

Sparse LDA with Network-Guided Block Covariance Matrix

Jin Hyun Nam

Medical University of South Carolina

E-mail: namj@musc.edu

Abstract: In the high-dimensional setting, linear discriminant analysis is faced with two challenges, namely singularity of the covariance matrix and difficulty of interpreting the resulting classifier. Although several methods have been proposed to address these problems, most of them did not take into account dependency between variables and efficacy of selected variables appropriately and they focused only on identifying a parsimonious set of variables maximizing classification accuracy. To address this limitation, here we propose a new approach that directly estimates the sparse discriminant vector without need of estimating the whole inverse covariance matrix, which can be formulated as a quadratic optimization problem. Furthermore, this approach allows to integrate external information to guide the structure of covariance matrix. We applied the proposed method to the transcriptomic study that aims to identify genomic markers predictive of the response to cancer immunotherapy, where the covariance matrix was constructed based on the communities identified from gene-gene networks.