Model-based clustering: an approach based on copulas

Dimitris Karlis

Department of Statistics, Athens University of Economics and Business

email: karlis@aueb,gr

Abstract

The majority of model-based clustering techniques is based on multivariate Gaussian models and their variants. This talk introduces and studies the framework of copula-based finite mixture models for clustering applications. In particular, the use of copulas in model-based clustering offers several advantages over current methods including the ability to obtain a range of shapes for the clusters, the choice of very flexible marginal models, the ability to jointly model different types of data, just to name few.

The talk presents the basic ideas and focus on issues of estimation and usage for different types of data. Estimation in the general case can be performed using standard EM, and, depending on the mode of the data, more efficient procedures can be used that can fully exploit the copula structure. The exposition of the methodology will be accompanied by the analysis of real and artificial data.

(Joint work with Ioannis Kosmidis, University College London)

Keywords: copulas; clustering; mixed mode data; parametric rotations; multivariate discrete data; mixed domain data;

Reference

Kosmidis, I. and Karlis, D. (2016) **Model-based clustering using copulas with applications.** To appear in *Statistics and Computing*