Speaker: Emmanuel J. Candès, Stanford University

Title: Statistical Estimation and Testing via the Sorted L1 Norm

Abstract:

This talk introduces a novel method for sparse regression and variable selection, which is inspired by modern ideas in multiple testing. Imagine we have observations from the linear model, then we suggest estimating the regression coefficients by means of a new estimator called SLOPE, an acronym for Sorted L-One Penalized Estimation. SLOPE trades off the residual sum of squares with the sorted L1 norm, which penalizes the regression coefficients according to their rank: the higher the rank, the larger the penalty. This is in analogy with the famous BHq procedure [Benjamini and Hochberg, 1995], which compares the value of a test statistic taken from a family to a critical thresholds that depends on its rank in the family.

SLOPE is a convex program and we demonstrate an efficient algorithm for computing the solution. For orthogonal designs, we prove that one can control the false discovery rate (FDR) for variable selection. When the design matrix is nonorthogonal we demonstrate inherent limitations on the FDR level and the power which can be obtained with model selection methods based on L1-like penalties. However, whenever the columns of the design matrix are not strongly correlated, we show empirically that it is possible to select the regularizer as to obtain FDR control at a reasonable level as long as the number of nonzero coefficients is not too large. At the same time, the procedure exhibits increased power over the lasso, which treats all coefficients equally, and over other multiple comparison procedures. We shall discuss some estimation properties of the new selection rule through diverse simulation studies.

Joint work with M. Bogdan, E. van den Berg, and W. Su.