Augmented Estimation of Principal Component Subspace in High Dimensions

Dongsun Yoon
Seoul National University

In this paper, we introduce a novel estimator, called the Augmented Principal Component Subspace, for estimating the principal component subspace for high-dimensional low-sample size data with spiked covariance structure. Our approach augments the naive sample principal component subspace by incorporating additional information from predefined reference directions. Augmented principal component subspace asymptotically reduces every principal angle between the estimated and the true subspaces, thereby outperforming the naive estimator regardless of the metric used. The estimator ' s efficiency is validated both analytically and through numerical studies, demonstrating significant improvements in accuracy when the reference directions contain substantial information about the true principal component subspace. Additionally, we suggest AugmentedPCA using this estimator, and explore connections between our method and the recently proposed James-Stein estimator for principal component directions.