Seminar

MODEL-BASED CLUSTERING USING COPULAS WITH APPLICATIONS

IOANNIS KOSMIDIS
UNIVERSITY COLLEGE LONDON (UK)

March 17, 2015
12.00
Aula Cucconi

Abstract
www.stat.unipd.it/fera-ricerca/seminari
Abstract

The majority of model-based clustering techniques is based on multivariate Normal 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 two direct advantages over current methods:

i) the appropriate choice of copulas provides the ability to obtain a range of exotic shapes for the clusters, and

ii) the explicit choice of marginal distributions for the clusters allows the modelling of multivariate data of various models (discrete, continuous, both discrete and continuous) in a natural way.

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 closure properties of the mixture models under marginalization will be discussed, and for continuous, real-valued data parametric rotations in the sample space will be introduced, with a parallel discussion on parameter identifiability depending on the choice of copulas for the components. The exposition of the methodology will be accompanied by the analysis of real and artificial data.