PhD program in Statistics
DSS Statistics Seminar
March 24, 2023, 12:00
https://uniroma1.zoom.us/j/86881977368?pwd=SWRFcVFjMDZTa0IXZk05TE1zNm5adz09
Passcode: 432940

A long history of model based clustering based on trimming and constraints

Agustín Mayo-Iscar

University of Valladolid

Model based clustering plays a major role in data analysis. Our interest focuses on approaches related to maximum likelihood estimation via EM/CEM algorithms. However, it is very common that input-datasets contain observations belonging to contaminating sources, out of the assumed family of distributions in the chosen model. It is well known that this contamination in the sample is able to break likelihood based estimators. Methodology based on the joint application of trimming and constraints, under the label TCLUST, has been developed for robustifying model based clustering proposals. Trimming tries to eliminate contaminating observations, however in order to achieve robust proposals in clustering, it is also needed to apply constraints to control the relative size of clusters' variability. There are TCLUST procedures available for estimating mixtures in different settings: linear models, factor analyzers and functional data among others. Statistical properties of TCLUST procedures, including consistency and a non negligible breakdown point are available. TCLUST's constraints have evolved in the last few years, providing an improved flexibility, in order to capture the patterns in the covariance matrix decomposition included in the classical parsimonious family of Celeux and Govaert. An important open issue in TCLUST procedures is related to their input parameters: the number of clusters, the level of trimming and the strength of the constraints. Exploratory tools and automatized procedures for assisting the users in choosing these input parameters have been developed. TCLUST procedures are available in CRAN ('tclust' package) and in MATLAB ('FSDA' toolbox).

SAPIENZA
UNIVERSITÀ DI ROMA

DSS - Dipartimento di Scienze Statistiche - www.dss.uniroma1.it