Inference on heterogeneity aware estimand

川田恵介

Table of contents

1 Weighted average difference 1
- 1.1 サブサンプル平均 . . . 1
- 1.2 Estimand: Weighted average difference . . . 2
- 1.3 Estimand: Proportion weight . . . 2
- 1.4 Moment condition . . . 2
- 1.5 Argumented inverse propensity score . . . 2
- 1.6 Algorithm . . . 3

2 Best Lienar Projection for CATE 3
- 2.1 Algorithm . . . 3
- 2.2 Best Linear projection . . . 3

3 Proportional weight の問題点と対策 3
- 3.1 Overlap weigh . . . 4
- 3.2 Propensity score weight weigh . . . 4

1 Weighted average difference

- ここまで $E[Y|D=1,X]-E[Y|D=0,X]=\tau$ を仮定

  - ほとんどの応用で X によって異なることが予想される

    `- $\tau(X) =$ Conditional average difference (function)`

    - 因果推論では `Conditional average treatment effect (CATE)` と呼ばれる

1.1 サブサンプル平均

- $X=x$ を満たすサブサンプルサイズが十分にあれば、ただのサブサンプル平均差で、応用上問題ない

– CATE を直接推定できる

- ほとんどの応用で、大量の X を用いるので、サブサンプルサイズは不足する

1.2 Estimand: Weighted average difference

- Estimand
$$\theta_0 = \int \tau(X) \times \omega(X) dX$$

- $\omega(X) =$ “研究者” が暗黙のうちに設定する集計用 Weight

1.3 Estimand: Proportion weight

- $\omega(X) = f(X)$
  – 因果推論では Average Treatment Effect と呼ばれる

- $$\theta_0 = \int \tau(X) \times f(X) dX$$

1.4 Moment condition

- 複数存在する: $\theta = E[m(Y, D, X)]$ where $m(O)$
  – $$= \mu_Y(1, X) - \mu_Y(0, X)$$
  – $$= \frac{DY}{\mu_D(X)} - \frac{(1-D)Y}{1-\mu_D(X)}$$

1.5 Argumented inverse propensity score

- おすすめの Moment condition

- $$m(O) = \mu_Y(1, X) - \mu_Y(0, X)$$
$$+ \underbrace{\frac{D(Y - \mu_Y(1, X))}{\mu_D(X)} - \frac{(1-D)(Y - \mu_Y(0, X))}{1-\mu_D(X)}}_{Adjustment}$$

- Neyman’s orthogonal condition を満たす

1.6 Algorithm

1. データ分割 (auxiliary/main data)
2. μ_Y, μ_D を auxiliary data (+ 機械学習) で推定する
3. Moment condition に代入 $m(O, g)$ し、main data を持ちいて平均値を推定、信頼区間とともに推定する

2 Best Lienar Projection for CATE

- 平均差では、X との関係性について、何もわからない
  - 信頼区間もしっかり推定しつつ、CATE の持つ特徴を、もう少し理解することを目指す

2.1 Algorithm

1. データ分割 (auxiliary/main data)
   - 交差推定も活用可能
2. μ_Y, μ_D を auxiliary data (+ 機械学習) で推定する
3. Moment condition に代入 $m(O, g)$ し、main data を持ちいて $m \sim Z$ を OLS で推定、信頼区間とともに推定する

2.2 Best Linear projection

- 以下の関数を推定する
$$\min E[(\tau(X) - g_\tau(Z))^2]$$
  where $Z \subset X$
- 線形モデルで近似する
$$g_\tau = \beta_0 + \beta_1 Z_1 + .. + \beta_L Z_L$$

3 Proportional weight の問題点と対策

- 完璧な対策はない

3.1 Overlap weigh

- Proportional weight 以外にも色々な weight が存在する
- R learner でかつ $D \in \{0, 1\}$ であれば、

$$\omega(X) = \frac{E[D|X](1 - E[D|X])f(X)}{\int E[D|X](1 - E[D|X])f(X)dX}$$

- Overlap weight と呼ばれる

3.2 Propensity score weight weigh

- $$\omega(X) = \frac{E[D|X]f(X)}{\int E[D|X]f(X)dX}$$

- 因果推論において、Average Treatment Effect on Treated において暗黙のうちに使用される