Unified Covariate Adjustment:

Estimating Multilinear Causal Estimands

Yonghan Jung

School of Information Science (iSchool)

UIUC

www.yonghanjung.me

2025 INFORMS

NeurIPS-24
Unified covariate adjustment for causal inference

Yonghan Jung
Thesis (2025)

Working papers

Unified Covariate Adjustment:

Estimating Multilinear Causal Estimands

Yonghan Jung

School of Information Science (iSchool)

UIUC

www.yonghanjung.me

2025 INFORMS

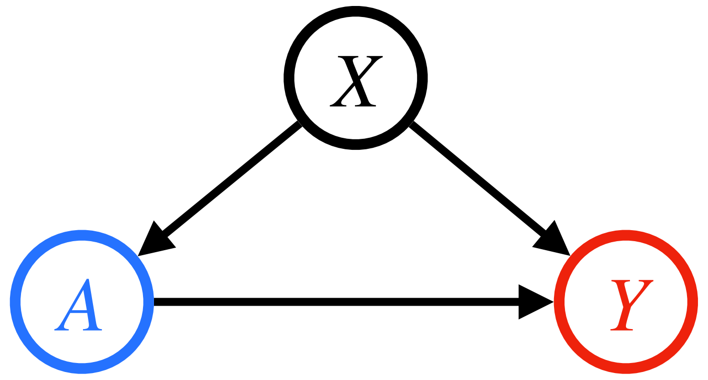
Motivation: Multilinear Causal Estimands

A causal effect $\mathbb{E}[Y \mid \mathrm{do}(\overline{\mathbf{a}})]$ is often identified as a multilinear functional.

Motivation: Multilinear Causal Estimands

A causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is often identified as a multilinear functional.

Back-Door (BD)

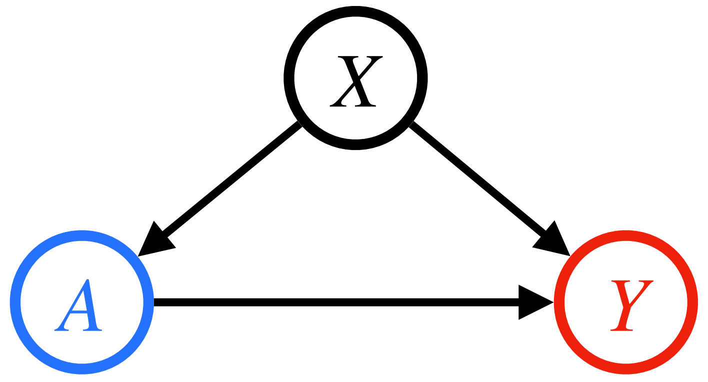


$$\sum_x \mathbb{E}[Y \mid \bar{a}, x] P(x)$$

Motivation: Multilinear Causal Estimands

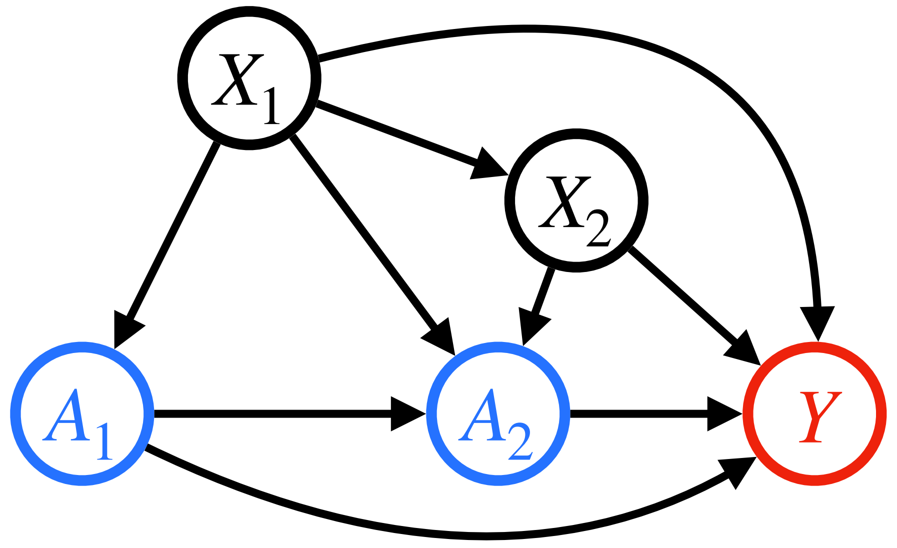
A causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is often identified as a multilinear functional.

Back-Door (BD)



$$\sum_x \mathbb{E}[Y \mid \bar{a}, x]P(x)$$

Sequential BD (SBD)

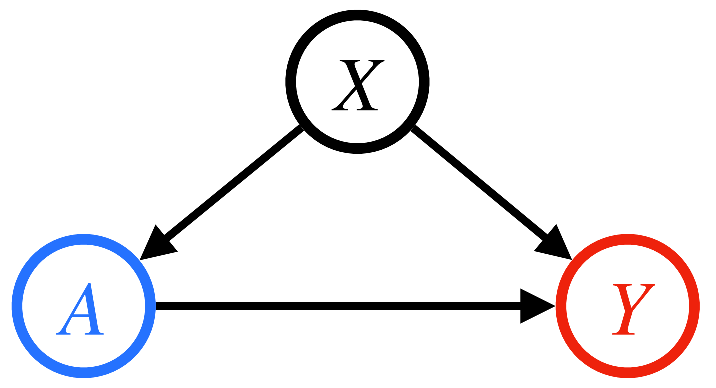


$$\sum_{x_1,x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1]P(x_2 \mid \bar{a}_1, x_1)P(x_1)$$

Motivation: Multilinear Causal Estimands

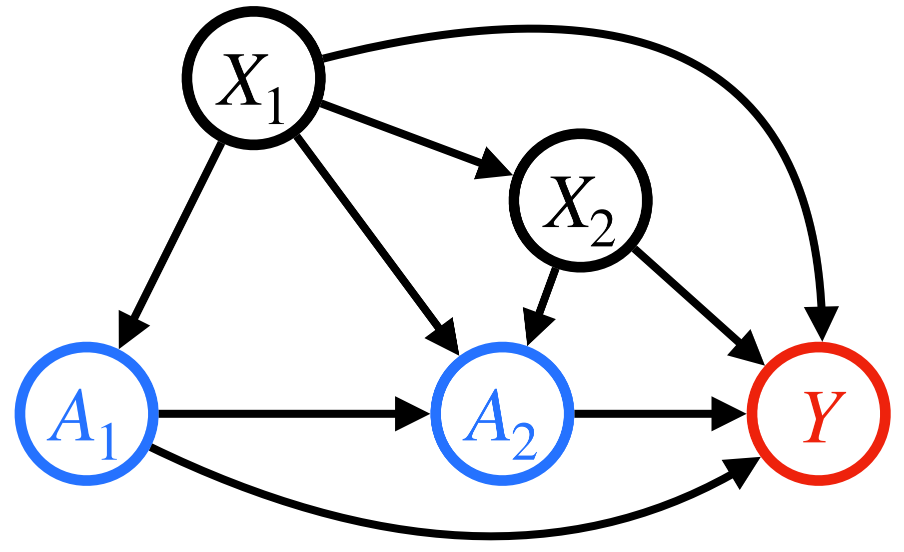
A causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is often identified as a multilinear functional.

Back-Door (BD)



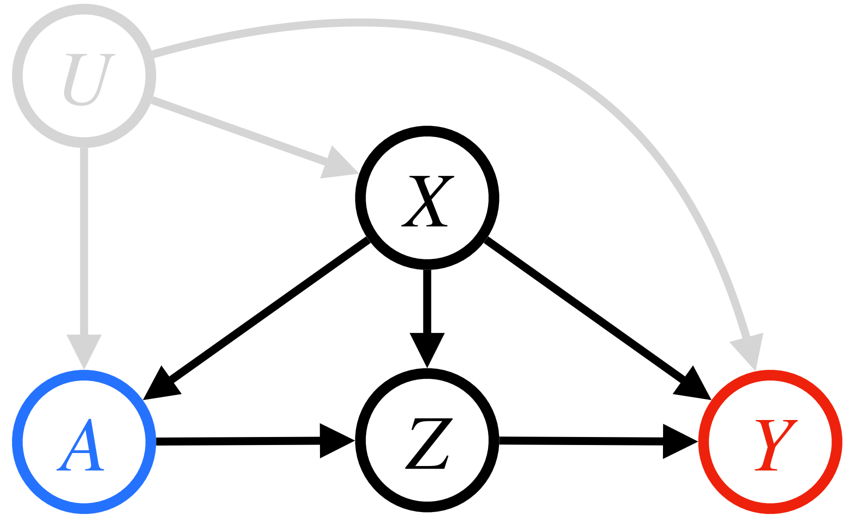
$$\sum_x \mathbb{E}[Y \mid \bar{a}, x]P(x)$$

Sequential BD (SBD)



$$\sum_{x_1,x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1]P(x_2 \mid \bar{a}_1, x_1)P(x_1)$$

Front-Door (FD)

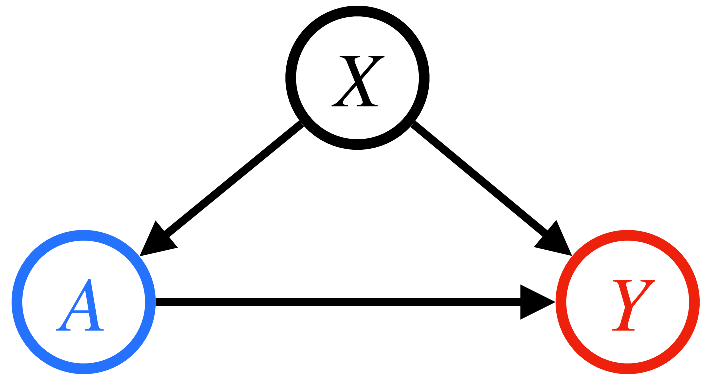


$$\sum_{x,a} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(ax)$$

Motivation: Multilinear Causal Estimands

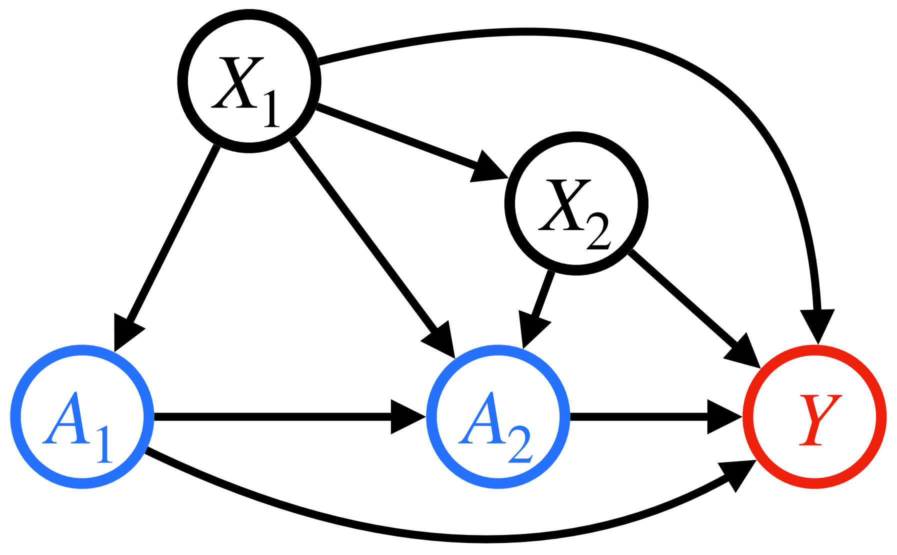
A causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is often identified as a multilinear functional.

Back-Door (BD)



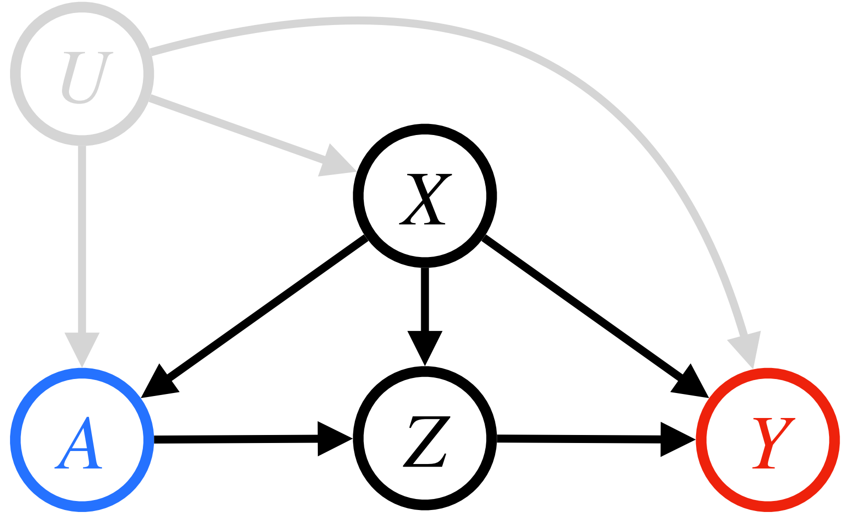
$$\sum_x \mathbb{E}[Y \mid \bar{a}, x]P(x)$$

Sequential BD (SBD)



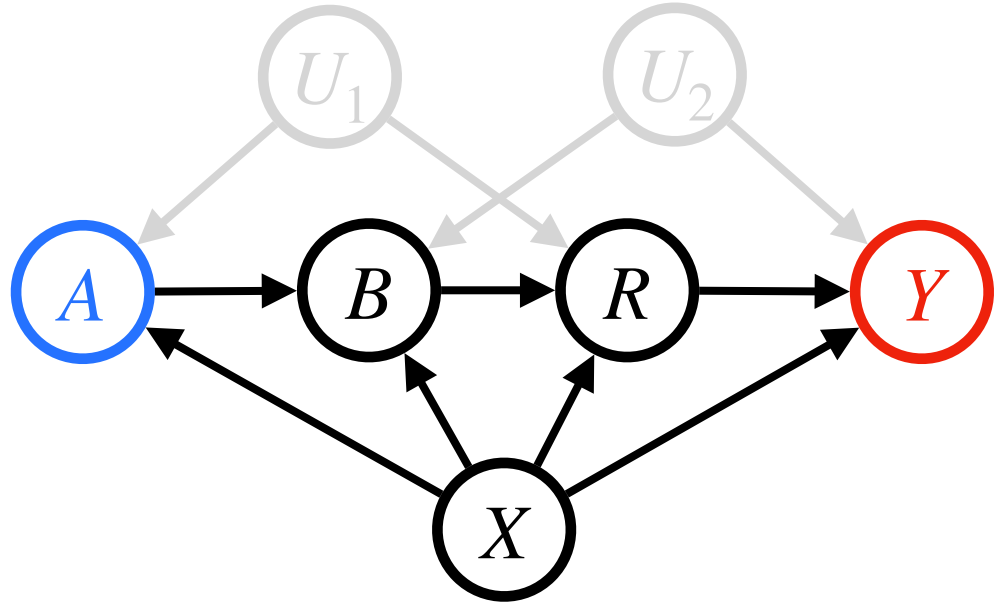
$$\sum_{x_1,x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1]P(x_2 \mid \bar{a}_1, x_1)P(x_1)$$

Front-Door (FD)



$$\sum_{x,a} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(ax)$$

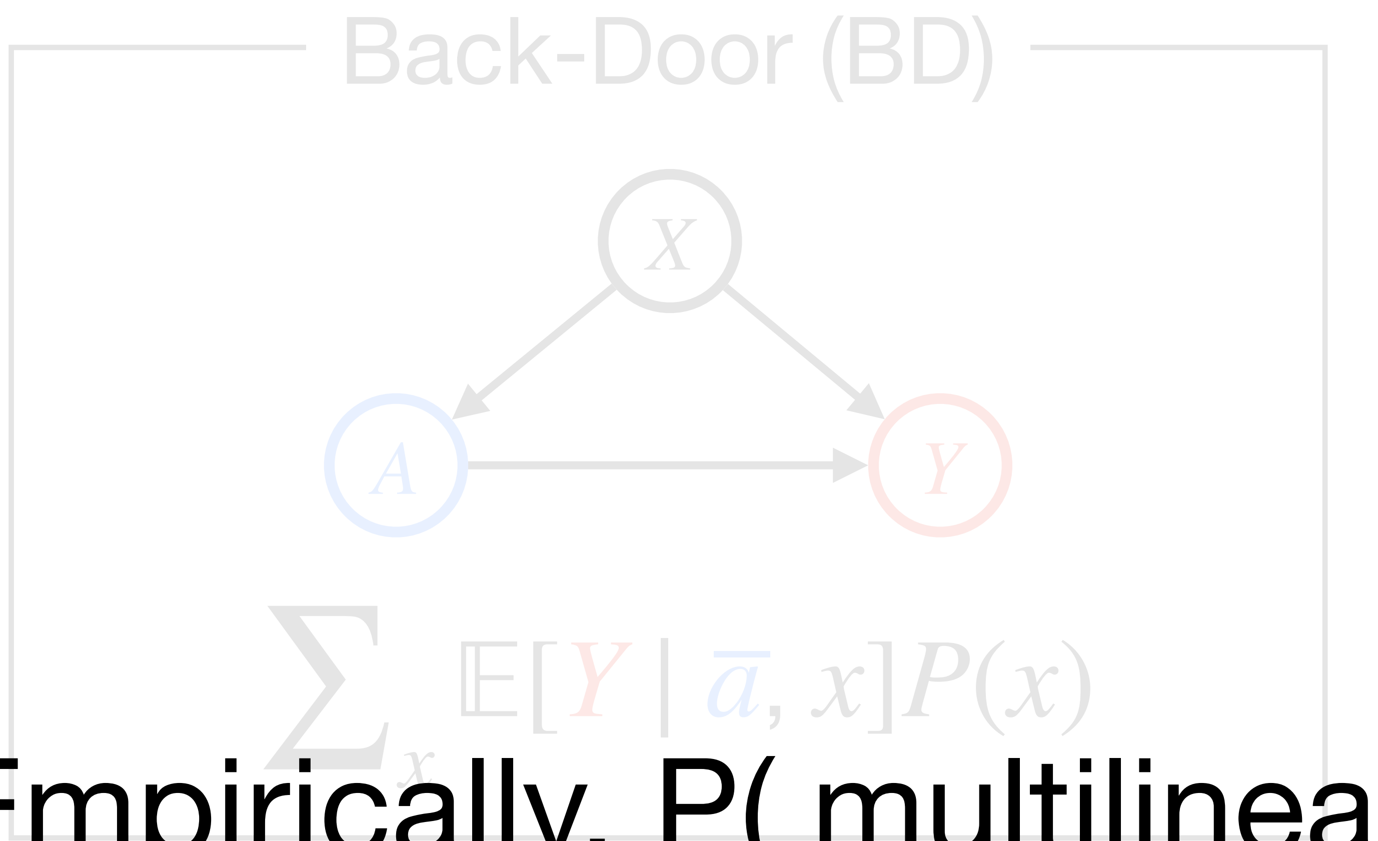
Verma

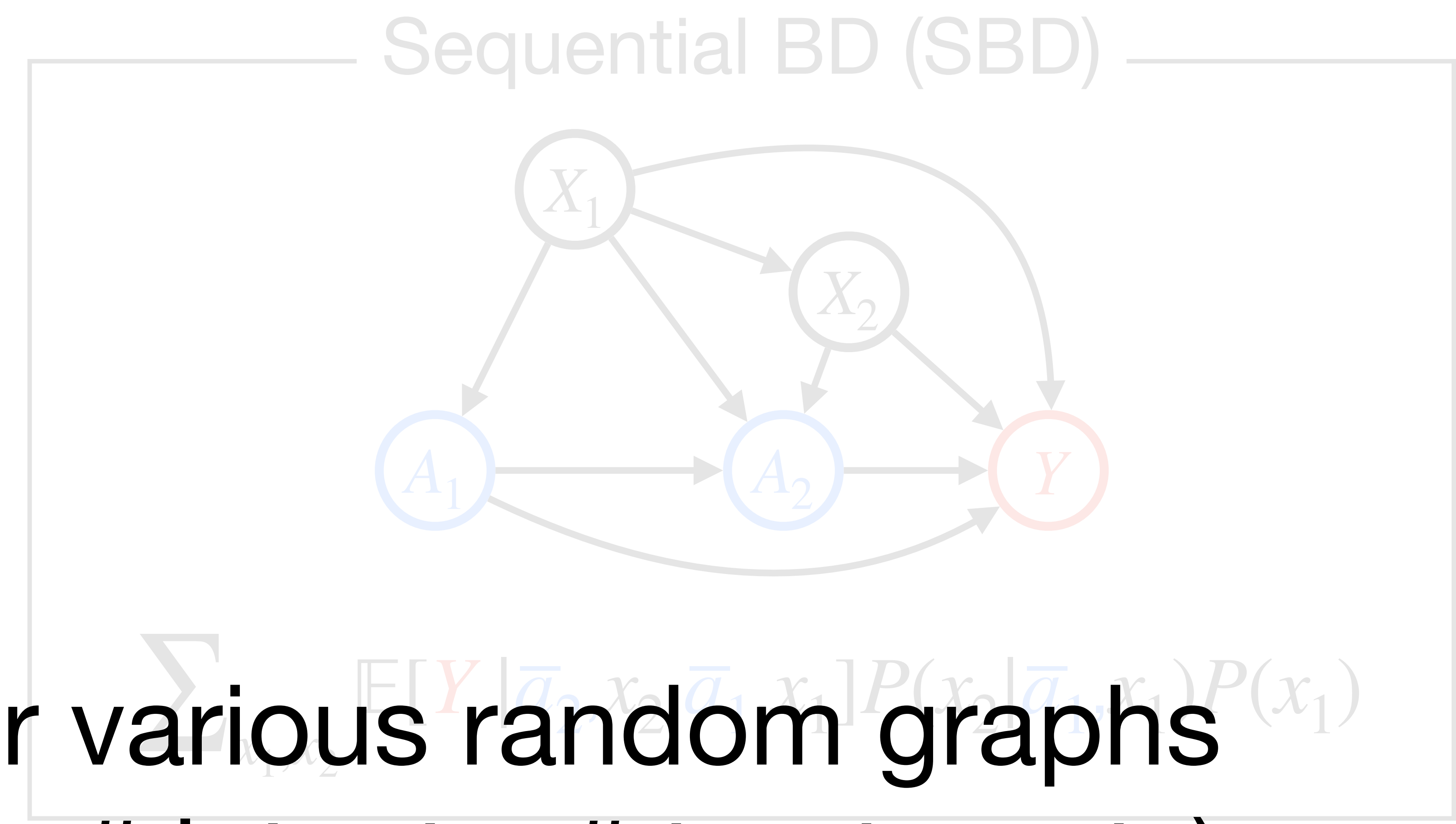


$$\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x]P(r \mid bax)P(b \mid \bar{a}x)P(ax)$$

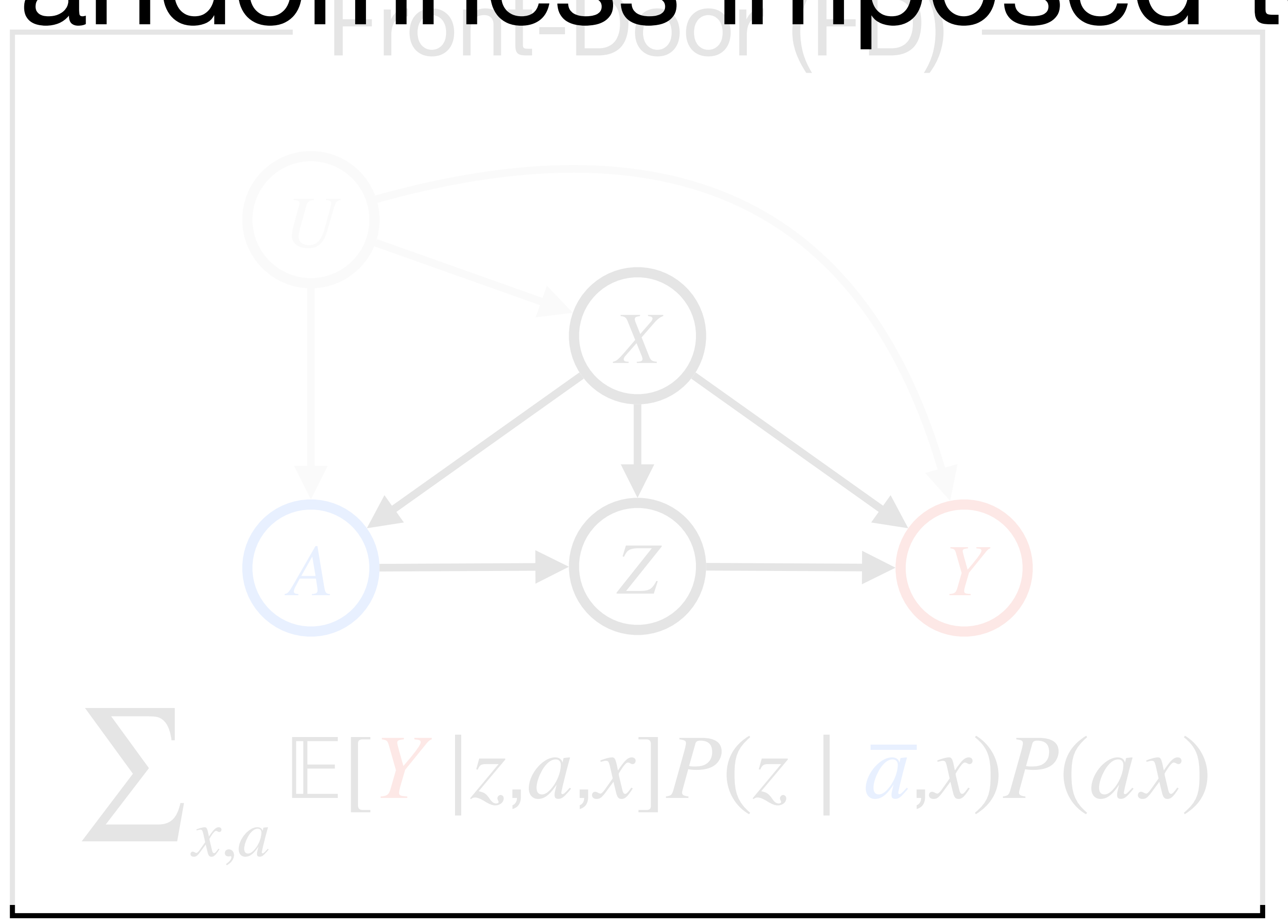
Motivation: Multilinear Causal Estimands

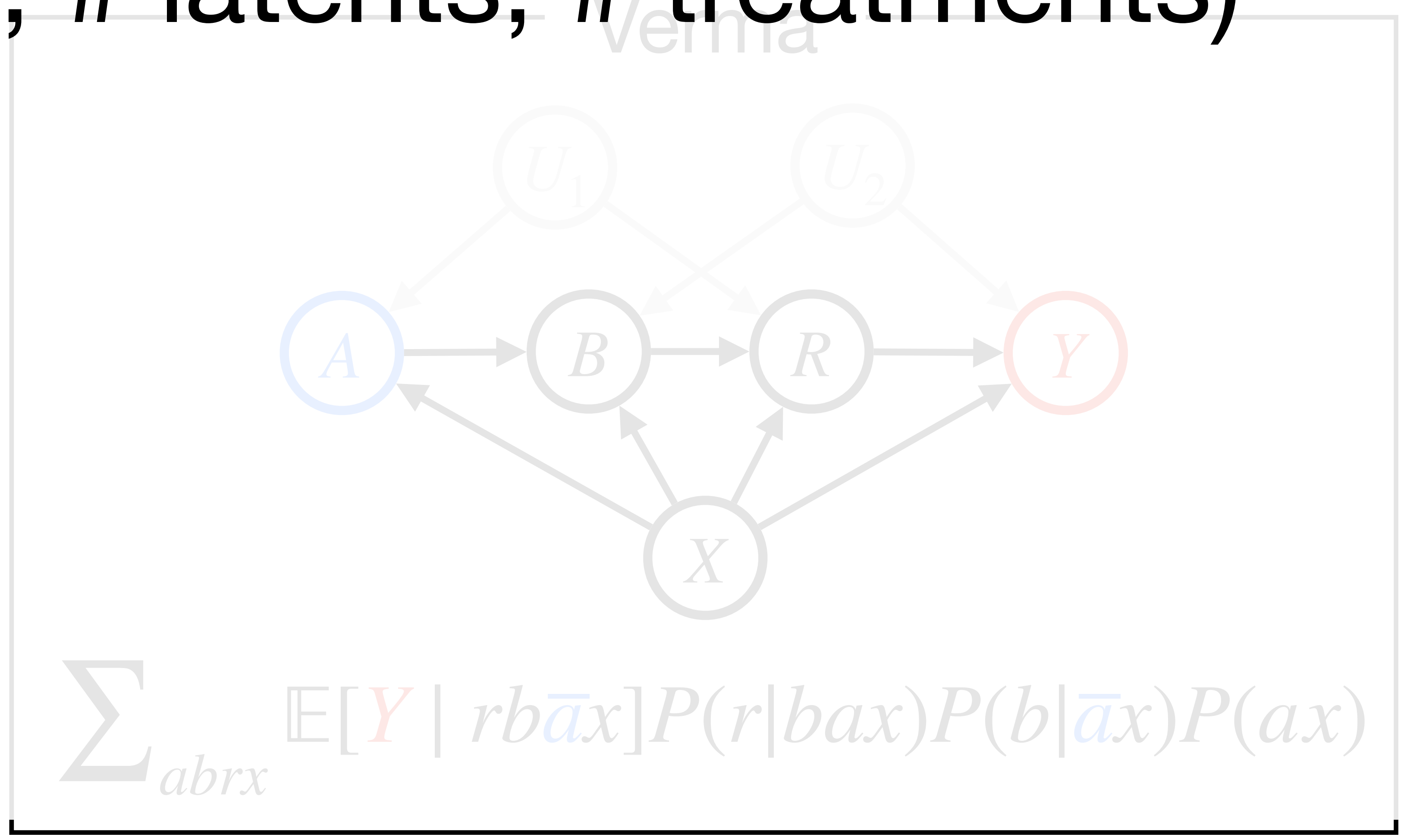
A causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is often identified as a multilinear functional.





Empirically, P(multilinear | ID) > 99% for various random graphs
(randomness imposed to # observables, # latents, # treatments)





Tasks

Tasks

1. **Identification**: Sound and complete graphical criterion for identifying causal effects as a multilinear causal estimand

Tasks

❶ **Identification**: Sound and complete graphical criterion for identifying causal effects as a multilinear causal estimand

❷ **Computationally efficiency**: A new formulation for multilinear causal estimands with computational efficiency.

Tasks

❶ **Identification**: Sound and complete graphical criterion for identifying causal effects as a multilinear causal estimand

❷ **Computationally efficiency**: A new formulation for multilinear causal estimands with computational efficiency.

❸ **Sample efficiency:** A doubly robust and sample efficient estimation framework

Identification Criterion for Multilniear Estimands

Identification Criterion for Multilniear Estimands

Implication of Multilinear Causal Estimands Criterion

Suppose a causal effect $\mathbb{E}[Y \mid \mathrm{do}(\overline{\mathbf{a}})]$ is identifiable. Its ID expression is multilinear, iff, it is given as a product of SBDs.

Identification Criterion for Multilniear Estimands

Implication of Multilinear Causal Estimands Criterion

Suppose a causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is identifiable. Its ID expression is multilinear, iff, it is given as a product of SBDs.

Special case when $|\mathbf{A}| = 1$

It's expressible as a multilinear causal estiands (non-ratio), iff, A and its children are not connected by bidirected paths (aka Tian's criterion)

Identification Criterion for Multilniear Estimands

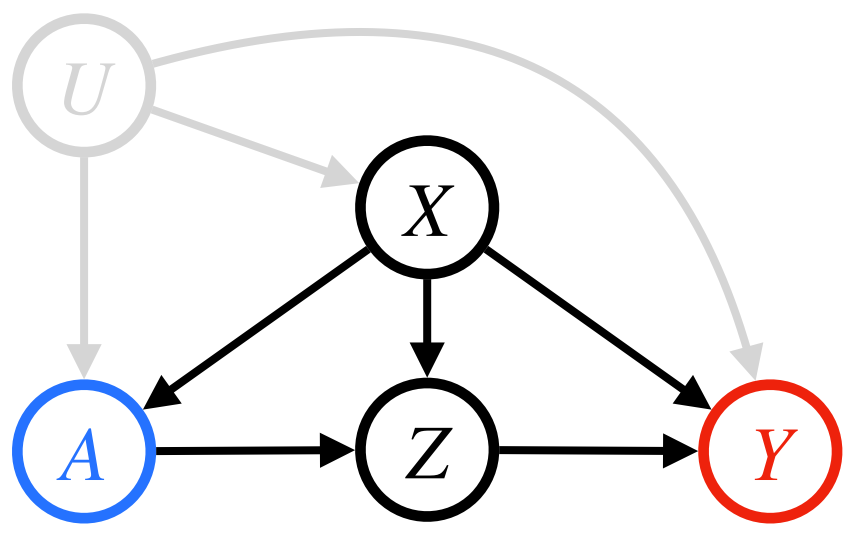
Implication of Multilinear Causal Estimands Criterion

Suppose a causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is identifiable. Its ID expression is multilinear, iff, it is given as a product of SBDs.

Special case when $|\mathbf{A}| = 1$

It's expressible as a multilinear causal estiands (non-ratio), iff, A and its children are not connected by bidirected paths (aka Tian's criterion)

Front-Door (FD)



$$\sum_{x,a} \mathbb{E}[Y \mid z, a, x] P(z \mid \bar{a}, x) P(x)$$

Identification Criterion for Multilniear Estimands

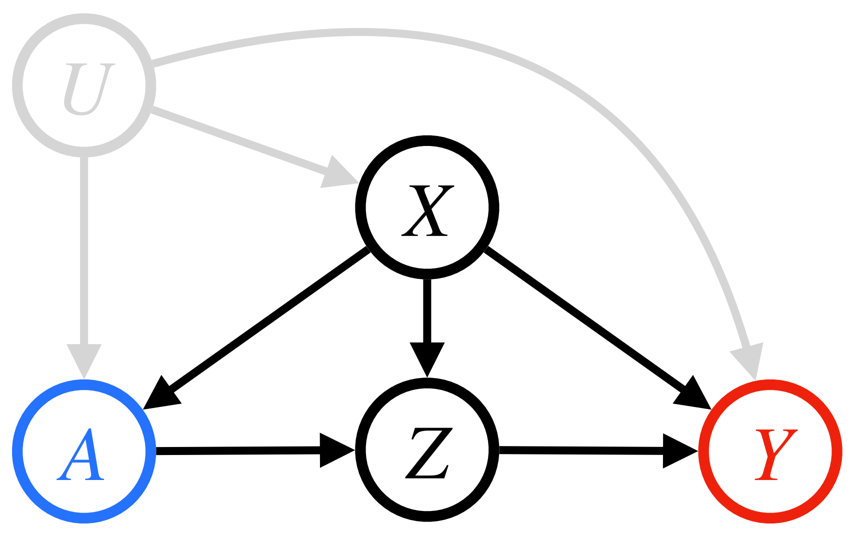
Implication of Multilinear Causal Estimands Criterion

Suppose a causal effect $\mathbb{E}[Y \mid \mathrm{do}(\bar{\mathbf{a}})]$ is identifiable. Its ID expression is multilinear, iff, it is given as a product of SBDs.

Special case when $|\mathbf{A}| = 1$

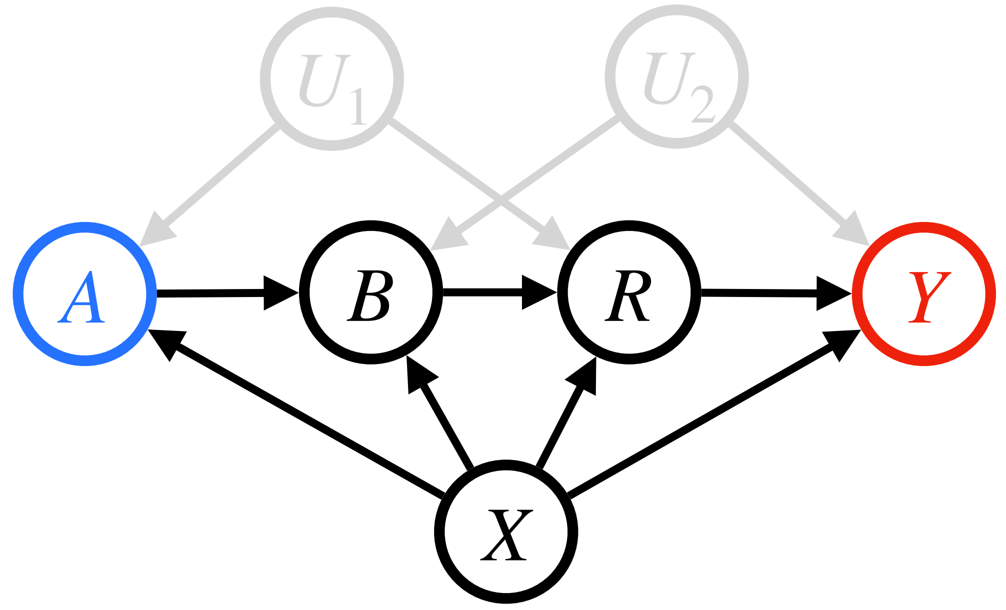
It's expressible as a multilinear causal estiands (non-ratio), iff, A and its children are not connected by bidirected paths (aka Tian's criterion)

Front-Door (FD)



$$\sum_{x,a} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(x)$$

Verma



$$\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x]P(r|bax)P(b|\bar{a}x)P(ax)$$

Nested Conditional Expectation for SBD

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\overline{a}_1, \overline{a}_2)] \triangleq \sum_{x_1, x_2} \mathbb{E}[Y \mid \overline{a}_2, x_2, \overline{a}_1, x_1] P(x_2 | \overline{a}_1, x_1) P(x_1)$$

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, x_1) P(x_1)$$

Evaluating $\sum_{x_1, x_2}$ is computationally expensive, but can be circumvented by nested expectation.

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, x_1) P(x_1)$$

1. $\mu_2(A_2, X_2, A_1, X_1) \triangleq \mathbb{E}[Y \mid A_2, X_2, A_1, X_1]$

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, x_1) P(x_1)$$

1. $\mu_2(A_2, X_2, A_1, X_1) \triangleq \mathbb{E}[Y \mid A_2, X_2, A_1, X_1]$

2. $\mu_1(A_1, X_1) \triangleq \mathbb{E}[\mu_2(\bar{a}_2, X_2, A_1, X_1) \mid A_1, X_1]$

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, x_1) P(x_1)$$

1. $\mu_2(A_2, X_2, A_1, X_1) \triangleq \mathbb{E}[Y \mid A_2, X_2, A_1, X_1]$

2. $\mu_1(A_1, X_1) \triangleq \mathbb{E}[\mu_2(\bar{a}_2, X_2, A_1, X_1) \mid A_1, X_1]$

$\boxed{\sum_{x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, A_1, X_1] P(x_2 \mid A_1, X_1)}$

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, x_1) P(x_1)$$

1. $\mu_2(A_2, X_2, A_1, X_1) \triangleq \mathbb{E}[Y \mid A_2, X_2, A_1, X_1]$

2. $\mu_1(A_1, X_1) \triangleq \mathbb{E}[\mu_2(\bar{a}_2, X_2, A_1, X_1) \mid A_1, X_1]$

$$\sum_{x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, A_1, X_1] P(x_2 \mid A_1, X_1)$$

3. $\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] = \mathbb{E}[\mu_1(\bar{a}_1, X_1)]$

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 | \bar{a}_1, x_1) P(x_1)$$

1. $\mu_2(A_2, X_2, A_1, X_1) \triangleq \mathbb{E}[Y \mid A_2, X_2, A_1, X_1]$

2. $\mu_1(A_1, X_1) \triangleq \mathbb{E}[\mu_2(\bar{a}_2, X_2, A_1, X_1) \mid A_1, X_1]$ $\quad \sum_{x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, A_1, X_1] P(x_2 \mid A_1, X_1)$

3. $\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] = \mathbb{E}[\mu_1(\bar{a}_1, X_1)]$ $\quad \sum_{x_1, x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, X_1) P(x_1)$

Nested Conditional Expectation for SBD

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] \triangleq \boxed{\sum_{x_1, x_2}} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, x_1) P(x_1)$$

1. $\mu_2(A_2, X_2, A_1, X_1) \triangleq \mathbb{E}[Y \mid A_2, X_2, A_1, X_1]$

2. $\mu_1(A_1, X_1) \triangleq \mathbb{E}[\mu_2(\bar{a}_2, X_2, A_1, X_1) \mid A_1, X_1]$ — $\sum_{x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, A_1, X_1] P(x_2 \mid A_1, X_1)$

3. $\mathbb{E}[Y \mid \mathrm{do}(\bar{a}_1, \bar{a}_2)] = \mathbb{E}[\mu_1(\bar{a}_1, X_1)]$ — $\sum_{x_1, x_2} \mathbb{E}[Y \mid \bar{a}_2, x_2, \bar{a}_1, x_1] P(x_2 \mid \bar{a}_1, X_1) P(x_1)$

The SBD estimand can be estimated in a computationally efficient manner using nested conditional expectations.

Limitation of Nested Expectation: FD

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \overline{a},x) P(x)$

Limitation of Nested Expectation: FD

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

❶ $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

Limitation of Nested Expectation: FD

Front-Door (FD): $\mathbb{E}[Y \mid do(\bar{a})] \triangleq \sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

❶ $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

❷ $\mu_1(A,X) \triangleq \mathbb{E}[\mu_2(Z, A, X) \mid A,X]$

$\sum_z \mathbb{E}[Y \mid z,A,X] P(z \mid A,X)$

Limitation of Nested Expectation: FD

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \overline{a},x) P(x)$

❶ $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

❷ $\mu_1(A,X) \triangleq \mathbb{E}[\mu_2(Z, A, X) \mid A,X]$

$\sum_z \mathbb{E}[Y \mid z,A,X] P(z \mid A,X)$

❸ $\mathbb{E}[\mu_1(\overline{a}, X)] \neq \mathbb{E}[Y \mid \mathrm{do}(\overline{a})]$

$\sum_{z,x} \mathbb{E}[Y \mid z,\overline{a},x] P(z \mid \overline{a},x) P(x)$

Limitation of Nested Expectation: FD

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

❶ $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

❷ $\mu_1(A,X) \triangleq \mathbb{E}[\mu_2(Z, A, X) \mid A,X]$

$\sum_z \mathbb{E}[Y \mid z,A,X] P(z \mid A,X)$

❸ $\mathbb{E}[\mu_1(\bar{a}, X)] \neq \mathbb{E}[Y \mid \mathrm{do}(\bar{a})]$

$\sum_{z,x} \mathbb{E}[Y \mid z,\bar{a},x] P(z \mid \bar{a},x) P(x)$

The standard nested expectation cannot represent multilinear estimands when treatments are both marginalized and fixed simultaneously.

Limitation of Nested Expectation: Multilinear Estimand

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a}, x) P(x)$

Verma $\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x] P(r \mid bax) P(b \mid \bar{a}x) P(ax)$

Treatments $\mathbf{A}$ are fixed to $\bar{\mathbf{a}}$ and marginalized $\mathbf{a}$ simultaneously.

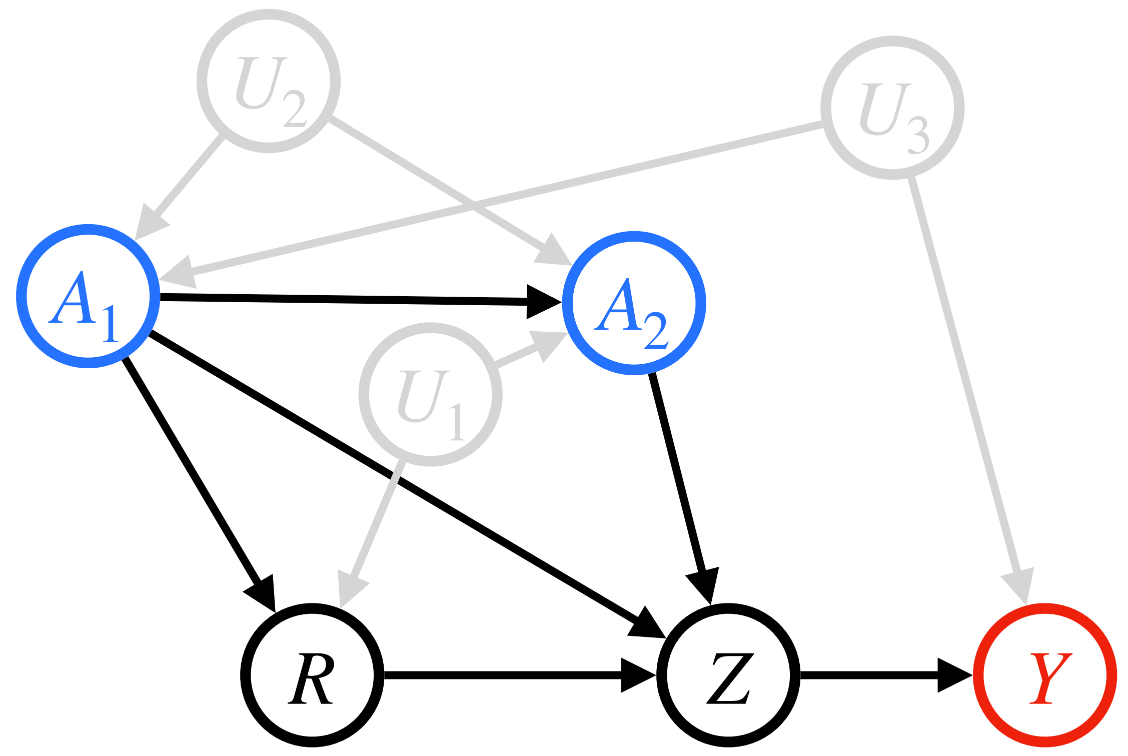
Limitation of Nested Expectation: Multilinear Estimand

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

Verma $\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x] P(r|bax) P(b|\bar{a}x) P(ax)$

Treatments $\mathbf{A}$ are fixed to $\bar{\mathbf{a}}$ and marginalized $\mathbf{a}$ simultaneously.

Example



$\sum_{r,z,a_1,a_2,r'} \mathbb{E}[Y \mid z, r, a_2, a_1] P(z \mid r,\bar{a}_1,\bar{a}_2) P(r \mid \bar{a}_1) P(r,a_1,a_2)$

$\exists$variable that is marginalized multiple times.

Kernel Policy Product: Representation of MCE

Kernel Policy Product: Representation of MCE

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

Kernel Policy Product: Representation of MCE

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

Kernel Policy Product: Representation of MCE

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

Copied Proxy: $\dot{A}$ is an *independent* copy of A s.t.
$P(Z|A,X) = P(Z|\dot{A},X)$.

Kernel Policy Product: Representation of MCE

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(x)$

= Expectation of Y over $P(Y \mid Z,A,X)P(Z \mid \dot{A},X)\mathbb{I}(\dot{A}=\bar{a})P(A,X)$

Copied Proxy: $\dot{A}$ is an *independent* copy of A s.t.
$P(Z|A,X) = P(Z|\dot{A},X)$.

Verma $\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x]P(r|bax)P(b|\bar{a}x)P(ax)$

Kernel Policy Product: Representation of MCE

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

= Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

Copied Proxy: $\dot{A}$ is an *independent* copy of A s.t.
$P(Z \mid A,X) = P(Z \mid \dot{A},X)$.

Verma $\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x] P(r|bax) P(b|\bar{a}x) P(ax)$

= Expectation of Y over $P(Y \mid R,B,A,X) \mathbb{I}(A=\bar{a}) P(R \mid B,\dot{A},X) P(B \mid A,X) P(\dot{A},X)$

Kernel Policy Product: Representation of MCE

Front-Door (FD) $\sum_{x,a} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

Copied Proxy: $\dot{A}$ is an *independent* copy of A s.t. $P(Z|A,X) = P(Z|\dot{A},X)$.

Kernel Policy Product: A product of conditional distributions and policies over variables & their copied proxies

Verma $\sum_{abrx} \mathbb{E}[Y \mid rb\bar{a}x] P(r|bax) P(b|\bar{a}x) P(ax)$

$=$ Expectation of Y over $P(Y \mid R,B,A,X) \mathbb{I}(A=\bar{a}) P(R \mid B,\dot{A},X) P(B \mid A,X) P(\dot{A},X)$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x]P(z \mid \overline{a},x)P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X)P(Z \mid \dot{A},X)\mathbb{I}(\dot{A}=\overline{a})P(A,X)$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \overline{a},x) P(x)$

$= \text{Expectation of } Y \text{ over } P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\overline{a}) P(A,X)$

❶ Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

1. Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

2. Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \text{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

❶ Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

❷ Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

❸ Learn $\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

❶ Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

❷ Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

❸ Learn $\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$ $\boxed{= \sum_z \mu_2(z,\dot{A},X) P(z \mid A,\dot{A},X) = \sum_z \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X] P(z \mid A,X)}$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

❶ Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

❷ Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

❸ Learn $\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$ $\boxed{= \sum_z \mu_2(z,\dot{A},X) P(z \mid A,\dot{A},X) = \sum_z \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X] P(z \mid A,X)}$

❹ Evaluate μ_1 on $(\bar{a}, A, X)$ to have $\mu_1(\bar{a}, A, X)$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

1. Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

2. Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

3. Learn $\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$ $\boxed{= \sum_z \mu_2(z,\dot{A},X) P(z \mid A,\dot{A},X) = \sum_z \mathbb{E}[Y \mid Z,A \leftarrow \dot{A},X] P(z \mid A,X)}$

4. Evaluate μ_1 on $(\bar{a}, A, X)$ to have $\mu_1(\bar{a}, A, X)$ $\boxed{= \sum_z \mathbb{E}[Y \mid Z,A,X] P(z \mid \bar{a},X)}$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

1. Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

2. Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

3. Learn $\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$ $\boxed{= \sum_z \mu_2(z,\dot{A},X) P(z \mid A,\dot{A},X) = \sum_z \mathbb{E}[Y \mid Z,A \leftarrow \dot{A},X] P(z \mid A,X)}$

4. Evaluate μ_1 on $(\bar{a}, A, X)$ to have $\mu_1(\bar{a}, A, X)$ $\boxed{= \sum_z \mathbb{E}[Y \mid Z,A,X] P(z \mid \bar{a},X)}$

5. $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] = \mathbb{E}[\mu_1(\bar{a}, A, X)]$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$

$= \text{Expectation of } Y \text{ over } P(Y \mid Z,A,X) P(Z \mid \dot{A},X) \mathbb{I}(\dot{A}=\bar{a}) P(A,X)$

1. Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

2. Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A). $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

3. Learn $\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$ $\boxed{= \sum_z \mu_2(z,\dot{A},X) P(z \mid A,\dot{A},X) = \sum_z \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X] P(z \mid A,X)}$

4. Evaluate μ_1 on $(\bar{a}, A, X)$ to have $\mu_1(\bar{a}, A, X)$ $\boxed{= \sum_z \mathbb{E}[Y \mid Z,A,X] P(z \mid \bar{a},X)}$

5. $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] = \mathbb{E}[\mu_1(\bar{a}, A, X)]$ $\boxed{= \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(a, x)}$

Computationally Efficiency Gain

Front-Door (FD): $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(x)$

$=$ Expectation of Y over $P(Y \mid Z,A,X)P(Z \mid \dot{A},X)\mathbb{I}(\dot{A}=\bar{a})P(A,X)$

1. Learn $\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$

2. Evaluate μ_2 on $(Z, \dot{A}, X)$ ($\dot{A}$ is a copied proxy of A): $\mu_2(Z, \dot{A}, X) := \mathbb{E}[Y \mid Z, A \leftarrow \dot{A}, X]$

Computational efficiency gain via replacing $\sum_{z,a,x}$ through KPP expression with copied proxies and nested conditional expectations

3. Learn $\mu_1(\dot{A}, A, X) \triangleq \mathbb{E}[\mu_2(Z, \dot{A}, X) \mid \dot{A}, A, X]$

4. Evaluate μ_1 on $(\bar{a}, A, X)$ to have $\mu_1(\bar{a}, A, X)$ $= \sum_z \mathbb{E}[Y \mid Z,A,X]P(z \mid \bar{a},X)$

5. $\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] = \mathbb{E}[\mu_1(\bar{a}, A, X)]$ $= \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(a,x)$

Sample Efficient Estimation

Sample Efficient Estimation

❶ DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

Sample Efficient Estimation

1 DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \overline{a},x) P(x)$$

Sample Efficient Estimation

1. DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \overline{a},x) P(x)$$

$$\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$$

Sample Efficient Estimation

❶ DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \overline{a},x) P(x)$$

$\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$ $\pi_2 > 0$ s.t. $\mathbb{E}[\mu_2 \pi_2] = \mathbb{E}[\mu_2(Z, \dot{A}, X)]$

Sample Efficient Estimation

1 DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid do(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x]P(z \mid \overline{a},x)P(x)$$

$\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$ $\quad\quad$ $\pi_2 > 0$ s.t. $\mathbb{E}[\mu_2\pi_2] = \mathbb{E}[\mu_2(Z, \dot{A}, X)]$

$\mu_1(A,\dot{A},X)=\mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$

Sample Efficient Estimation

1 DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x]P(z \mid \bar{a},x)P(x)$$

$$\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X] \qquad \pi_2 > 0 \text{ s.t. } \mathbb{E}[\mu_2\pi_2] = \mathbb{E}[\mu_2(Z, \dot{A}, X)]$$

$$\mu_1(A,\dot{A},X)=\mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X] \qquad \pi_2 > 0 \text{ s.t. } \mathbb{E}[\mu_1\pi_1] = \mathbb{E}[\mu_1(\bar{a}, A, X)]$$

Sample Efficient Estimation

1. DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid \mathrm{do}(\overline{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x]P(z \mid \overline{a},x)P(x)$$

$\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X]$ $\pi_2 > 0$ s.t. $\mathbb{E}[\mu_2\pi_2] = \mathbb{E}[\mu_2(Z, \dot{A}, X)]$

$\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X]$ $\pi_2 > 0$ s.t. $\mathbb{E}[\mu_1\pi_1] = \mathbb{E}[\mu_1(\bar{a}, A, X)]$

2. UCA($\hat{\mu}$,$\hat{\pi}$) is an AIPW-style estimator tailored to KPP.

Sample Efficient Estimation

1. DML($\boldsymbol{\mu}$,$\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is a set of the 1st-order error correction parameters.

$$\mathbb{E}[Y \mid \mathrm{do}(\bar{a})] \triangleq \sum_{z,a,x} \mathbb{E}[Y \mid z,a,x] P(z \mid \bar{a},x) P(x)$$

$$\mu_2(Z, A, X) \triangleq \mathbb{E}[Y \mid Z, A, X] \qquad \pi_2 > 0 \text{ s.t. } \mathbb{E}[\mu_2\pi_2] = \mathbb{E}[\mu_2(Z, \dot{A}, X)]$$

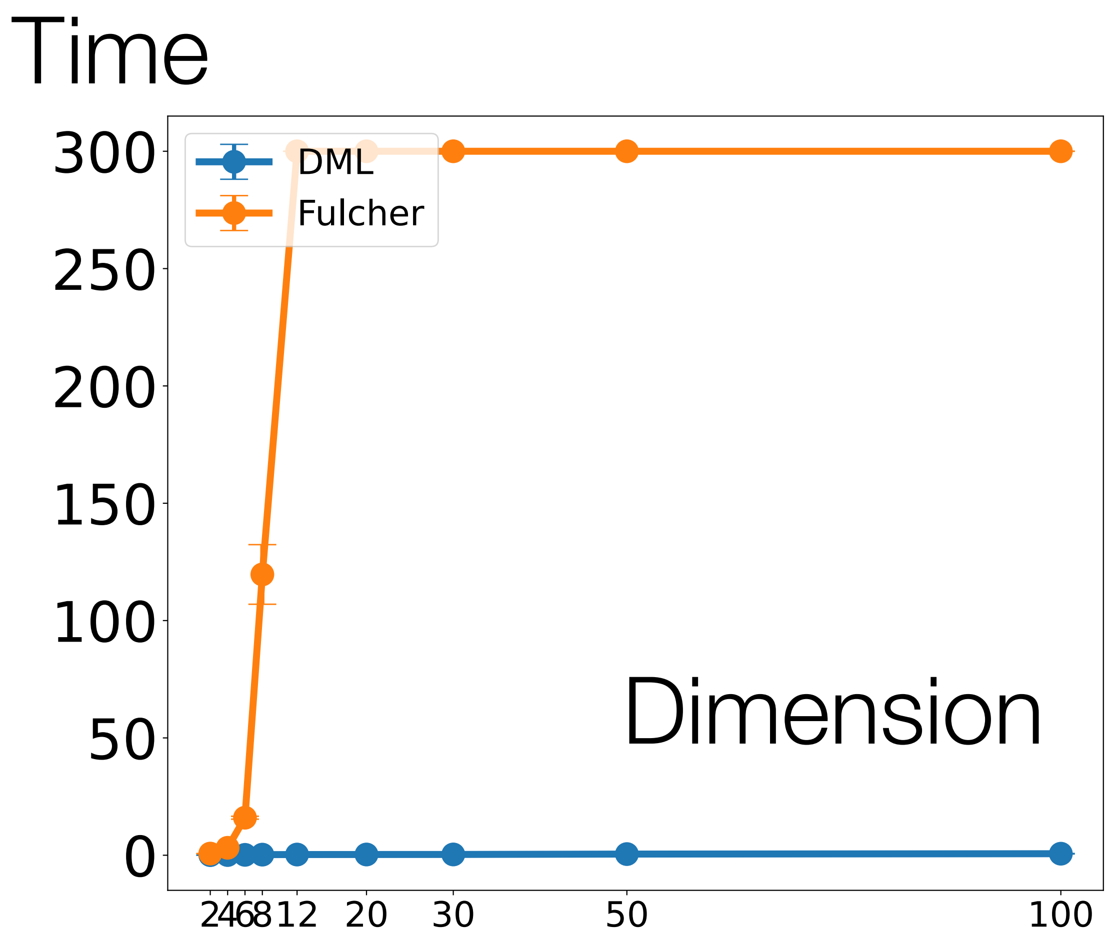
$$\mu_1(A,\dot{A},X) = \mathbb{E}[\mu_2(Z,\dot{A},X) \mid A,\dot{A},X] \qquad \pi_2 > 0 \text{ s.t. } \mathbb{E}[\mu_1\pi_1] = \mathbb{E}[\mu_1(\bar{a}, A, X)]$$

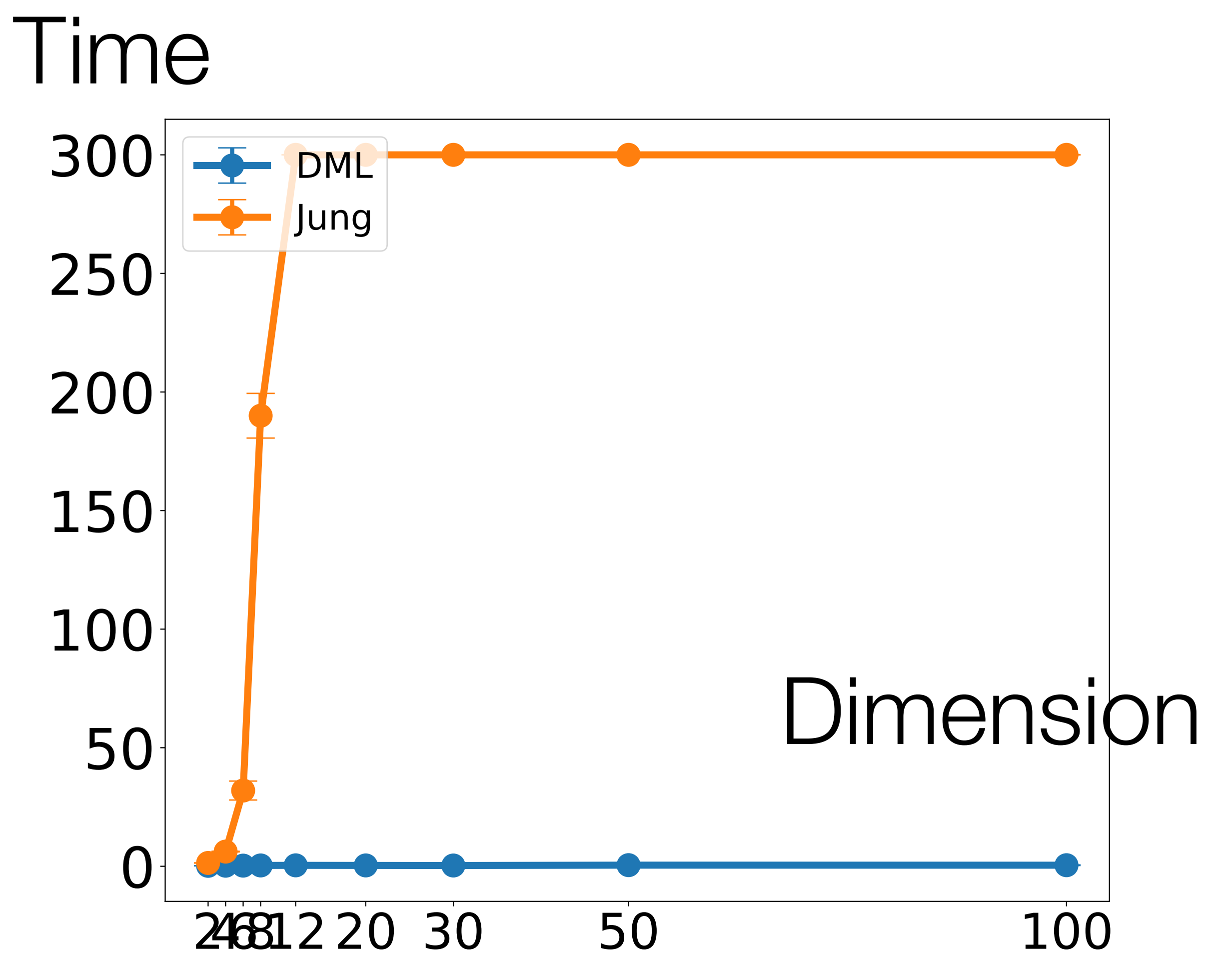
2. UCA($\hat{\mu}$,$\hat{\pi}$) is an AIPW-style estimator tailored to KPP.

$$\mathrm{Error}(\mathrm{UCA}(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\pi}}), \mathrm{DML}(\boldsymbol{\mu},\boldsymbol{\pi})) = \sum_i \mathrm{Error}(\hat{\mu}_i,\mu_i) \times \mathrm{Error}(\hat{\pi}_i,\pi_i)$$

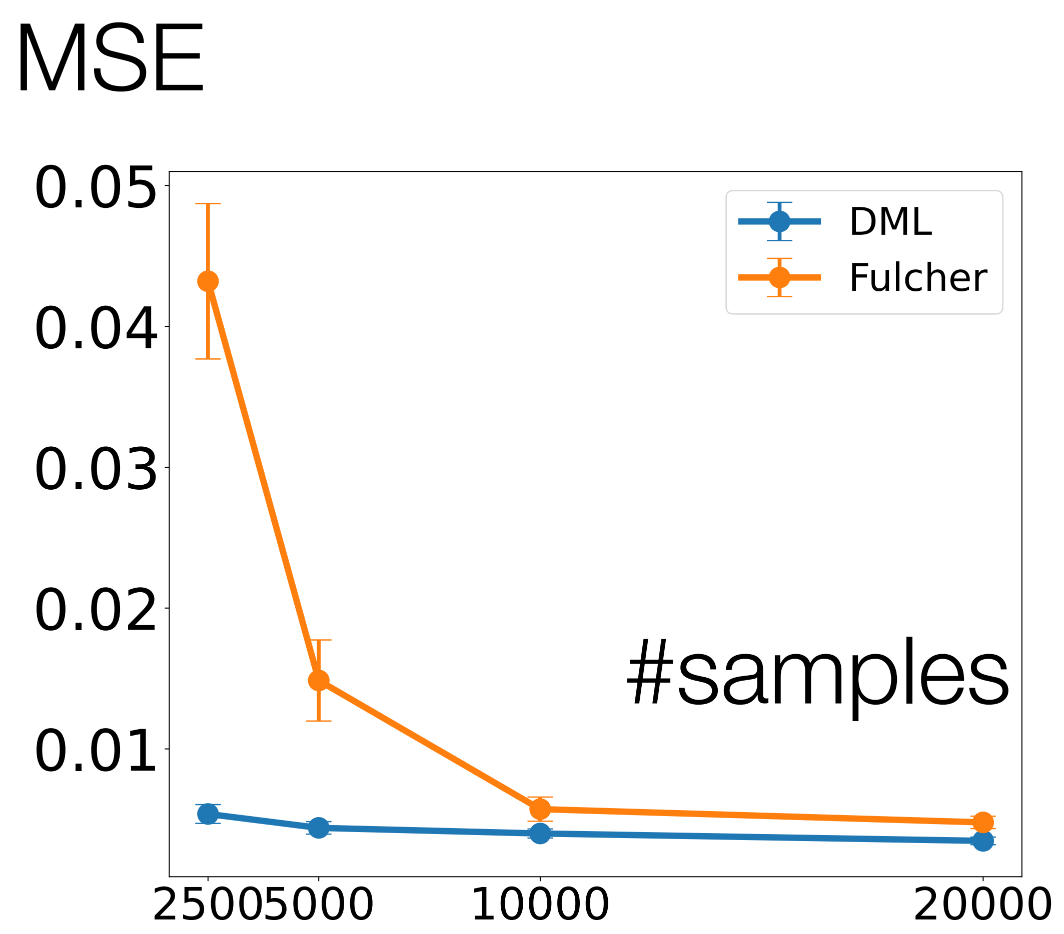
Simulation Results

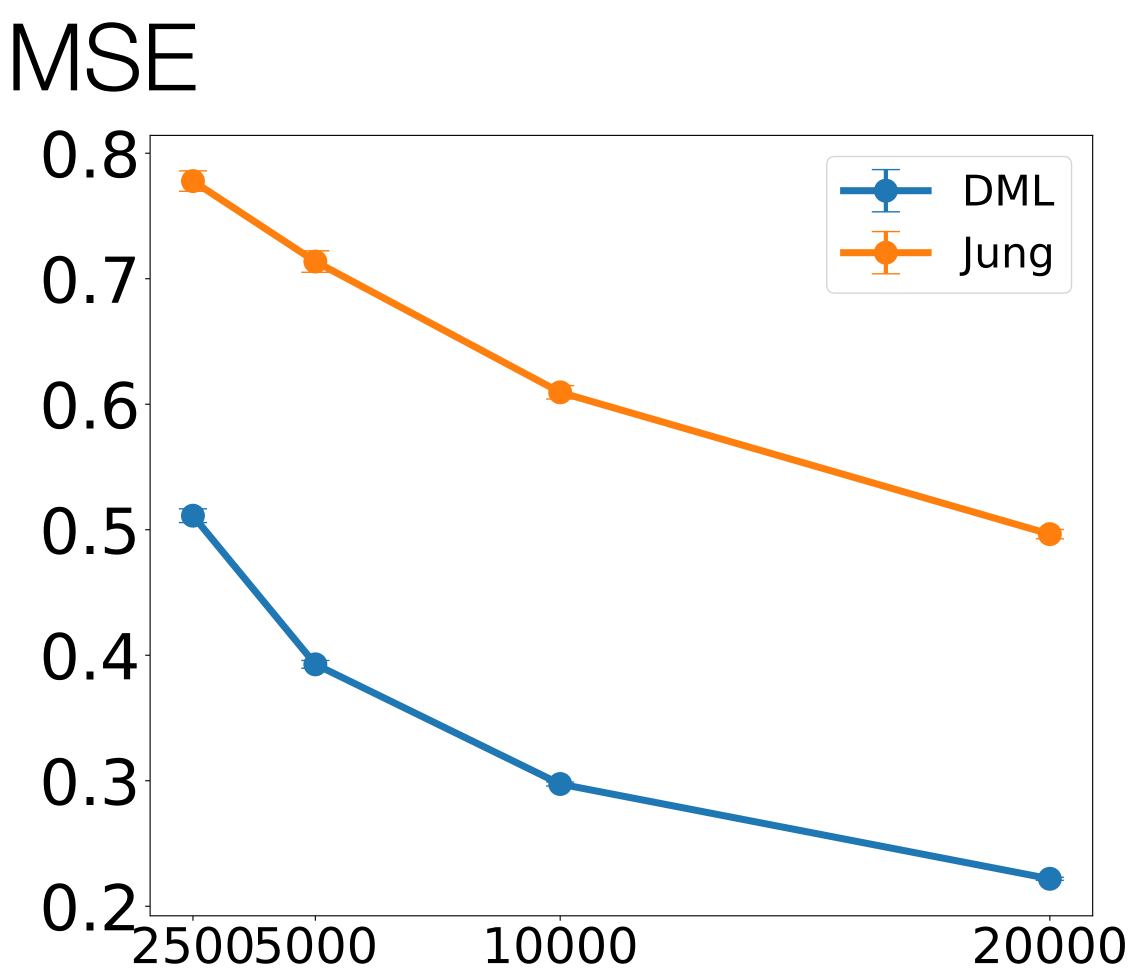
FD | **Verma**





- Existing estimators' evaluation time increase as dimensions increases
- UCA estimator exhibits computational efficiency gains.





- UCA estimator exhibits sample efficiency.

Summary

Summary

1. **Identification**: Sound and complete graphical criterion for identifying causal effects as a multilinear causal estimand

Summary

1. **Identification**: Sound and complete graphical criterion for identifying causal effects as a multilinear causal estimand

2. **Computationally efficiency**: A new formulation for multilinear causal estimands with computational efficiency.

Summary

❶ **Identification**: Sound and complete graphical criterion for identifying causal effects as a multilinear causal estimand

❷ **Computationally efficiency**: A new formulation for multilinear causal estimands with computational efficiency.

❸ **Sample efficiency:** A doubly robust and sample efficient estimation framework