

MICHIGAN STATE UNIVERSITY
Department of Statistics and Probability

COLLOQUIUM

Yi Li

University of Michigan

Ultrahigh Dimensional Time Course Feature Selection

Tuesday, November 25, 2014
10:20 a.m. - 11:10 am
Refreshments 10:00 am
C405 Wells Hall

Abstract

Statistical challenges arise from modern biomedical studies that produce time course genomic data with ultrahigh dimensions. In a renal cancer study that motivated this paper, the pharmacokinetic measures of a tumor suppressor (CCI-779) and expression levels of 12625 genes were measured for each of 33 patients at 8 and 16 weeks after the start of treatments, with the goal of identifying predictive gene transcripts and the interactions with time in peripheral blood mononuclear cells for pharmacokinetics over the time course. The resulting dataset defies analysis even with regularized regression. Although some remedies have been proposed for both linear and generalized linear models, there are virtually no solutions in the time course setting. As such, a novel GEE-based screening procedure is proposed, which only pertains to the specifications of the first two marginal moments and a working correlation structure. Different from existing methods that either fit separate marginal models or compute pairwise correlation measures, the new procedure merely involves making a single evaluation of estimating functions and thus is extremely computationally efficient. The new method is robust against the mis-specification of correlation structures and enjoys theoretical readiness, which is further verified via Monte Carlo simulations. The procedure is applied to analyze the aforementioned renal cancer study and identify gene transcripts and possible time-interactions that are relevant to CCI-779 metabolism in peripheral blood.

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