On Positivity Condition for Causal Inference

Inwoo Hwang* Yesong Choe* Yeahoon Kwon Sanghack Lee

Seoul National University
Graduate School of Data Science

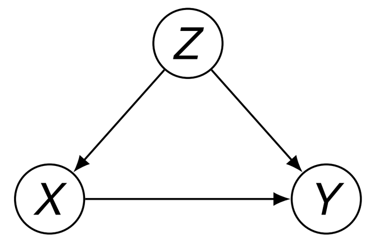
ICML
International Conference On Machine Learning

Introduction and Background

- Identifying and estimating a **causal effect** is a fundamental task when inferring a causal effect using observational study *without experiments*.
- **Strict positivity** ($P(\mathbf{V}) > 0$) of the given distribution is a long-standing critical assumption for causal inference, which is often unrealistic in many practical scenarios.
- We examine the graphical counterpart of the conventional positivity condition to license the use of identification formula **without strict positivity**.

Motivating Examples

Backdoor formula (existing results):

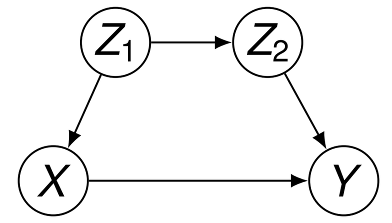


$$P_x(y) = \sum_{z} P(y \mid x, \mathbf{z})P(\mathbf{z})$$

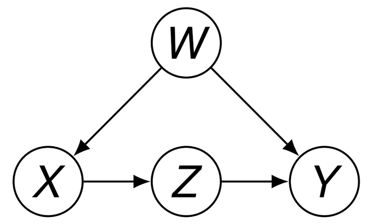
$$\Rightarrow \forall \mathbf{z}(P(\mathbf{z}) = 0 \vee P(x \mid \mathbf{z}) > 0) \equiv \text{adj}(\mathbf{x}; \mathbf{Z})$$

- To estimate the average treatment effect, there must exist some subjects that received the treatment for each value of the covariate in the population—i.e., $P(X \mid \mathbf{z}) > 0$ for all $\mathbf{z}$ with $P(\mathbf{z}) \neq 0$.
- Under the strict positivity, we can identify the causal effect—i.e., we can get the intervened distribution of y ($P_x(y)$) from the observed distribution $P(\mathbf{V})$.

Multiplicity of identification formulae and conditions:



$$\sum_{z_1} P(y \mid x, z_1)P(z_1) \quad \text{or} \quad \sum_{z_2} P(y \mid x, z_2)P(z_2)$$



$P_x(y) = \sum_w P(y \mid x, w)P(w)$	Backdoor
$P_x(y) = \sum_z P(z) \sum_{x'} P(y \mid x', z)P(x')$	Front-door
$P_x(y) = \sum_z P(z \mid x)P(y \mid z, w)$	IDENTIFY

- Without strict positivity, one may estimate the causal effect with a formula but ***not with the other***.

Causal Identification with Strict Positivity

- The causal effect $P_x(y)$ is **identifiable** if it can be uniquely computed from $P(\mathbf{V})$ in any causal model which induces $\mathcal{G}$.
- Two main tools for eliciting the identification formula, **do-calculus** and ***Q*-decomposition**, are established under the strict positivity assumption.

Do-Calculus

- The following transformation are valid for any positive do-distribution induced by a model:
  - Rule 1 (addition/deletion of observation): $P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{z}, \mathbf{w}) = P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X}}}}$
  - Rule 2 (exchange of action and observation): $P_{\mathbf{x},\mathbf{z}}(\mathbf{y} \mid \mathbf{w}) = P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{z}, \mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}}$
  - Rule 3 (addition/deletion of action): $P_{\mathbf{x},\mathbf{z}}(\mathbf{y} \mid \mathbf{w}) = P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X}},\overline{\mathbf{Z(W)}}}}$, where $\mathbf{Z}(\mathbf{W}) = \mathbf{Z} \setminus An(\mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X}}}}$.

Q-decomposition

- Given $\mathbf{H} \subseteq \mathbf{V}$, let $\mathbf{H}_1, \ldots, \mathbf{H}_k$ be the c-components of $\mathcal{G}[\mathbf{H}]$. Let $\prec$ be a topological order over $\mathcal{G}[\mathbf{H}]$. Let $\mathbf{H}^{\preceq i}$ be the variables in $\mathbf{H}$ that come before $V^{(i)}$ including $V^{(i)}$. Given $Q[\mathbf{H}] > 0$, where $Q[\mathbf{H}^{\preceq i}] = \sum_{\mathbf{h}^{\succ i}} Q[\mathbf{H}]$,

$$Q[\mathbf{H}_j] = \prod_{V^{(i)} \in \mathbf{H}_j} \frac{Q[\mathbf{H}^{\preceq i}]}{Q[\mathbf{H}^{\preceq i-1}]}.$$

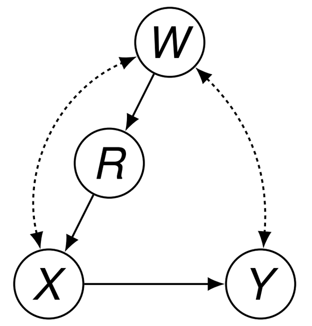
- (Napkin)

$$Q[W, X, Y] = \frac{Q[W, R, X, Y]}{Q[W, R, X]} \cdot \frac{Q[W, R, X]}{Q[W, R]} \cdot \frac{Q[W]}{Q[\emptyset]}$$

- These two well-known methods of identification heavily rely on $P(\mathbf{V}) > 0$.

Post-hoc Analysis

- **(Prop 7.1)** Post-hoc analysis yields a sufficient positivity condition for the identification formula derived through `Identify+`.



$$\exists r \frac{\sum_w P(y, x \mid r, w)P(w)}{\sum_w P(x \mid r, w)P(w)} \geq 0 \quad \Leftarrow \exists r(① \geq 0 \wedge ② > 0)$$

$① \geq 0 \Leftarrow \text{adj}(r; W)$

$② > 0 \Leftarrow \text{adj}(r; W) \wedge P(x, r) > 0$

$\therefore \exists r(\text{adj}(r; W) \wedge P(x, r) > 0)$

Causal Identification with Relaxed Positivity

✓ Do-Calculus

- **(Prop 4.2)** We develop a general and principled approach for deriving a sufficient positivity condition by examining the conditions for do-calculus.
  - Rule 1: $P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{z}, \mathbf{w}) = P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W})_{(\mathcal{G}\setminus\mathbf{X})}$ and $P_{\mathbf{x}}(\mathbf{z}, \mathbf{w}) > 0$
  - Rule 2: $P_{\mathbf{x},\mathbf{z}}(\mathbf{y} \mid \mathbf{w}) = P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{z}, \mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W})_{(\mathcal{G}\setminus\mathbf{X})_{\underline{\mathbf{Z}}}}$ and $P_{\mathbf{x}}(\mathbf{z}, \mathbf{w}) > 0$
  - Rule 3: $P_{\mathbf{x},\mathbf{z}}(\mathbf{y} \mid \mathbf{w}) = P_{\mathbf{x}}(\mathbf{y} \mid \mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{W})_{(\mathcal{G}\setminus\mathbf{X})_{\overline{\mathbf{Z(W)}}}}$ and $P_{\mathbf{x}}(\mathbf{w}) > 0$
- (Napkin)

$$\begin{aligned} P_x(y) &= P_{w,r}(y \mid x) && \text{if } P_{w,r}(x) > 0 \\ &= P_{w,r}(y, x)/P_{w,r}(x) && \text{if } P_{w,r}(x) > 0 \\ &= \frac{\sum_{w'} P(y, x \mid r, w')P(w')}{\sum_{w'} P(x \mid r, w')P(w')}. && \text{if adj}(r; W) \end{aligned}$$

$\therefore \exists r(\text{adj}(r; W) \wedge P(x, r) > 0)$

✓ *Q*-decomposition

- **(Thm 5.1)** We modify Q-decomposition so that it does not rely on the strict positivity.
- (Napkin)

$$Q[W, X, Y] = \frac{Q[W, R, X, Y]}{Q[W, R]} \cdot \frac{Q[W]}{Q[\emptyset]} \quad \text{if } Q[W, R] > 0$$

$$Q[W, X, Y] = 0 \quad \text{if } Q[W] = 0$$

- **(Thm 6.1)** We devise `Identify+` that returns a positivity and identification formula which is sound.

$P_x(y) =$

$$\frac{\sum_w Q[\{W, X, Y\}](x, y, w, r)}{\sum_{y', w} Q[\{W, X, Y\}](x, y', w, r)} \geq 0$$

$\therefore \exists r(\text{adj}(r; W) \wedge P(x, r) > 0)$

Conclusion

- We provide positivity conditions for **do-calculus** and **generalized Q-decomposition**, forming a basis for causal effect identification **without** $P(\mathbf{V}) > 0$.
- We devise `Identify+` algorithm, incorporating a relaxed version of **generalized Q-decomposition** into an existing identification method.
- We hope this research sparks further investigation into the development of an identification algorithm that adapts to the positivity.

References

- Pearl, J. Causal diagrams for empirical research. Biometrika, 82(4):669–688, 1995.
- Tian, J. and Pearl, J. On the identification of causal effects. Technical Report R-290-L, Department of Computer Science, University of California, Los Angeles, CA, 2003.