Robust Simultaneous Confidence Interval Estimation of Principal Component Loadings

By

Martin Augustine B. Borlongan

A thesis submitted in partial fulfillment
Of the requirements for the degree of

Master of Science in Statistics

School of Statistics
University of the Philippines
Diliman, Quezon City

September 2015
Abstract

In principal components analysis, outlying observations cause perturbations in the eigenvectors that consequently influence the principal components (PCs). Estimation of PC loadings are sensitive to outliers and hence, subsequent inference are not optimal. We propose a method that integrates bootstrap into a fully-automated forward search in the construction of robust confidence intervals for the elements of the eigenvectors of the correlation matrix in the presence of outliers. The coverage probability of the proposed bootstrap simultaneous confidence intervals was compared to the coverage probabilities of regular asymptotic confidence region and asymptotic confidence region via Minimum Covariance Determinant (MCD) approach through a simulation study. The proposed method generally yielded more stable coverage probabilities for datasets with or without outliers across several sample sizes compared to approaches based on asymptotic results.

Keywords: principal components analysis, bootstrap, forward search algorithm, outliers