

ランダムフォレストによる因果推論

慶應義塾大学 研究員
中村知繁

Regression tree 及び Random forest から得られる推定量の性質が Wager and Wether(2014) 及び Wager and Athey (2018) で明らかにされて以来、Random forest を用いた異質性考慮した因果効果の推定に注目が集まっている。特に、Wager and Athey(2018) によって提案された causal forest は、傾向スコアに対するモデル及び、潜在結果変数に対する直接的なモデリングを必要とせず、ノンパラメトリックに異質性を考慮した因果効果を推定できるという点で画期的であった。さらに、Athey, Tibshirani and Wager (2019) によって Generalized Random Forest (GRF) が提案され、Random forest によって得られる forest kernel を用いて、適当な正則条件を満たす局所推定方程式を解くことができることが示された。その結果、より広いクラスの因果推論の問題における異質性を考慮した因果効果を推定できるようになった。この他、近年は Neyman-orthogonality を用いた GRF である orthogonal random forest (Oprescu, Syrgkanis and Wu 2019) なども提案され、積極的に推定量の改善が行われている。

本発表では、Tree 及び Random forest について説明した後、これらを用いた因果効果の推定法である causal forest について詳しく説明する。また、近年注目を集めている GRF や、GRF と Neyman-orthogonality の組み合わせによる推定量の改善などについても紹介する予定である。