

COLLOQUIUM

Department of Statistics and Probability
Michigan State University

Pengsheng Ji

Assistant Professor
University of Georgia

Optimality in High Dimensional Variable Selection

Tuesday, April 16, 2013
Wells Hall, Room C405
10:20 a.m. - 11:10 a.m.

Abstract

Consider the linear regression model $Y = X_{n,p}\beta + z$ where both p and n are large but $p \gg n$. We find for the setting where both the signal β and the Gram matrix $X'X$ are sparse, neither the lasso nor the subset selection is asymptotically optimal in terms of the Hamming errors, or the phase diagram. We propose the Univariate Penalization Screening (UPS) for variable selection. This is a screen and clean method where we screen with univariate thresholding, and clean with penalized MLE. It has two important properties: sure screening and separable after screening. These properties enable us to reduce the original regression problem to many small-size regression problems that can be fitted separately. The UPS is effective both in theory and in computation. Some connections with optimality in multiple testing are established.

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