MULTIPLE PRINCIPAL COMPONENT ANALYSES AND PROJECTIVE CLUSTERING

Nittaya Kerdprasop and Kittisak Kerdprasop
Data Engineering and Knowledge Discovery Research Unit
School of Computer Engineering, Suranaree University of Technology, Thailand
E-mail: nittaya, kerdpras@sut.ac.th

Abstract

Projective clustering is a clustering technique for high dimensional data with the inherent sparsity of the data points. To overcome the unreliable measure of similarity among data points in high dimensions, all data points are projected to a lower dimensional sub-space. Principal component analysis (PCA) is an effi-cient method to dimensionality reduction by projecting all points to a lower dimensional subspace so that the information loss is minimized. However, PCA does not handle well the situation that different clusters are formed in different subspaces. We propose a method of multiple principal component analysis for iteratively computing projective clusters. The objective function is designed to determine the subspace associated with each cluster. Some experiments have been carried out to show the effectiveness of the proposed method.

Published in: Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA), Copenhagen, Denmark, , August 22-26, 2005, pp.1132-1136.