

Fuzzy Equivalence Relations as Similarity Measure in Agglomerative Clustering Algorithm

Valerijs Mihailovs^[0000-0001-6943-493X]

University of Latvia,
3 Jelgavas street, Riga, LV-1004, Latvia
tykon@inbox.lv

Abstract. In the clustering process, one of the important measures for grouping data into different clusters is the similarity and dissimilarity of two objects. The similarity of two objects is inversely proportional to dissimilarity and can be expressed via the distance between them in various ways:

$$s(x, y) = e^{-d(x, y)}, \quad s(x, y) = \frac{1}{1 + d(x, y)}.$$

Such similarity measures are specific instances of fuzzy equivalence relations, particularly those where transitivity holds for the product and Hamacher t-norm.

This paper studies hierarchical clustering, particularly agglomerative clustering, where fuzzy equivalence relations are considered as a measure of the similarity of objects and clusters, in which transitivity is defined on the basis of various t-norms: Lukasiewicz, product, and Hamacher's t-norm. In construction of fuzzy equivalence relation is involved a tool called an additive generator. For objects and clusters internal and external similarities are introduced using aggregation of corresponding equivalence relations. In the paper, a quality measure of the cluster system is presented using fuzzy equivalence relations. Additionally, the paper provides a comparative analysis between standard clustering methods (K-Means, C-Means) and the proposed approach, with fuzzy equivalence relations, where transitivity is defined on the basis of t-norms.

Keywords: Hierarchical clustering · T-norms · Additive generator · Aggregation of fuzzy equivalence relations.