

Clustering Based on Principal Components and Technical Efficiency Estimates

Richard N. Capuyan

School of Statistics
University of the Philippines
Diliman, Quezon City

April 2009

Abstract

Cluster analysis searches for natural grouping among observations based on the characteristics of some variables. Labeling of the resulting clusters becomes more complicated as more variables are involved. Dimension reduction through principal components analysis and stochastic frontier modeling are proposed as pre-clustering process. Using principal components scores or technical efficiency estimates, computing time and labeling of resulting clusters becomes simpler.