Statistica Sinica **36** (2026), 1-17
doi:https://doi.org/10.5705/ss.202023.0143

SUFFICIENT DIMENSION REDUCTION FOR CLASSIFICATION

Xin Chen, Jingjing Wu, Zhigang Yao and Jia Zhang*

Southern University of Science and Technology, University of Calgary, National University of Singapore and Southwestern University of Finance and Economics

Abstract: We propose a new sufficient dimension reduction approach designed deliberately for high-dimensional classification problems. This novel method is named as Maximal Mean Variance (MMV), inspired by the mean variance index first proposed by Cui, Li and Zhong (2015). MMV requires reasonably mild restrictions on the predictors, and keeps the model-free advantage without the need to estimate the link function. The consistency of the MMV estimator is established under regularity conditions with possibly diverging number of predictors and categories of the response. We also construct the asymptotic normality for the estimator when the dimension of the predictors keeps fixed. The relationship between MMV and several classical classification algorithms are further elaborated. Moreover, although without any definite theoretical guarantee, our method works pretty well when the sample size is far less than the problem dimension. The surprising classification efficiency gain of MMV is demonstrated by simulation studies and real data analysis.

Key words and phrases: Classification, consistency, mean variance index, sufficient dimension reduction.

1. Introduction

Sufficient dimension reduction fits into what is currently quite a hot area in research of high dimensional data. Large quantities of related articles and studies have appeared in recent decades. However, most of the literature focuses on the regression problem where the response Y is a continuous variable, while little is designed specially for the problem of classification with a categorical response.

The slice-based methods, including but not limited to the seminal sliced inverse regression (Li, 1991), sliced average variance estimation (Cook and Weisberg, 1991), directional regression (Li and Wang, 2007) and sliced regression (Wang and Xia, 2008), can be naturally applied to the classification problem with the slices determined directly by the categories of the response. It seems to work nicely, but the number of the slices is strictly restricted by the number of the categories, which can be problematic when there are only a few categories. More specifically, faced with a common binary classification problem, the number

*Corresponding author. E-mail: zhangjia@swufe.edu.cn