FORWARD PROJECTION FOR HIGH-DIMENSIONAL DATA

T. ORTNER[1], P. FILZMOSER[2], S. BRODINOVA[3],
M. ZAHARIEVA[4], C. BREITENEDER[5]

[1,3,4,5]*Interactive Media Systems Group, TU Wien*

[1,2]*Institute of Statistics and Mathematical Methods in Economics, TU Wien*

[4]*Multimedia Information Systems Group, University of Vienna*

Vienna, AUSTRIA

e-mail: [1]`thomas.ortner@tuwien.ac.at`

Abstract

We provide a novel view on group structure in data. Projecting observations onto a subspace spanned by a small selection of observations, we calculate orthogonal distances as a measure for dissimilarity. Sequentially exchanging the observations, used to span the subspace, we recieve a series of distances. Observations, taken from a similar group structure will behave similar along those projections.

This leads to a visualisation of high dimensional data providing some basic diagnostic on group structures and outliers. The series of distances can be further utilized to perform cluster algorithms, leading to significant improvement when facing clusters located in different subspaces.