Matching and ATT Estimation via Random Recursive Partitioning

Stefano Iacus and Guiseppe Porro

Average Treatment Effect estimation is the problem of measuring the difference in outcome of two groups of treated and control units. In contexts where randomized experiment are difficult to be carried out (like in econometrics, sociology, political analysis, etc), the experimenter is only able to select observations to be to exposed a treatment and subsequently search for a control group from other data sources. In such observational studies, to reduce the bias due to different pre-treatment conditions between treated and control units, it is usually applied some matching method.

Due to the curse of dimensionality, direct matching on the space of covariates X is usually hard to carry out. Approaches based on different notions of similarity have been developed, like propensity score methods and distance based methods. In these cases, matching is based on a one-dimensional quantity instead of looking for similar individuals in the k-dimensional space X.

In this talk we present the Random Recursive Partitioning (RRP) method, which operates directly on the space of covariates X. The method randomly generates non-empty recursive partitions of the space X and evaluates how frequently two individuals lie in the same subset of X. RRP is indeed a Monte Carlo method on the set of possible partitions of X for estimating the likelihood of the "proximity" of two individuals. We also propose some tools to measure the extent of the common support between treated and control units. No average treatment effect estimation can in fact carried out without a sufficient overlap between the two groups.

Applications to real and simulated data and corresponding R code is presented.