Asymptotics of empirical eigenstructure for high dimensional spiked covariance

Weichen Wang

Two Sigma Investments
E-mail: nickweichwang@gmail.com

Abstract: We derive the asymptotic distributions of the spiked eigenvalues and eigenvectors under a generalized and unified asymptotic regime, which takes into account the magnitude of spiked eigenvalues, sample size and dimensionality. This regime allows high dimensionality and diverging eigenvalues and provides new insights into the roles that the leading eigenvalues, sample size and dimensionality play in principal component analysis. Our results are a natural extension of those in [Statist. Sinica 17 (2007) 1617–1642] to a more general setting and solve the rates of convergence problems in [Statist. Sinica 26 (2016) 1747–1770]. They also reveal the biases of estimating leading eigenvalues and eigenvectors by using principal component analysis, and lead to a new covariance estimator for the approximate factor model, called Shrinkage Principal Orthogonal complEment Thresholding (S-POET), that corrects the biases. Our results are successfully applied to outstanding problems in estimation of risks for large portfolios and false discovery proportions for dependent test statistics and are illustrated by simulation studies.