Complete Graphical Criterion for Sequential Covariate Adjustment in Causal Inference

Yonghan Jung[1*] Min Woo Park[2] Sanghack Lee[2*] ([1]Purdue University [2]Seoul National University)

PURDUE UNIVERSITY
Seoul National University Graduate School of Data Science
NEURAL INFORMATION PROCESSING SYSTEMS

Contribution

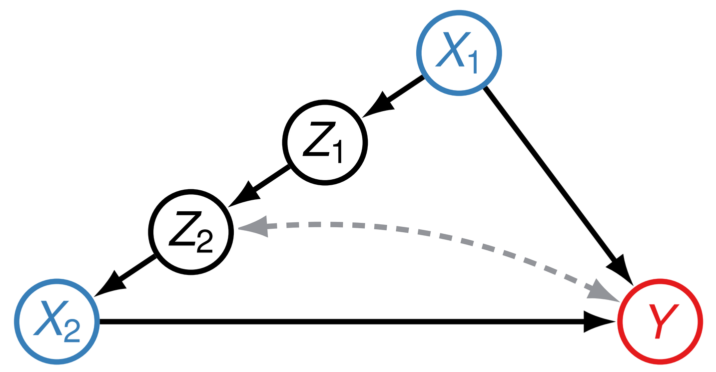
- We presents **Sequential Adjustment Criterion (SAC)**, a **sound and complete** criterion for sequential covariate adjustment.

Comparison with other graphical criteria for covariate adjustments.

Criterion	Static	Sequential	Multi-outcome	Completeness
Back-Door (BD)	✓	✗	N/A	✗
Adjustment Criterion (AC)	✓	✗	N/A	✓
Sequential Back-Door (SBD)	✓	✓	✗	✗
multi-outcome SBD (mSBD)	✓	✓	✓	✗
SAC (Ours)	✓	✓	✓	✓

Motivation & Background

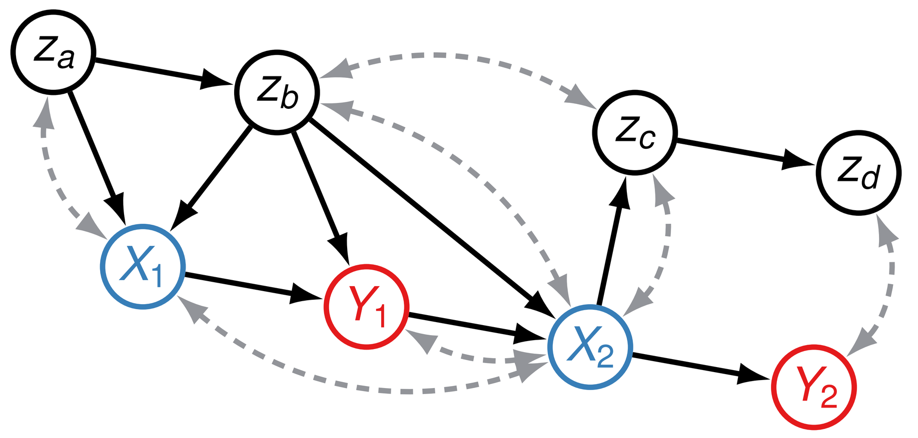
- **Covariate Adjustment (CA)** The causal effect from treatment **X** to outcome **Y** can be expressed in terms of observational data, using covariates **Z**.



- There is a **complete** graphical criteria **Adjustment Criterion (AC)** for CA.

$$P(y \mid do(x_1, x_2)) = \sum_{z_1, z_2} P(y \mid x_1, x_2, z_1, z_2) P(z_1, z_2)$$

- **Sequential Covariate Adjustment (SCA)** is one of the most prevalent methods for estimating multi-outcome causal effects from observational data.
- Previously existing graphical criteria for SCA are **not complete**.
  ⇒ Not satisfied by existing criterion (mSBD),



but identified as SCA:

$$P(y_1, y_2 \mid do(x_1, x_2)) = \sum_{z_a, z_b, z_c, z_d} P(y_2 \mid x_1, x_2, y_1, z_a, z_b, z_c, z_d) P(z_c, z_d, y_1 \mid x_1, z_a, z_b) P(z_a, z_b)$$

Complete Criterion for SCA (Theory)

- **Definition** (Sequential Covariate Adjustment). Let $(\mathbf{X}, \mathbf{Y})$ denote a pair of ordered sets such that $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ and $\mathbf{Y} = (\mathbf{Y}_0, \ldots, \mathbf{Y}_m)$. Let $\mathbf{Z} \subseteq \mathbf{V} \setminus (\mathbf{X} \cup \mathbf{Y})$ denote vertices ordered as $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_m)$. Define $\mathbf{H}_i := \mathbf{X}^{(i)} \cup \mathbf{Y}^{(i)} \cup \mathbf{Z}^{(i)}$.

$$P(\mathbf{y} \mid do(\mathbf{x})) = \sum_{\mathbf{z}} \prod_{j=0}^{m} P(\mathbf{z}_{j+1}, \mathbf{y}_j \mid \mathbf{h}_{j-1}, \mathbf{x}_j, \mathbf{z}_j).$$

◎ Sequential Adjustment Criterion (SAC).

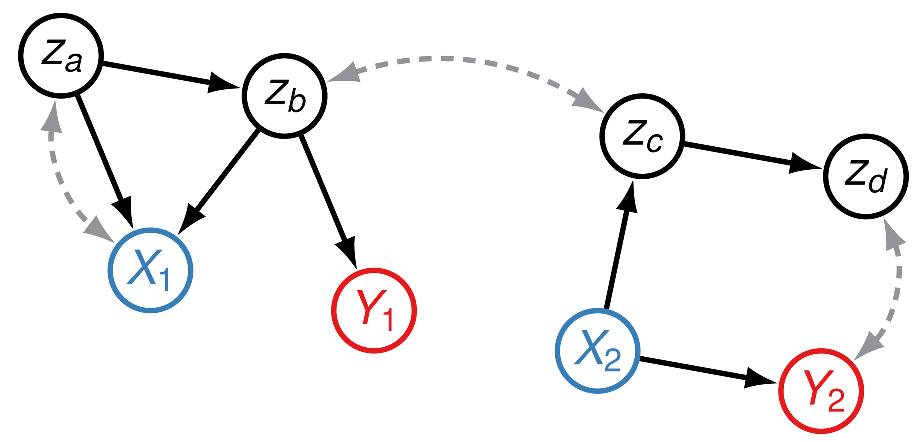
Given $\mathbf{Z} := (\mathbf{Z}_1, \cdots, \mathbf{Z}_m)$ where each $\mathbf{Z}_i$ is non-descendant of $\mathbf{X}^{\geq i+1}$, $\mathbf{Z}$ is said to satisfy sequential adjustment criterion (SAC) if the following conditions are satisfied for $i = 1, \cdots, m$:

1. X_i and $\mathbf{Y}^{\geq i}$ are d-seperated given $\mathbf{Z}_i \cup \mathbf{H}_{i-1}$ in subgraph, *proper sequential backdoor graph* $(\mathcal{G}_{\overline{\mathbf{X}^{\geq i+1}}})^{X_i, \mathbf{Y}^{\geq i}}_{\text{pbd}}$.
2. $\mathbf{Z}_i$ is not a *descendant of proper causal path set* from X_i to $\mathbf{Y}^{\geq i}$.

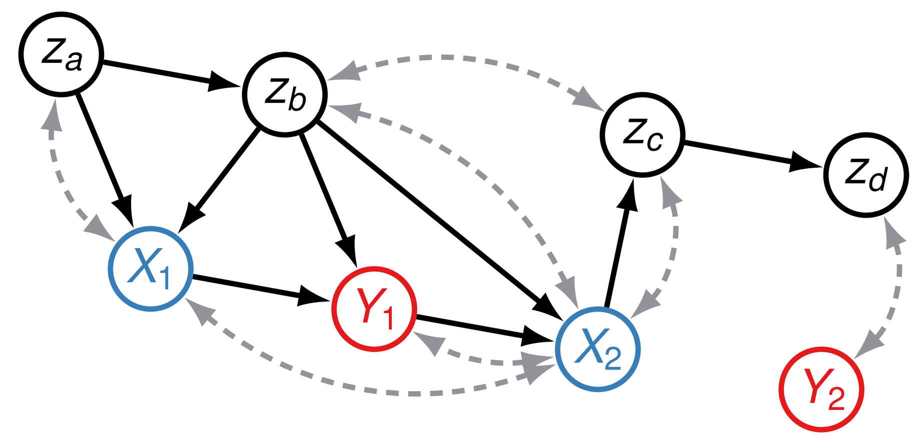
- **Key Example**

Check back-door in subgraph according to the topological order of **X**.

① Subgraph for X_1:
Given $\mathbf{Z}_1 = \{Z_a, Z_b\}$, X_1 and $\{Y_1, Y_2\}$ are **d-seperated** in the subgraph.



② Subgraph for X_2:
Given history $\mathbf{H}_1 = \{Z_a, Z_b, X_1, Y_1\}$ and $\mathbf{Z}_2 = \{Z_c, Z_d\}$, X_2 and Y_2 are **d-seperated** in the subgraph.



∴ $\mathbf{Z} = (\{Z_a, Z_b\}, \{Z_c, Z_d\})$ satisfies SAC.

$$P(y_1, y_2 \mid do(x_1, x_2)) = \sum_{z_a, z_b, z_c, z_d} P(y_2 \mid x_1, x_2, y_1, z_a, z_b, z_c, z_d) P(z_c, z_d, y_1 \mid x_1, z_a, z_b) P(z_a, z_b)$$

✓ **Soundness and Completeness**

Z satisfies SAC ⟺ **Z** is expressible as SCA.

- The soundness and completeness of SAC provide extensive coverage, incorporating existing covariate adjustment criteria.
  - ✓ mSBD ⟹ SAC.
    - The SAC can encompass the mSBD.
  - ✓ Adjustment Criterion (AC) ⟹ SAC.
    - The AC is a special case of the SAC where $m = 1$.
    - $P(\mathbf{y} \mid do(\mathbf{x})) = \sum_{\mathbf{z}} P(\mathbf{y} \mid \mathbf{x}, \mathbf{z}) P(\mathbf{z})$.

Constructive SAC (Algorithm)

Construction of Sequential Adjustment Set:
A method to construct a sequential adjustment set.

- **Input**: A disjoint pair of ordered set $(\mathbf{X}, \mathbf{Y})$ and a causal graph $\mathcal{G}$.
- **Output**: An ordered set $\mathbf{Z}^{\text{an}} := (\mathbf{Z}^{\text{an}}_1, \cdots, \mathbf{Z}^{\text{an}}_m)$.

$\exists (\mathbf{Z}_1, \cdots, \mathbf{Z}_m)$ satisfying SAC ⇔ $\mathbf{Z}^{\text{an}}$ satisfies SAC.

minSCA outputs the smallest subset of $\mathbf{Z}^{\text{an}}$ without sacrificing the validity of the adjustment.

$\mathbf{Z}^{\text{min}} = (\{Z_a, Z_b\}, \emptyset)$ satisfies SAC.

$$P(y_1, y_2 \mid do(x_1, x_2)) = \sum_{z_a, z_b} P(y_2 \mid x_1, x_2, y_1, z_a, z_b) P(y_1 \mid x_1, z_a, z_b) P(z_a, z_b)$$

Conclusion

Sequential Adjustment Criterion (SAC), a sound and complete criterion for sequential covariate adjustment.

Constructive Sequential Adjustment Criterion identifies a set that satisfies the sequential adjustment criterion *if and only if* the causal effect can be expressed as a sequential covariate adjustment.

An algorithm **minSCA** for identifying a minimal sequential covariate adjustment set ensuring that no unnecessary vertices are included.