COLLOQUIUM

Balancing Flexibility and Interpretability: A Conditional Linear Model Estimation via Random Forest

Professor Marcelo C. Medeiros,
Department of Economics, University of Illinois at Urbana-Champaign

5-1-25

Neckers 156 - Time: 3:00-4:00 pm

Reception immediately following in the math library.

Abstract:
Traditional parametric econometric models often rely on rigid functional forms, while nonparametric techniques, despite their flexibility, frequently lack interpretability. This paper proposes a parsimonious alternative by modeling the outcome Y as a linear function of a vector of variables of interest X, conditional on additional covariates Z. Specifically, the conditional expectation is expressed as $E[Y|X,Z]=X^T\beta(Z)$, where $\beta()$ is an unknown Lipschitz-continuous function. We introduce an adaptation of the Random Forest (RF) algorithm to estimate this model, balancing the flexibility of machine learning methods with the interpretability of traditional linear models. This approach addresses a key challenge in applied econometrics by accommodating heterogeneity in the relationship between covariates and outcomes. Furthermore, the heterogeneous partial effects of X on Y are represented by $\beta()$ and can be directly estimated using our proposed method. Our framework effectively unifies established parametric and nonparametric models, including varying-coefficient, switching regression, and additive models. We provide theoretical guarantees, such as pointwise and Lp-norm rates of convergence for the estimator, and establish a pointwise central limit theorem through subsampling, aiding inference on the function $\beta()$. We present Monte Carlo simulation results to assess the finite-sample performance of the method.