Lasso adjustments of treatment effect estimates in randomized experiments

Bin Yu
Departments of Statistics and EECS, UC Berkeley
statistics.berkeley.edu/~binyu
binyu@berkeley.edu

In randomized experiments, linear regression is often used to adjust for imbalances in covariates between treatment groups, yielding an estimate of the average treatment effect with lower asymptotic variance than the unadjusted estimator. If there is large number of covariates, many of which are irrelevant to the potential outcomes, the Lasso can be used to both select relevant covariates and perform the adjustment. We study the resulting estimator under the Neyman model for randomization, and show that it is more efficient than the unadjusted estimator, and that it is possible to give a conservative estimate of the asymptotic variance. Simulations show that Lasso-based adjustment can be advantageous even when $p < n$. Moreover, when the covariates selected by the Lasso indicate the presence of heterogenous treatment effects, our method can yield conditional treatment effect estimates for subpopulations.
(This talk is based on joint work with Adam Bloniarz, Hanzhong Liu, Jas Sekhon and Cunhui Zhang.)