

# Summary and Future Reading

## 機械学習の経済学への応用

川田恵介

### まとめ

#### 教師付き学習

- 本講義: 以下を可能な限り削減する  $f(X)$  を推定する手法群

$$E[(E[Y|X] - f(X))^2]$$

- 注: 削減したい Population Risk に応じて、変わってくる

#### 応用: 予測

- $E[(Y - \text{Pred}Y(X))^2]$  最小化する予測関数  $\text{Pred}Y(X)$  を推定したい
  - $f(X)$  は優れた予測モデル
  - 機械学習のそもそもの問題意識
  - [Introduction to Statistical Learning](#)

#### 応用: 母パラメータの推論

- 母集団のパラメータ推定: 母分布  $f(Y, D, X)$  を”要約”するパラメトリックな関数  $\beta_D(X)$  を推定誤差つきで推定したい
  - 一般に Nuisance 関数を含む Moment 法として定式化できる
- $f(X)$  は,
  - 一致性は満たしやすい  $\iff$  パラメトリック推定
  - 収束は遅い

## Neyman's Orthogonality & 交差推定 & 教師付き学習

- ここ 10 年ほどで大きく発展
- Nuisance 関数を機械学習で推定し、一致性を担保
- 収束速度の遅さをカバーするように
  - Nuisance 関数の局所的な推定誤差の影響を受けないようなモーメント (Neyman's Orthogonality)
  - 交差推定で推定された Nuisance 関数と誤差項の相関を断ち切る

### 例: Partial Linear Model

- モーメント推定:  $\bar{\beta}_D = \text{スカラー}$

$$0 = \sum (D_i - f_D(X_i)) \times (D_i - f_D(X_i) - \bar{\beta}_D (Y_i - f_Y(X_i))^2)$$

- $f_Y(X_i), f_D(X_i)$  が  $n^{1/4}$  以上の速度で  $E[Y|X], E[D|X]$  に収束すれば OK
  - Neyman's orthogonality を満たすモーメントを推定、Nuisance 関数を交差推定

## Y と D の関係性把握

- $E[Y|D = d', X] - E[Y|D = d, X]$  の特徴を推論
  - 平均ないし BestLinearPredictor (Chernozhukov et al. 2018; Semenova and Chernozhukov 2021)
  - ノンパラメトリック (Nie and Wager 2021; Athey, Tibshirani, and Wager 2019; Wager and Athey 2018)
  - Group Average Difference (Kallus 2022; Chernozhukov et al. 2020)

## 他の応用

- Neyman's orthogonality (local robustness, efficient influence function) は色々応用されている

### 他の応用: Panel Data

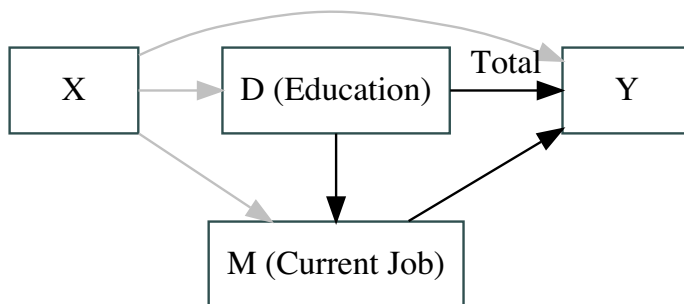
- 2 期間パネルであれば、FirstDifference をとった後に普通に適用

$$E[\Delta Y | \Delta D, X]$$

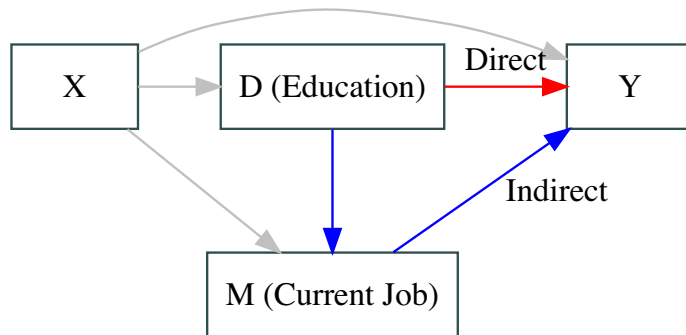
- $\Delta Z = Z_{i1} - Z_{i0}$
- 発展: Chang (2020); Nie, Lu, and Wager (2019)
  - 新しい識別問題も含めたサーベイ Chaisemartin and D'Haultfoeuille (n.d.)

## 他の応用: Mediation Analysis

- 潜在結果で定義する因果効果は、全ての”経路”を含む



## 他の応用: Mediation Analysis



- Farbmacher et al. (2022); Díaz et al. (2021)
  - $X$  だけでなく,  $M$  が多次元でも OK

## 他の応用: Sampling Weight

- 条件つき平均差を推定するデータ (例えば、実験データ) と適用した母集団を代表するデータ (例えば、国勢調査) を用いて、Sampling Weight  $p(X)$  を推定
- $p(X) = \text{Nuisance 関数}$ 
  - Neyman's orthogonality & 交差推定
- Yang, Kim, and Song (2020)
  - Survey (Yang and Kim 2020)

## 他の応用: Omitted Variable

- $E[Y|D = d, X, \underbrace{U}_{\text{Unobservable}}] - E[Y|D = d', X, U]$

- Sensitivity Analysis:  $U$  の影響についてシナリオを想定し、結果がどの程度変化するか推定
  - $X$  の情報を使って、シナリオを作成可能
- Cinelli and Hazlett (2020) : Linear Model
  - Chernozhukov et al. (2022) : 一般化

## Suggestion

- 機械学習の応用: 大量の変数を利用した分析を容易に
- コントロール変数を使用している研究であれば、すぐに可能
  - 何が仮定されているのか?
  - 伝統的な手法と比べた利点は何か?
- 信頼できる母集団の推論を行うためには、現状、
  - Nuisance を意識
  - 研究関心となる変数の関係性を絞り込む
- 動機付けと識別は、人間が丁寧に議論

## Reference

- Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. “Generalized Random Forests.” *The Annals of Statistics* 47 (2): 1148–1178. <https://doi.org/10.1214/18-AOS1709>.
- Chaisemartin, Clément de, and Xavier D’Haultfoeuille. n.d. “Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey.” *The Econometrics Journal*.
- Chang, Neng-Chieh. 2020. “Double/Debiased Machine Learning for Difference-in-Differences Models.” *The Econometrics Journal* 23 (2): 177–91.
- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. “Double/Debiased Machine Learning for Treatment and Structural Parameters.” *The Econometrics Journal* 21 (1): C1C68. <https://doi.org/10.1111/ectj.12097>.
- Chernozhukov, Victor, Carlos Cinelli, Whitney Newey, Amit Sharma, and Vasilis Syrgkanis. 2022. “Long Story Short: Omitted Variable Bias in Causal Machine Learning.” <https://doi.org/10.3386/w30302>.
- Chernozhukov, Victor, Mert Demirer, Esther Duflo, et al. 2020. “Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments.” *Econometrica Forthcoming*.
- Cinelli, Carlos, and Chad Hazlett. 2020. “Making Sense of Sensitivity: Extending Omitted Variable Bias.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (1): 39–67.
- Díaz, Iván, Nima S Hejazi, Kara E Rudolph, and Mark J van Der Laan. 2021. “Nonparametric Efficient Causal Mediation with Intermediate Confounders.” *Biometrika* 108 (3): 627–41.

- Farbmacher, Helmut, Martin Huber, Lukáš Lafférs, Henrika Langen, and Martin Spindler. 2022. “Causal Mediation Analysis with Double Machine Learning.” *The Econometrics Journal* 25 (2): 277–300.
- Kallus, Nathan. 2022. “Treatment Effect Risk: Bounds and Inference.” *arXiv Preprint arXiv:2201.05893*, (MR by Management Science).
- Nie, Xinkun, Chen Lu, and Stefan Wager. 2019. “Nonparametric Heterogeneous Treatment Effect Estimation in Repeated Cross Sectional Designs.” arXiv. <https://doi.org/10.48550/ARXIV.1905.11622>.
- Nie, Xinkun, and Stefan Wager. 2021. “Quasi-Oracle Estimation of Heterogeneous Treatment Effects.” *Biometrika* 108 (2): 299–319.
- Semenova, Vira, and Victor Chernozhukov. 2021. “Debiased Machine Learning of Conditional Average Treatment Effects and Other Causal Functions.” *The Econometrics Journal* 24 (2): 264289. <https://doi.org/10.1093/ectj/utaa027>.
- Wager, Stefan, and Susan Athey. 2018. “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests.” *Journal of the American Statistical Association* 113 (523): 1228–42. <https://doi.org/10.1080/01621459.2017.1319839>.
- Yang, Shu, and Jae Kwang Kim. 2020. “Statistical Data Integration in Survey Sampling: A Review.” *Japanese Journal of Statistics and Data Science* 3 (2): 625–50.
- Yang, Shu, Jae Kwang Kim, and Rui Song. 2020. “Doubly Robust Inference When Combining Probability and Non-Probability Samples with High Dimensional Data.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (2): 445–65.