L'inférence causale en présence de dépendance et de structure de réseau
« Atelier sur la découverte de la structure causale à haute dimension »
25 au 27 juin 2018

Causal Inference in the Presence of Dependence and Network Structure
"Workshop on Discovery of causal structure in high dimensions"
June 25-27, 2018

High-Dimensional propensity score estimation via covariate balancing

Yang Ning [*]

yn265@cornell.edu

In this paper, we address the problem of estimating the causal effects in observational studies when the number of potential confounders is possibly much greater than the sample size. In particular, we develop a robust method to estimate the propensity score via covariate balancing in high-dimensional settings. Since it is usually impossible to obtain the exact covariate balance in high dimension, we propose to estimate the propensity score by balancing a carefully selected subset of covariates that are predictive of the outcome under the assumption that the outcome model is linear and sparse. The estimated propensity score is, then, used for the Horvitz-Thompson estimator to infer the treatment effects. We prove that the proposed methodology has the desired properties such as sample boundedness, root-n consistency, asymptotic normality, and semiparametric efficiency. We then extend these results to the case where the outcome model is a sparse generalized linear model. In addition, we show that the proposed estimator remains root-n consistent and asymptotically normal even when the model is misspecified. Finally, we show that our proposal is able to estimate the treatment effect more accurately than the existing proposals and apply it to estimate the causal effects of college attendance on adulthood political participation. Open-source software is available for implementing the proposed methodology.

This is the joint work with Peng and Imai.

[*]Department of Statistical Science, Cornell University, Comstock Hall 1188, Ithaca, NY 14853, USA