Statistica Sinica **33** (2023), 1-22
doi:https://doi.org/10.5705/ss.202021.0047

AN EFFICIENT CONVEX FORMULATION FOR REDUCED-RANK LINEAR DISCRIMINANT ANALYSIS IN HIGH DIMENSIONS

Jing Zeng[1], Xin Zhang[2] and Qing Mai[2]

[1] *University of Science and Technology of China and* [2] *Florida State University*

Abstract: In this paper, we propose a parsimonious reduced-rank linear discriminant analysis model for high-dimensional sparse multi-class discriminant analysis. We construct a sparse dimension reduction subspace to contain all the information necessary for a linear discriminant analysis. We show explicitly the connections between our model and two well-studied models in the literature: the principal fitted component model in sufficient dimension reduction, and the multivariate reduced-rank regression model. The likelihood-inspired efficient estimator is then recast from a convex optimization perspective. A doubly penalized convex optimization is proposed to unite sparsity and low-rankness in high dimensions, and is then solved efficiently using a three-operator splitting algorithm. We establish the rank selection consistency and classification error consistency of the proposed method when the number of variables grows very fast with the sample size. The effectiveness of the proposed method is demonstrated by means of extensive simulation studies and an application to facial recognition data sets.

Key words and phrases: Dimension reduction, discriminant analysis, linear discriminant analysis, nuclear norm penalty, variable selection.

1. Introduction

High-dimensional linear discriminant analysis (LDA) methods have been widely studied and applied (e.g., Bickel and Levina (2004); Cai and Liu (2011), Shao et al. (2011); Mai, Zou and Yuan (2012)). We consider multi-category classification with $K \geq 2$ classes, where an LDA can identify at most $K-1$ linearly independent discriminant directions. When the dimension of the subspace spanned by all discriminant directions is less than $K-1$, this is known as the reduced-rank LDA problem (Hastie, Tibshirani and Friedman (2009, Chap. 4.3.3)). There are two popular approaches to this problem. The first approach includes methods such as the penalized LDA (Witten and Tibshirani (2011)) and sparse optimal scoring (Clemmensen et al. (2011)). These methods are high-dimensional ex-

Corresponding author: Qing Mai, Department of Statistics, Florida State University, Tallahassee, FL 32306, USA. E-mail: qmai@fsu.edu.