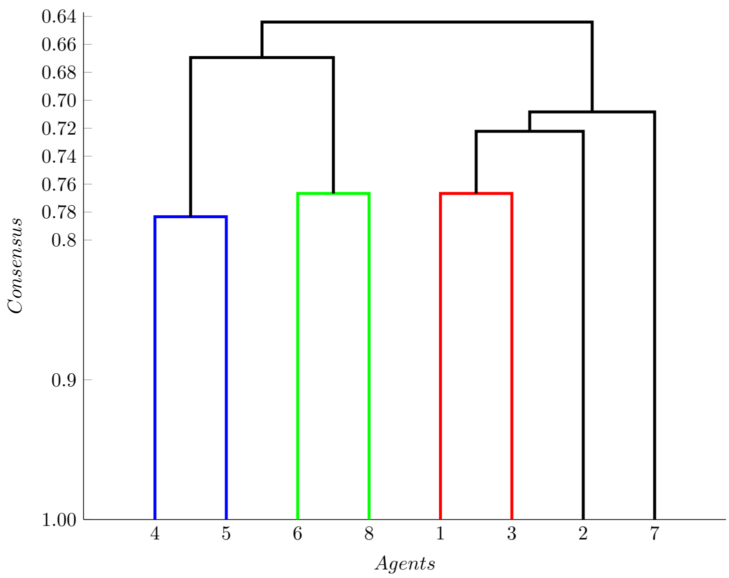

Figure 1.1. Dendrogram obtained with the arithmetic mean.

Figure 1.3 shows the dendrogram that corresponds to consider the EOWA operator whose weights are given by applying Equation (1) to the BUM function generated by Equation (2) with the piecewise linear centering function g_1 defined as

$$g_1(x) = \begin{cases} 2x, & \text{if } 0 \le x \le 0.5, \\ 2-2x, & \text{if } 0.5 \le x \le 1. \end{cases}$$

In this case, the weights are $w_i^k = \dfrac{2(2i-1)}{k^2}$ for $i \le \dfrac{k+1}{2}$, and $w_i^k = w_{k+1-i}$ if $i \ge \dfrac{k+1}{2}$.

Similarly, Figure 1.4 shows the dendrogram that corresponds to consider the EOWA operator whose weights are given by applying Equation (1) to the BUM function generated by Equation (2) with the parabolic centering function

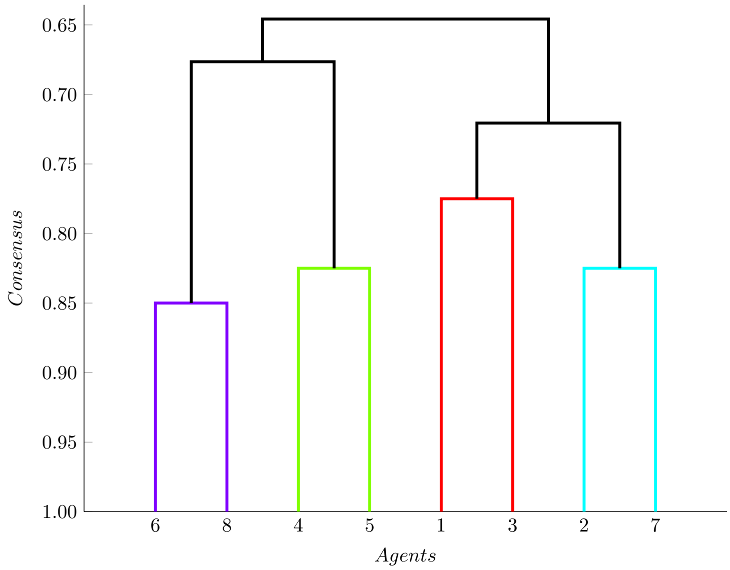

Figure 1.2. Dendrogram obtained with the 1-trimmed mean.

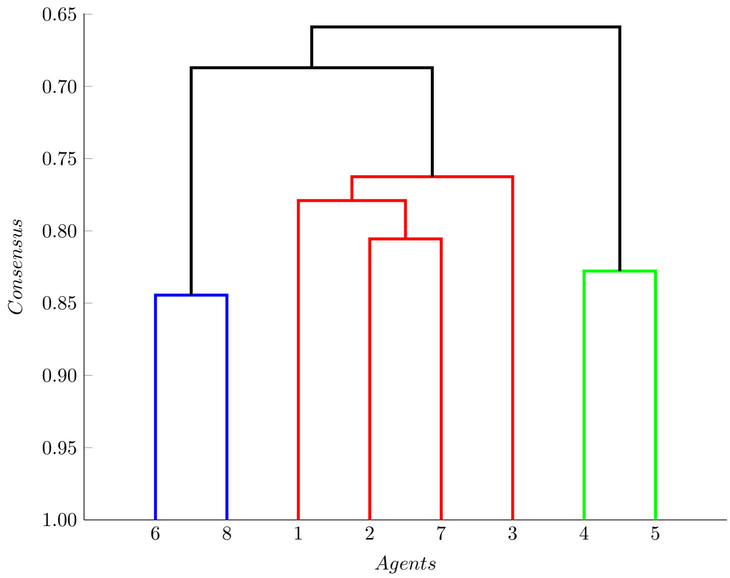

Figure 1.3. Dendrogram obtained with the EOWA operator generated by g_1.

$g_2(x) = 4(x - x^2)$. Now the weights are $w_i^k = \dfrac{3k(2i-1) - 6i(i-1) - 2}{k^3}$ for $i \leq \dfrac{k+1}{2}$, and $w_i^k = w_{k+1-i}$ if $i \geq \dfrac{k+1}{2}$.

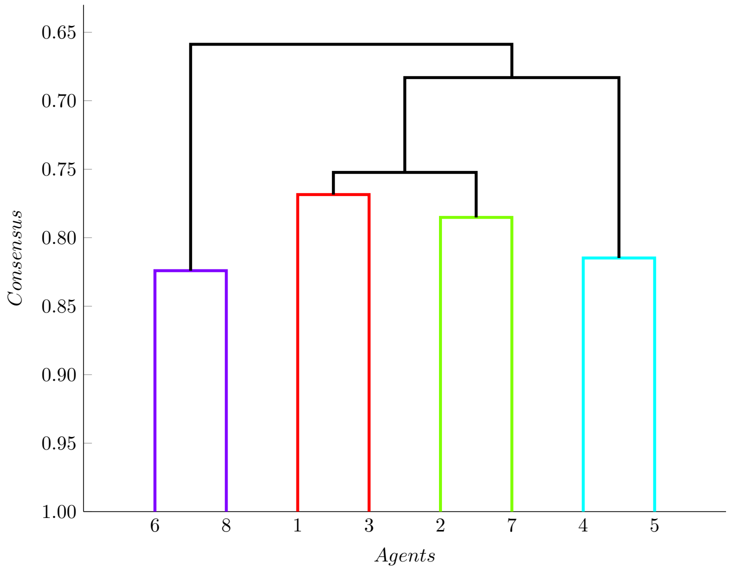

Figure 1.4. Dendrogram obtained with the EOWA operator generated by g_2.

6. Concluding remarks

When a group of agents show their opinions about a set of alternatives, an interesting problem is to know what is the consensus in the whole group or in a subset of agents with respect to one or several alternatives. The most important is not to know the corresponding degrees of consensus, but to compare the consensus in different subsets of agents and alternatives. With our proposal, this information can be easily achieved. Even more, the proposed consensus-based clustering and the corresponding dendrograms provide a rich and visual picture of the homogeneity in the individual opinions.

The mentioned consensus and clustering procedures are static. However, in consensus reaching processes the degree of consensus in a specific situation is only the starting point of a dynamic and iterative process that pursues to increase the agreement among agents. A consensus reaching process consists of several rounds where a human or virtual moderator may invite some agents to modify their opinions in order to increase the collective agreement (see Fedrizzi *et al.* [10], Saint and Lawson [26], Martínez and Montero [22] and Palomares *et al.* [25], among others).

These consensus reaching processes and the corresponding clustering analyses can be carried out in the setting of this contribution. The fact that the proposed consensus measure is associated with an aggregation function provides flexibility to the process. Once are determined the alternatives x_i where the degree of consensus in the whole group of agents, $C_F(V,A,\{x_i\})$, is smaller than the overall degree of consensus $C_F(V,A,X)$, the moderator may invite those agents whose opinions over x_i are quite different to the median assessment to properly modify their assessments. If these agents move their assessments on the selected alternatives towards the corresponding median assessments, then the degree of consensus increases. Due to the monotonicity of the aggregation function, the overall degree of consensus increases as well. All these changes can be visualized through the corresponding dendrograms.

Acknowledgments

The authors gratefully acknowledge the funding support of the Spanish *Ministerio de Economía y Competitividad* (project ECO2012-32178) and *Consejería de Educación de la Junta de Castilla y León* (project VA066U13).

References

[1] Alcalde-Unzu, J., Vorsatz, M. (2013). Measuring the cohesiveness of preferences: an axiomatic analysis. *Social Choice and Welfare* 41, pp. 965-988.

[2] Alcantud, J.C.R., de Andrés, R., Cascón, J.M. (2013). On measures of cohesiveness under dichotomous opinions: some characterizations of Approval Consensus Measures. *Information Sciences* 240, pp. 45-55.

[3] Beliakov, G., Bustince Sola, H., Calvo Sánchez, T. (2016). *A Practical Guide to Averaging Functions*. Springer, Heidelberg.

[4] Beliakov, G., Pradera, A., Calvo, T. (2007). *Aggregation Functions: A Guide for Practitioners*. Springer, Heidelberg.

[5] Bosch, R. (2005). *Characterizations of Voting Rules and Consensus Measures*. Ph. D. Dissertation, Tilburg University.

[6] Calvo, T., Kolesárova, A., Komorníková, M., Mesiar, R. (2002). Aggregation operators: Properties, classes and construction methods. In: T. Calvo, G. Mayor, R. Mesiar (eds) *Aggregation Operators: New Trends and Applications*, pp. 3-104. Physica-Verlag, Heidelberg.

[7] Calvo, T., Mayor, G. (1999). Remarks on two types of extended aggregation functions. *Tatra Mountains Mathematical Publications* 16, pp. 235-253.

[8] Erdamar, B., García-Lapresta, J.L., Pérez-Román, D., Sanver, M.R. (2014). Measuring consensus in a preference-approval context. *Information Fusion* 17, pp. 14-21.

[9] Everitt, B.S., Landau, S., Leese, M. (2001). *Cluster Analysis*, 4th Edition. Oxford University Press, New York.

[10] Fedrizzi, M., Kacprzyk, J., Owsińnski, J.W., Zadrożny, S. (1994). Consensus reaching via a GDSS with fuzzy majority and clustering of preference profiles. *Annals of Operations Research* 51, pp. 127-139.

[11] Fodor, J., Marichal, J.L., Roubens, M. (1995). Characterization of the ordered weighted averaging operators. *IEEE Transtactions on Fuzzy Systems* 3, pp. 236-240.

[12] Fodor, J., Roubens, M. (1994). *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer Academic Publishers, Dordrecht.

[13] García-Lapresta, J.L., Llamazares, B. (2001). Majority decisions based on difference of votes. *Journal of Mathematical Economics* 45, pp. 463-481.

[14] García-Lapresta, J.L., Martínez-Panero, M. (2009). Linguistic-based voting through centered OWA operators. *Fuzzy Optimization and Decision Making* 8, pp. 381-393.

[15] García-Lapresta, J.L., Pérez-Román, D. (2011). Measuring consensus in weak orders. In: E. Herrera-Viedma, J.L. García-Lapresta, J. Kacprzyk, H. Nurmi, M. Fedrizzi, S. Zadrożny (eds.), *Consensual Processes*, STUDFUZZ, vol. 267. Springer-Verlag, Berlin, pp. 213-234.

[16] García-Lapresta, J.L., Pérez-Román, D. (2015). Ordinal proximity measures in the context of unbalanced qualitative scales and some applications to consensus and clustering. *Applied Soft Computing* 35, pp. 864-872.

[17] García-Lapresta, J.L., Pérez-Román, D. (2016). Consensus-based clustering under hesitant qualitative assessments. *Fuzzy Sets and Systems* 292, pp. 261-273.

[18] García-Lapresta, J.L., Pérez-Román, D., Falcó, E. (2014). Consensus reaching processes under hesitant linguistic assessments. In: P. Angelov et al. (eds.), *Intelligent Systems'2014. Advances in Intelligent Systems and Computing* 322, pp. 257-268.

[19] Gini, C. (1912). *Variabilità e Mutabilità*. Tipografia di Paolo Cuppini, Bologna.

[20] Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E. (2009). *Aggregation Functions*. Cambridge University Press, Cambridge.

[21] Jain, A.K., Murty, M.N., Flynn, P.J. (1999). Data clustering: A review. *ACM Computing Surveys* 31 (3), pp. 264-323.

[22] Martínez, L., Montero, J. (2007). Challenges for improving consensus reaching process in collective decisions. *New Mathematics and Natural Computation* 3, pp. 203-217.

[23] Martínez-Panero, M. (2011). Consensus perspectives: Glimpses into theoretical advances and applications. In: E. Herrera-Viedma, J.L. García-Lapresta, J. Kacprzyk, H. Nurmi, M. Fedrizzi, S. Zadrożny (eds.), *Consensual Processes*, STUDFUZZ, vol. 267, Springer-Verlag, Berlin, pp. 179-193.

[24] Mayor, G., Calvo, T. (1997). On extended aggregation functions. *Proceedings of IFSA 97*, vol. I, Prague, pp. 281-285.

[25] Palomares, I., Estrella, F.J., Martínez, L., Herrera, F. (2014). Consensus under a fuzzy context: taxonomy, analysis framework AFRYCA and experimental case of study. *Information Fusion* 20, pp. 252-271.

[26] Saint, S., Lawson, J.R. (1994). *Rules for Reaching Consensus. A Modern Approach to Decision Making*. Jossey-Bass, San Francisco.

[27] Torra, V., Narukawa, Y. (2007). *Modeling Decisions: Information Fusion and Aggregation Operators*. Springer, Berlin.

[28] Ward Jr., J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58, pp. 236-244.

[29] Yager, R.R. (1988). On ordered weighted averaging operators in multicriteria decision making. *IEEE Transactions on Systems Man and Cybernetics* 8, pp. 183-190.

[30] Yager, R.R. (1996). Quantifier guided aggregation using OWA operators. *International Journal of Intelligent Systems* 11, pp. 49-73.

[31] Yager, R.R. (2007). Centered OWA operators. *Soft Computing* 11, pp. 631-639.

[32] Yitzhaki, S. (1998). More than a dozen alternative ways of spelling Gini. *Research on Economic Inequality* 8, pp. 13-30.