Title:
Statistical Inference of Linear Parameters in High-Dimensional Regression

Speaker:
Dr. Cun-Hui Zhang
Department of Statistics & Biostatistics
Rutgers University

Time & Place:
Wednesday, October 15, 2014, 4:00–5:00pm
Room 140 Bardeen
(Cookies & Coffee @ 3:30 in Rm 1210 MSC)

Abstract:

We consider efficient estimation of a linear function of regression coefficients at $n^{-1/2}$ rate in high-dimensional regression models. Let $s$ be the complexity of the regression coefficient vector as measured by the size of its support or a capped-$\ell_2$ norm. A low-dimensional projection estimator provides efficient point and interval estimation of the linear function when $s \log p \ll n^{1/2}$ under certain regularity conditions on the design matrix. However, for i.i.d. designs with a known population Gram matrix, the sample size condition $s \log p \ll n$ suffices. We close this gap by proving that the condition $s \log p = O(n^{-1/2})$ is necessary for regular inference at $n^{-1/2}$ rate for Gaussian designs when the population Gram matrix is unknown. This demonstrates an important benefit of unsupervised data in linear regression. Moreover, a sufficient and nearly necessary condition on $\{s, n, p\}$ for efficient estimation at $n^{-1/2}$ rate is obtained for semi-supervised data and/or under sparsity conditions on the population Gram matrix of the design.