

MICHIGAN STATE UNIVERSITY
Department of Statistics and Probability

COLLOQUIUM

Hyokyoung Grace Hong
Michigan State University

A data-driven approach to conditional screening of high
dimensional variables

Tuesday, September 15, 2015
10:20 a.m. - 11:10 am
Refreshments 10:00 am
C405 Wells Hall

Abstract

Marginal screening is a widely applied technique to handily reduce the dimensionality of the data when the number of potential features overwhelms the sample size. Due to the nature of the marginal screening procedures, they are also known for their difficulty in identifying the so-called hidden variables that are jointly important but have weak marginal associations with the response variable. Failing to include a hidden variable in the screening stage has two undesirable consequences: (1) important features are missed out in model selection; and (2) biased inference is likely to occur in the subsequent analysis. Motivated by some recent work in conditional screening, we propose a data-driven conditional screening algorithm, which is computationally efficient, enjoys the sure screening property under weaker assumptions on the model, and works robustly in a variety of settings to reduce false negatives of hidden variables. Numerical comparison with alternatives screening procedures are also made to shed light on the relative merit of the proposed method. We illustrate the proposed methodology using a leukemia microarray data example.

To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.