MICHIGAN STATE UNIVERSITY
Department of Statistics and Probability

COLLOQUIUM

Qi Long
University of Pennsylvania

Knowledge-Guided Scalable Bayesian Variable Selection for High-dimensional Data

Tuesday, December 5, 2017
10:20 a.m. - 11:10 am
Refreshments 10:00 am
C405 Wells Hall

Abstract

Variable selection for structured covariates lying on an underlying known graph is a problem motivated by practical applications, and has been a topic of increasing interest. However, most of the existing methods may not be scalable to high dimensional settings involving tens of thousands of variables lying on known pathways such as the case in genomics studies. We propose an adaptive Bayesian shrinkage approach which incorporates prior network information by smoothing the shrinkage parameters for connected variables in the graph, so that the corresponding coefficients have a similar degree of shrinkage. We fit our model via a computationally efficient expectation maximization algorithm which scalable to high dimensional settings. Theoretical properties for fixed as well as increasing dimensions are established, even when the number of variables increases faster than the sample size. We demonstrate the advantages of our approach in terms of variable selection, prediction, and computational scalability in a simulation study, and apply the method to a cancer genomics study. Joint work with Changgee Chang and Suprateek Kundu.

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