



dmata

universidade de aveiro
departamento de matemática

CIDMA

CENTRO DE I&D EM MATEMÁTICA E APLICAÇÕES
CENTER FOR R&D IN MATHEMATICS AND
APPLICATIONS

Seminário

Grupo de Probabilidades e Estatística

11 de dezembro de 2019 15:45

Sala 11.3.21

CDPCA for clustering of variables

Adelaide Freitas

Department of Mathematics & CIDMA

University of Aveiro

Abstract

A constrained principal component analysis for multivariate numerical data, called Clustering and Disjoint Principal Component Analysis (CDPCA), was proposed in 2009 by Vichi and Saporta to detect clusters of objects and, simultaneously, to find a partitioning of variables such that the between cluster deviance in the reduced space of such partition is maximized. Since the partition is formed by a disjoint set of original variables, it will render easier interpretations of the sparse components that define the reduced space of the variables.

A heuristic iterative procedure based on an alternating least square method has been suggested to perform CDPCA and is already available in R in the package `biplotbootGUI`. In this talk, we focus our attention on some recent theoretical developments of the CDPCA methodology. A (high dimensional) genomic data set is considered to analyze the ability of CDPCA for producing clusterings of variables when multiple runs of CDPCA are executed.

FCT Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

This seminar is supported by CIDMA - Center for Research and Development in Mathematics and Applications and by FCT-Fundação para a Ciência e a Tecnologia), within project UID/MAT/04106/2019 (CIDMA).