Estimating Identifiable Causal Effects through Double Machine Learning

Yonghan Jung
PURDUE UNIVERSITY

Jin Tian
IOWA STATE UNIVERSITY

Elias Bareinboim
COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

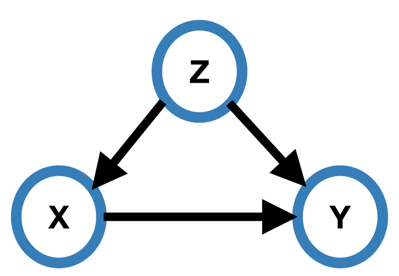
Summary

- Causal effect estimators robust to biases have been developed for very specific settings such as back-door criterion / ignorability assumption.
- *Double Machine Learning (DML)* [1] has been proposed for devising ***debiased*** estimators, which converges at $\sqrt{N}$ rate even when parameters composing estimators ('nuisance') converge slowly.
- We propose ***doubly robust*** & ***debiased*** causal estimator for ***any identifiable causal effect*** based on *DML*.

Example — DML for Back-door

Roughly, a DML estimator is based on the function called '*Influence function*' and the technique called '*Cross-fitting*'.

Causal diagram (G)



Data (D)

$D = \{\mathbf{V}_{(i)}\}_{i=1}^{N}$ for

$\mathbf{V} = \{Z, X, Y\}$

Causal functional.

$\Psi(P) = P(y \mid do(x))$

$= \sum_{z} P(y \mid x, z)P(z)$

Constructing DML estimator

1. **Influence function (IF)** $\phi(\mathbf{V}; \psi, \eta)$**:** (Roughly) A sensitivity of a causal functional $\Psi(P)$ with respect to the small changes in P, that depends on the target parameter $\psi \equiv P(y \mid do(x))$ and *nuisance* η.

$$\phi(\mathbf{V}; \psi, \eta = (\eta_0, \eta_1)) \equiv \frac{I_x(X)}{P(X \mid Z)}(I_y(Y) - P(y \mid X, Z)) + P(y \mid x, Z) - \psi$$

2. **Uncentered influence function (UIF)** $\mathcal{V}(\mathbf{V}; \eta) \equiv \phi(\mathbf{V}; \psi, \eta) + \psi$, that depends on the nuisance η .

3. **DML estimator** T_N **(Cross-fitting):** Training & evaluating estimators based on the UIF are done with distinct samples; For distinct samples (D_a, D_b), and $(\widehat{\eta}_a, \widehat{\eta}_b)$ nuisances trained using data (D_a, D_b):

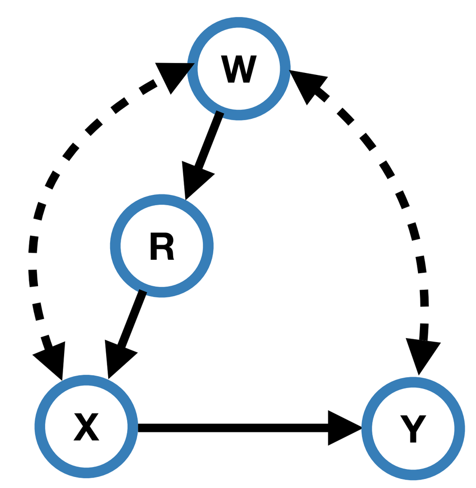
$$T_N \equiv \frac{2}{N}\sum_{\mathbf{V}_{(i)} \in D_a} \mathcal{V}(\mathbf{V}_{(i)}; \eta_b) + \frac{2}{N}\sum_{\mathbf{V}_{(i)} \in D_b} \mathcal{V}(\mathbf{V}_{(i)}; \eta_a).$$

[1] Chernozhukov, Victor, et al. "Double/debiased machine learning for treatment and structural parameters." *The Econometrics Journal* 21.1 (2018): C1-C68.

DML Properties of T_N

1. **Doubly Robust:** T_N converges to ψ whenever η_0 or η_1 is correct.
2. **Debiasedness:** T_N converges at $\sqrt{N}$ rate to ψ even when η_0 and η_1 converges $N^{-1/4}$ rate.

Example — DML for non-Backdoor case



[Example 1]

- The causal effect is identifiable and expressible as a function of two back-door adjustments:

$$P(y \mid do(x)) = \frac{\sum_w P(x, y \mid r, w)P(w)}{\sum_w P(x \mid r, w)P(w)} = \frac{M_1}{M_2}$$

- Then, UIF $\mathcal{V}$ is given as a function of $\mathcal{V}_{M_i}(\mathbf{V}; \eta_i = (\eta_{0,i}, \eta_{1,i}))$, UIFs for M_i and $\mu_{M_i} = \mathbb{E}_P[\mathcal{V}_{M_i}]$, for i=1,2.

$$\mathcal{V}(\mathbf{V}; \eta) \equiv A(\mathcal{V}_{M_1}, \mathcal{V}_{M_2}) = \frac{1}{\mu_{M_2}}(\mathcal{V}_{M_1} - \frac{\mu_{M_1}}{\mu_{M_2}}(\mathcal{V}_{M_2} - \mu_{M_2}))$$

- The DML estimator T_N based on $\mathcal{V}$ achieves ***doubly robustness*** and ***debiasedness*** with respect to $\eta_i = (\eta_{0,i}, \eta_{1,i})$ where

$$\eta_1 = \{\underbrace{P(x, y \mid r, w)}_{\eta_{1,0}}, \underbrace{P(r \mid w)}_{\eta_{1,1}}\} \text{ and } \eta_2 = \{\underbrace{P(x \mid r, w)}_{\eta_{2,0}}, \underbrace{P(r \mid w)}_{\eta_{2,1}}\}.$$

DML for *any* identifiable causal estimands

1. (**DML-ID** in Algo. 1) represents a causal effect as a function of multi-outcome sequential back-doors (mSBD) adjustment M_i [2]:

Theorem 1: Expressibility

Any identifiable causal effect can be expressed as a function of mSBD adjustments through **DML-ID;** i.e., $\psi = A(\{M_i\})$.

2. (**DeriveUIF** in Algo. 2) derives/expresses an UIF of as a function of UIFs of mSBDs.

Theorem 2: Derivation of UIF

An UIF for a causal effect $\psi = A(\{M_i\})$ can be expressed as a function of UIFs of mSBD M_i; $\mathcal{V} = B(\{\mathcal{V}_{M_i}(\mathbf{V}; \eta_i = (\eta_{i,0}, \eta_{i,1})\})$.

3. (**DML estimator** in Def. 4) constructs T_N based on the UIF and cross fitting technique.

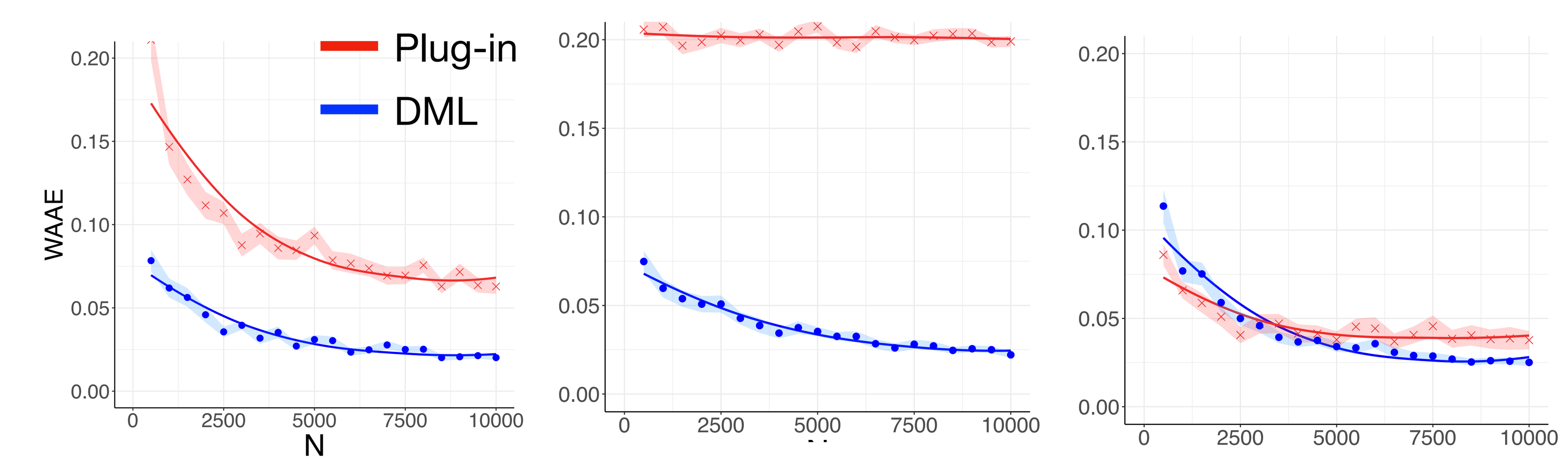
Theorem 3: DML Properties

A DML estimator T_N achieves doubly robustness and debiasedness, with respect to $\eta_i = (\eta_{i,0}, \eta_{i,1})$.

[2] Jung, Yonghan, Jin Tian, and Elias Bareinboim. "Estimating causal effects using weighting-based estimators." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 06. 2020.

Simulation for Example 1

- A proposed DML estimator is compared with the plug-in estimator, only viable estimator working for identifiable causal functional.



- **(Debiasedness; Left)** DML converges (i.e., the error 'WAAE' decreases) faster even when nuisances converge slower rate ($N^{-1/4}$).
- **(Doubly Robustness; (Center, Right))** DML converges even when models for either $\eta_{i,0}$ (center) or $\eta_{i,1}$ (right) is misspecified.

Conclusion

- (**DML-ID & DeriveUIF**) We devised algorithms to represent any identifiable causal functional as a function of multi-outcome sequential back-doors (mSBD) and derive corresponding UIFs.
- (**DML estimator**) We introduced a general purpose causal estimators achieving ***doubly robustness*** and ***debiasedness*** properties based on the derived UIF.