

COMPUTATIONAL STATISTICS

LAB VII - UNSUPERVISED LEARNING

Luca Bortolussi

Department of Mathematics and Geosciences
University of Trieste

Office 238, third floor, H2bis
`luca@dmi.units.it`

Trieste, Winter Semester 2015/2016

OUTLINE

1 UNSUPERVISED LEARNING

TASK 1 - CLUSTERING AND EM

- Download data from moodle2 (use the datasets for density estimation).
- Implement k-means clustering
- Implement k-Gaussian mixtures clustering, using the EM algorithm to train model parameters (initialise the method using k-means).

TASK 2 - PCA

- Download data from moodle2 (use the new datasets for PCA).
- Implement the PCA algorithm, for a fixed dimension m of the subspace, or fixing m to explain $\alpha\%$ of the variance.