Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 292 (2016) 261–273
FUZZY sets and systems
www.elsevier.com/locate/fss

Consensus-based clustering under hesitant qualitative assessments

José Luis García-Lapresta [a,*], David Pérez-Román [b]

[a] PRESAD Research Group, IMUVA, Dept. de Economía Aplicada, Universidad de Valladolid, Spain
[b] PRESAD Research Group, Dept. de Organización de Empresas y Comercialización e Investigación de Mercados, Universidad de Valladolid, Spain

Received 20 January 2014; received in revised form 7 April 2014; accepted 12 May 2014

Available online 20 May 2014

Abstract

In this paper, we consider that agents judge the feasible alternatives through linguistic terms – when they are confident in their opinions – or linguistic expressions formed by several consecutive linguistic terms – when they hesitate. In this context, we propose an agglomerative hierarchical clustering process where the clusters of agents are generated by using a distance-based consensus measure.

Keywords: Clustering; Consensus; Linguistic assessments

1. Introduction

In different decision-making problems, agents show their opinions about a set of alternatives and then an aggregation procedure generates an outcome (a winning alternative, several winning alternatives, a ranking on the set of alternatives, etc.). The opinions given by the agents may be provided in different ways: the favorite alternative, a subset of acceptable alternatives, a ranking on the set of alternatives, an assessment for each alternative, etc.

In the case of agents assess independently each alternative, the corresponding assessments can be of different nature depending of the context: numerical values, intervals of real numbers, fuzzy numbers, linguistic terms, etc.

Usually people prefer to handle the imprecision through linguistic terms rather than with exact numerical values. Since opinions are imprecise, trying to represent them by using a precise number is meaningless (see Zimmer [38]). Wallsten et al. [34] have shown empirically how most people are more comfortable using words rather than numbers to describe probabilities. On the other hand, for evaluating qualitative aspects, as the comfort of a car, is more appropriate to use linguistic terms than quantitative values (see Levrat et al. [23]).

In the context of voting theory, it is worth mentioning that Balinski and Laraki [4,5] have proposed a voting system called *Majority Judgment* where voters assess the alternatives through linguistic terms (‘excellent’, ‘very good’, ‘good’, ‘acceptable’, ‘poor’ and ‘to reject’). See Balinski and Laraki [6] for an experimental analysis of their proposal.

* Corresponding author.
E-mail addresses: lapresta@eco.uva.es (J.L. García-Lapresta), david@emp.uva.es (D. Pérez-Román).

http://dx.doi.org/10.1016/j.fss.2014.05.004

Clearly, the use of linguistic assessments allows the agents to show their imprecise opinions in a suitable way (see Zadeh [36,37]). However, in some situations agents, even experts assessing alternatives within their field of expertise may have doubts about which feasible linguistic term would assign to an alternative (see Agell et al. [2] and García-Lapresta et al. [16]). For this reason, it is interesting to allow the agents to judge in a more imprecise way, giving them the option of assigning several linguistic terms.[1]

Our approach concerning the imprecision is based on an adaptation of the *absolute order of magnitude spaces* introduced by Travé-Massuyès and Dague [31] and Travé-Massuyès and Piera [32]; more specifically in the extensions devised by Roselló et al. [25–27] (see also Agell et al. [1] and Falcó et al. [13,14]).

In many contexts and disciplines, observations or objects are grouped in clusters in such a way that elements within each cluster are similar to one another with respect to an attribute. Then, objects are classified in homogeneous clusters and the objects in a cluster are more similar to each other than they are to an object belonging to a different cluster (see Jain et al. [21] and Everitt et al. [12], among others).

In this paper, we have devised a clustering procedure in the context of hesitant qualitative assessments.[2] More specifically, we have analyzed how agents can be grouped into clusters when such agents rate alternatives by means of linguistic terms from a predetermined linguistic scale,[3] if the agents are confident in their opinions. If they are not confident about which term to use, they are allowed to use a linguistic expression generated by several consecutive linguistic terms.

In the cluster formation, we have considered that the similarity between two groups of agents with respect to an alternative is the degree of consensus[4] in the merged group. That consensus is measured taking into account the distances between all the pairs of individual assessments over the alternative that is been evaluated (according to a previously fixed metric on the set of linguistic expressions).

Given an alternative, the agglomerative hierarchical clustering process starts by joining the two agents that are more similar with respect to that alternative (ties are broken in a lexicographic manner). The next cluster is created by merging a new agent to the previous cluster or by joining two agents in such a way that the similarity is maximized (again ties are broken in a lexicographic manner). The process continues until the last cluster join all the agents.

The overall agglomerative hierarchical clustering process follows the same pattern that the previous one, but now taking all the alternatives into account. The (overall) degree of consensus in a group of agents is defined as the outcome given by an aggregation function to the degrees of consensus in that group of agents with respect to all the alternatives. The overall similarity between two groups of agents is defined as the (overall) degree of consensus in the merged group.

We have proven some properties that both consensus measures satisfy. These properties ensure that the similarities that generate the clustering processes work in a suitable way.

The information provided by the clustering procedure is interesting for knowing what is the agreement of the agents over each alternative and also what is the overall agreement. It is also useful for having a picture on what agents are more similar regarding a specific alternative as well as from an overall perspective. In this way, possible outliers are easily detected.

A clear application of our proposal is in consensus reaching processes, where a human or virtual moderator invites the agents to modify their opinions for increasing the agreement before applying a decision-making procedure (see, for instance, Saint and Lawson [28] and Eklund et al. [9]).

In order to show how our proposal works, we illustrate the procedure from the data of the field experiment carried out by Agell et al. [2].

---

[1] In this way, Torra [29] introduced the notion of *hesitant fuzzy set* as a generalization of the intuitionistic fuzzy sets and fuzzy multisets, by allowing the agents to assign several values for the membership function to each alternative.

[2] Other clustering procedures have been used in decision-making. For instance, Valls and Torra [33] proposed one for grouping alternatives taking into account the assessments obtained in different criteria; García-Lapresta and Pérez-Román [18] proposed another one for grouping individuals who rank alternatives through weak orders.

[3] We have assumed that the linguistic scale is uniformly and symmetrically distributed. Thus, the distance between consecutive linguistic terms is assumed to be constant.

[4] The notion of *consensus measure* was introduced by Bosch [8] in the context of linear orders. Additionally, Bosch [8] and Alcalde-Unzu and Vorsatz [3] provided axiomatic characterizations of several consensus measures in the context of linear orders. García-Lapresta and Pérez-Román [17] extended that notion to the context of weak orders and they analyzed a class of consensus measures generated by distances.