Cornell University Library

We gratefully acknowledge support from
the Simons Foundation
and member institutions

arXiv.org > math > arXiv:1206.1024

Search or Article-id

(Help | Advanced search)

All papers   Go!

Mathematics > Statistics Theory

Conditional Sure Independence Screening

Emre Barut, Jianqing Fan, Anneleen Verhasselt

(Submitted on 5 Jun 2012 (v1), last revised 31 Oct 2012 (this version, v2))

Independence screening is a powerful method for variable selection for `Big Data' when the number of variables is massive. Commonly used independence screening methods are based on marginal correlations or variations of it. In many applications, researchers often have some prior knowledge that a certain set of variables is related to the response. In such a situation, a natural assessment on the relative importance of the other predictors is the conditional contributions of the individual predictors in presence of the known set of variables. This results in conditional sure independence screening (CSIS). Conditioning helps for reducing the false positive and the false negative rates in the variable selection process. In this paper, we propose and study CSIS in the context of generalized linear models. For ultrahigh-dimensional statistical problems, we give conditions under which sure screening is possible and derive an upper bound on the number of selected variables. We also spell out the situation under which CSIS yields model selection consistency. Moreover, we provide two data-driven methods to select the thresholding parameter of conditional screening. The utility of the procedure is illustrated by simulation studies and analysis of two real data sets.

Subjects: **Statistics Theory (math.ST)**

Cite as: **arXiv:1206.1024 [math.ST]**

(or **arXiv:1206.1024v2 [math.ST]** for this version)

Submission history

From: Emre Barut [view email]
[v1] Tue, 5 Jun 2012 19:06:58 GMT (80kb)
[v2] Wed, 31 Oct 2012 23:49:12 GMT (86kb)

Which authors of this paper are endorsers?

Download:

- PDF
- PostScript
- Other formats

Current browse context:
math.ST
< prev | next >
new | recent | 1206

Change to browse by:

math
stat

References & Citations

- NASA ADS

Bookmark(what is this?)