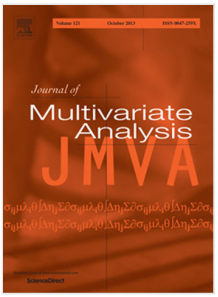
Accepted Manuscript



Model-free sure screening via maximum correlation

Qiming Huang, Yu Zhu

PII: S0047-259X(16)00058-0
DOI: http://dx.doi.org/10.1016/j.jmva.2016.02.014
Reference: YJMVA 4093

To appear in: *Journal of Multivariate Analysis*

Received date: 6 May 2014

Please cite this article as: Q. Huang, Y. Zhu, Model-free sure screening via maximum correlation, *Journal of Multivariate Analysis* (2016), http://dx.doi.org/10.1016/j.jmva.2016.02.014

Model-Free Sure Screening via Maximum Correlation

Qiming Huang* and Yu Zhu†

Department of Statistics, Purdue University

West Lafayette, IN 47907, U.S.A.

Abstract

For screening features in an ultrahigh-dimensional setting, we develop a maximum correlation-based sure independence screening (MC-SIS) procedure, and show that MC-SIS possesses the sure screen property without imposing model or distributional assumptions on the response and predictor variables. MC-SIS is a model-free method in contrast with some other existing model-based sure independence screening methods in the literature. Simulation examples and a real data application are used to demonstrate the performance of MC-SIS and to compare MC-SIS with other existing sure screening methods. The results show that MC-SIS can outperform those methods when their model assumptions are violated, and remain competitive when the model assumptions are satisfied.

Key Words: B-spline; Distance correlation; Optimal transformation; Sure screening property; Variable selection.

*Electronic address: hqm@purdue.edu
†Electronic address: yuzhu@purdue.edu