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)
⭐ Yonghan is on the academic job market! Visit www.yonghanjung.me.

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

Contribution

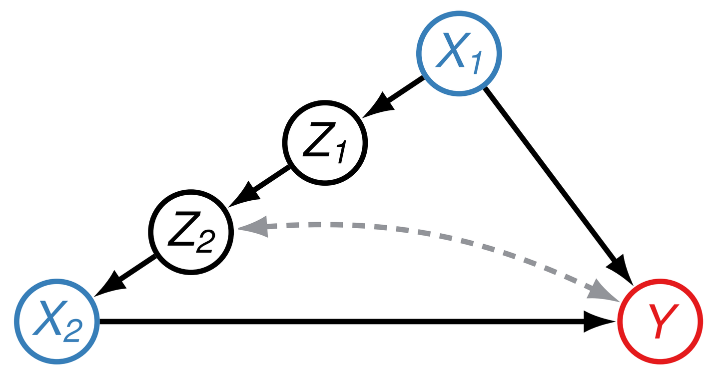
- We present **Sequential Adjustment Criterion (SAC)**, a **sound and complete** criterion for sequential covariate adjustment.

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 $\mathbf{X}$ to outcome $\mathbf{Y}$ can be expressed in terms of observational data, using covariates $\mathbf{Z}$.

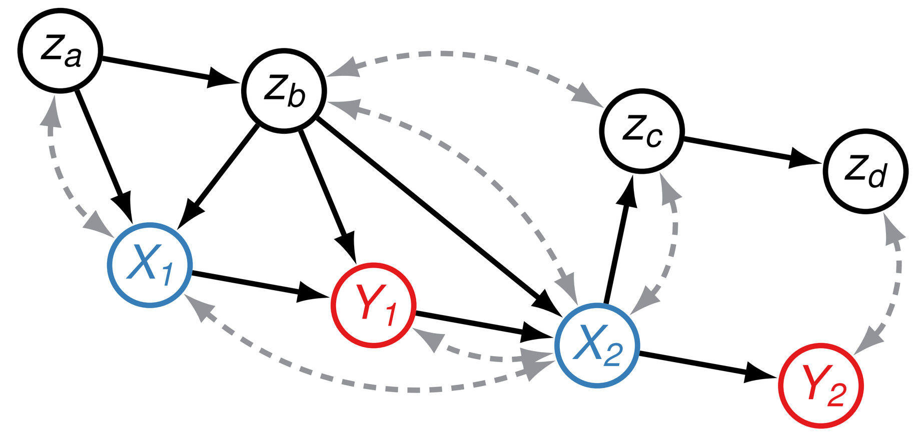


- There is a **complete** graphical criteria **Adjustment Criterion (AC)** for CA.

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

- **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** (i.e., $\exists$ graphs s.t. the causal effect is expressible through SCA while AC and mSBD are not satisfied).

Example: Not satisfied by existing criterion (mSBD),



but identified as SCA:

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

Complete Criterion for SCA

- **Definition** (Sequential Covariate Adjustment (SCA)). 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 $\mathbf{H}_i := \mathbf{X}^{(i)} \cup \mathbf{Y}^{(i)} \cup \mathbf{Z}^{(i)}$.

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

◎ Sequential Adjustment Criterion (SAC).

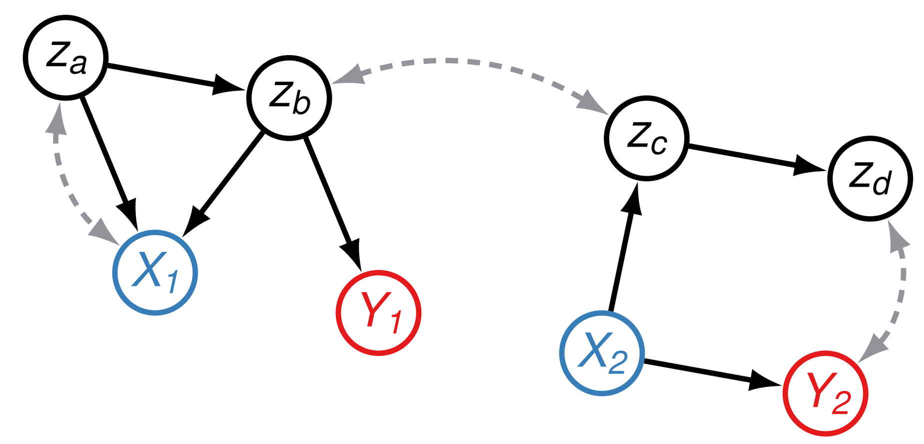
Given $\mathbf{Z} := (\mathbf{Z}_1, \cdots, \mathbf{Z}_m)$ where each $\mathbf{Z}_i$ is non-descendant of $\mathbf{X}^{\geq i+1}$, $\mathbf{Z}$ is said to satisfy *sequential adjustment criterion (SAC)* if the following conditions are satisfied for $i = 1, \cdots, m$:

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

Example (continued)

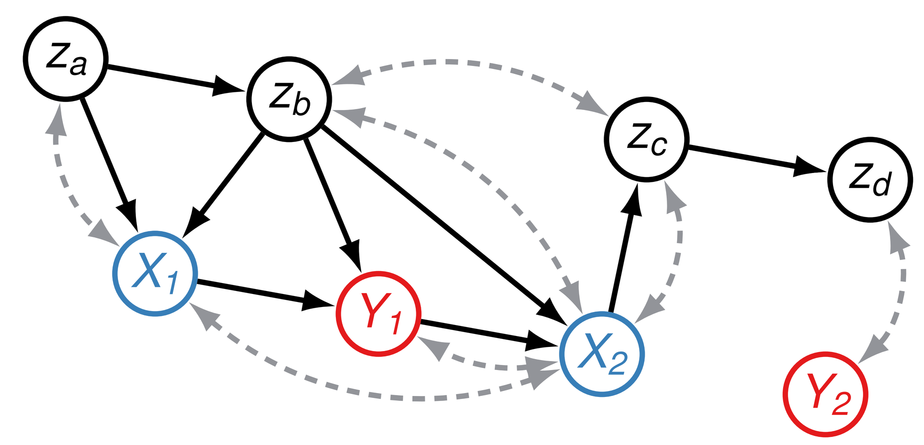
① Subgraph for X_1: $(\mathcal{G}_{\overline{X_2}})^{X_1, (Y_1, Y_2)}_{\mathrm{pbd}}$:

∗ Given $\mathbf{Z}_1 = \{Z_a, Z_b\}$, X_1 and $\{Y_1, Y_2\}$ are **d-separated** in the subgraph.
∗ $\mathbf{Z}_1$ is not a descendent of proper causal path set from X_1 to (Y_1, Y_2).



② Subgraph for X_2: $\mathcal{G}^{X_2, Y_2}_{\mathrm{pbd}}$

∗ 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-separated** in the subgraph.
∗ $\mathbf{Z}_2$ is not a descendent of proper causal path set from X_2 to Y_2.



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

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

◎ Soundness and Completeness of SAC

$\mathbf{Z}$ satisfies SAC $\Longleftrightarrow$ $P(\mathbf{y} \mid \mathrm{do}(\mathbf{x}))$ is expressible as SCA by adjusting over $\mathbf{Z}$

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

Constructive SAC (Algorithm)

Construction of Sequential Adjustment Set

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

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

Construction of Minimal Sequential Adjustment Set

`minSCA`$(\mathbf{X}, \mathbf{Y}, \mathcal{G})$ outputs $\mathbf{Z}^{\min}$, the smallest subset of $\mathbf{Z}^{\mathrm{an}}$ without sacrificing the validity of the adjustment.

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

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

Conclusion

1. **Sequential Adjustment Criterion (SAC)** is a sound and complete criterion for sequential covariate adjustment.
2. **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.
3. An algorithm **minSCA** for identifying a minimal sequential covariate adjustment set ensuring that no unnecessary vertices are included.