Title: Group Feature Screening via F-test

Speaker: Dr. Won Chul Song, Purdue University

Date: Tuesday, March 7, 2017

Time and Location: 12:00 p.m.-1:00 p.m, 3-1637

Abstract:

Feature screening is crucial in analysis of ultrahigh dimensional data, where the number of variables (features) is exponentially larger than the number of observations. In various ultrahigh dimensional data, variables are naturally grouped with correlation within groups. In this article, we propose a new feature screening method for data with grouped variables and conduct group screening via the F-test. This is in clear contrast with the existing literature on feature screening, which screens variables one by one. Under mild technical conditions, we prove that our group screening procedure possesses a sure screening property that it selects all active groups with probability approaching one at an exponential rate, meanwhile it controls the number of false positives. We use simulations to demonstrate the advantages of the proposed method and show its application in a genome-wide association study. We conclude that grouping method is very effective in large-scale inference for high dimensional data and a good choice of optimal test can detect true signals with desired properties.