Department of Mathematics and Statistics Colloquium

A sequential test for variable selection in high dimensional complex data

Kofi Placid Adragni

University of Maryland Baltimore County

Abstract: Given a high dimensional p -vector of continuous predictors X and a univariate response Y, principal fitted components (PFC) provide a sufficient reduction of X that retains all regression information about Y in X while reducing the dimensionality. The reduction is a set of linear combinations of all the p predictors, where with the use of a flexible set of basis functions, predictors related to Y via complex, nonlinear relationship could be detected.

In the presence of possibly large number of irrelevant predictors, the accuracy of the sufficient reduction is hindered. The proposed method adapts a sequential test to the PFC to obtain a pruned sufficient reduction that shed off the irrelevant predictors. The sequential test is based on the likelihood ratio which expression is derived under different covariance structures of $X|Y$. The resulting reduction has an improved accuracy and also allows the identification of the relevant variables.

Monday, October 12 at 3:45 in Roop 103

refreshments at 3:30