Sparse PCA for high-dimensional data based on
the noise-reduction methodology and its application

Kazuyoshi Yata*;   Makoto Aoshima

Institute of Mathematics, University of Tsukuba, Ibaraki, Japan

E-mail address:
yata@math.tsukuba.ac.jp (Kazuyoshi Yata)
aoshima@math.tsukuba.ac.jp (Makoto Aoshima)

Abstract:
In this talk, we consider the thresholded sparse principal component analysis (TSPCA) in high-dimensional settings. We first illustrate that the TSPCA gives a preferable performance for high-dimensional data if one can find a suitable threshold value. In other words, the performance of the TSPCA heavily depends on a selected threshold value. In this talk, we propose a new sparse estimator of PC directions by using the noise-reduction methodology. We show that the proposed sparse estimation holds consistency properties without depending on any threshold values. we investigate the performance of the proposed TSPCA in simulations. We also give the estimation of shrinkage PC directions and its application to the clustering. Finally, we investigate the performance of the estimated shrinkage PC directions in actual data analyses.

Keywords:
Clustering; High-dimensional PCA; Large p small n; Noise-reduction methodology