A Unified Optimization Framework for Regularized Multivariate Analysis

Genevera I. Allen*

gallen@rice.edu

Department of Statistics and Electrical & Computer Engineering,

Rice University, Houston, TX, USA;

Department of Pediatrics-Neurology, Baylor College of Medicine,

& Jan and Dan Duncan Neurological Research Institute,

Texas Children's Hospital, Houston, TX, USA.

Abstract

Regularized principal components analysis, especially Sparse PCA and Functional PCA, has become widely used for dimension reduction in high-dimensional settings. Many examples of massive data, however, may benefit from estimating both sparse AND functional factors. These include neuroimaging data where there are discrete brain regions of activation (sparsity) but these regions tend to be smooth spatially (functional). Here, we introduce a framework for regularization of PCA that can encourage both sparsity and smoothness of the row and/or column PCA factors. This framework generalizes many of the existing optimization problems used for Sparse PCA, Functional PCA and two-way Sparse PCA and Functional PCA, as these are all special cases of our method. In particular, our method permits flexible combinations of sparsity and smoothness that lead to improvements in feature selection and signal recovery as well as more interpretable PCA factors. We demonstrate the utility of these methods on simulated data and a neuroimaging example on electroencephalography (EEG) data. This work provides a unified optimization framework for regularized PCA that can form the foundation for a cohesive approach to regularization in high-dimensional multivariate analysis.

Keywords: principal components analysis (PCA); sparse PCA; functional PCA; two-way regularized PCA; neuroimaging.