

Accepted Manuscript

Robust estimation and variable selection in sufficient dimension reduction

Hossein Moradi Rekabdarkolae, Edward Boone, Qin Wang

PII: S0167-9473(16)30271-7

DOI: <http://dx.doi.org/10.1016/j.csda.2016.11.007>

Reference: COMSTA 6383

To appear in: *Computational Statistics and Data Analysis*

Received date: 8 March 2016

Revised date: 10 November 2016

Accepted date: 11 November 2016



Please cite this article as: Rekabdarkolae, H.M., Boone, E., Wang, Q., Robust estimation and variable selection in sufficient dimension reduction. *Computational Statistics and Data Analysis* (2016), <http://dx.doi.org/10.1016/j.csda.2016.11.007>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Robust Estimation and Variable Selection in Sufficient Dimension Reduction

Hossein Moradi Rekabdarkolae, Edward Boone, Qin Wang*

Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, Richmond, Virginia 23284, U.S.A.

Abstract

Dimension reduction and variable selection play important roles in high dimensional data analysis. Minimum Average Variance Estimation (MAVE) is an efficient approach among many others. However, because of the use of least squares criterion, MAVE is not robust to outliers in the dependent variable or errors with heavy tailed distributions. A robust extension of MAVE through modal regression is proposed. This new approach can adapt to different error distributions and thus brings robustness to the contamination in the response variable. The estimator is shown to have the same convergence rate as the original MAVE. Furthermore, the proposed method is combined with adaptive LASSO to select informative variables. The efficacy of this new solution is illustrated through simulation studies and a data analysis on Hong Kong air quality.

Keywords: Robust estimation, shrinkage estimation, sufficient dimension reduction, variable selection.

1. Introduction

Regression analysis is a standard tool for studying the relationship between the response $y \in R$ and the predictor vector $\mathbf{X} \in R^p$. Two important goals

*Corresponding author

Email addresses: moradirekabh@vcu.edu (Hossein Moradi Rekabdarkolae), elboone@vcu.edu (Edward Boone), qwang3@vcu.edu (Qin Wang)

Download English Version:

<https://daneshyari.com/en/article/4949392>

Download Persian Version:

<https://daneshyari.com/article/4949392>

[Daneshyari.com](https://daneshyari.com)