

Some remarks on kurtosis projection pursuit for clustering

Cinzia Franceschini¹ and Nicola Loperfido²

¹ *University of Sassari, Italy*

² *University of Urbino “Carlo Bo”, Italy*

Abstract

In cluster analysis, data projections with either maximal or minimal kurtosis are used to gain a better insight into the cluster structure. The method, known as kurtosis projection pursuit for clustering, is theoretically motivated for samples generated from a mixture of two multivariate normal distributions with either equal or proportional covariances. When more than two clusters are present, both heuristic arguments and limited empirical findings suggest that the method tends to detect just two clusters. Theoretically wise, we provide conditions which support the use of kurtosis-optimizing projections in the presence of more than two clusters. Empirically wise, we successfully apply kurtosis minimizations to a real data set where more than two clusters are present. Our empirical findings suggest that kurtosis projection pursuit may accommodate for outliers and that may be used for variable selection in clustering.

Keywords

Dimension reduction, Finite mixtures, K-means, Kurtosis optimization, Principal components, Tandem analysis.