Annual Report 3

January 31, 2013

Adam Sales

Title: Propensity scores with many more adjustment variables than observations: Evaluation of a math enrichment progam via school-level matching

Abstract: The rapid proliferation of data presents an opportunity, and a challenge for causal inference: how can researchers adjust causal estimates for a large number of variables? In the course of evaluating an educational-enrichment program, we suggest a method for incorporating an arbitrary number of pre-treatment covariates into a propensity score matching scheme. Propensity-score matching compares treated units against non-treated units with similar estimated probabilities of treatment assignment, conditional on a set of covariates. In this scenario, there were tens of treated schools, but thousands of covariates. This paper uses principal components analysis to both construct the match and test its validity. It also attempts to compare this approach with other high-dimensional regression techniques such as ridge regression and the LASSO.

Sumanta Basu

Title: Network Granger Causality with Inherent Grouping Structure

Abstract: The problem of estimating high-dimensional network models arises naturally in the analysis of many biological and socio-economic systems. In this work, we aim to learn a network structure from temporal panel data, employing the framework of Granger causal models under the assumptions of sparsity of its edges and inherent grouping structure among its nodes. To that end, we introduce a group lasso regression regularization framework, and also examine a thresholded variant to address the issue of group misspecification. Further, the norm consistency and variable selection consistency of the estimates is established, the latter under the novel concept of direction consistency. The performance of the proposed methodology is assessed through an extensive set of simulation studies and comparisons with existing techniques. The study is illustrated on two motivating examples coming from functional genomics and financial econometrics.

Nirupam Chakrabarty

Title: Semiparametric Estimation of Target Locations

Abstract: Detection, identification and tracking of spatial phenomena are important tasks in various environmental and infrastructure applications. Wireless sensor networks are widely used for monitoring natural phenomena in space and over time, as well as for target detection and tracking. Sensors acquire signals emitted from the target that are corrupted by noise, and then try to detect the location of the target based on the information received. In most cases, the form of the signal generating model is assumed to be known for subsequent analysis. But the assumption of a known signal model can be restrictive, since different types of targets may emit different kinds of signals. Here we introduce a semi parametric model for signal propagation where both the location of the target and the signal generating function are treated as unknown for subsequent inference. In the proposed model, the parametric part (target location) and the nonparametric part (signal function) are "bundled" together, which makes the problem more challenging. We propose a two-stage estimation method for detecting the target location.

Xi Lu

Title: Assisted Policy Search

Abstract: Treatment policies are multi-stage decision rules that assign treatments according to patients' characteristics and diagnostic courses, which aim to control chronic diseases and mental disorders. Oftentimes clinicians and scientists have knowledge about useful tailoring variables, hence parameterized policy classes could be considered as the candidate pool to search for optimal treatment policy, which is defined as the policy that yields the highest mean value of response outcome. We propose an estimator for optimal policy that is assisted by the modelling of treatment effects, leaving the components of data-generating process that are irrelevant of policy search unparameterized. Augmented version of the assisted estimator is also proposed to further improve efficiency. When the policies of interest take special form of linear decision rules, by viewing the policy search problem as minimizing weighted misclassification error rate, we propose efficient algorithm based on convex upper bound of 0-1 loss. Simulation results are provided to investigate computational efficiency of surrogate algorithm and robustness with respect to mis-specifying treatment effects models.