Causal Inference with Graphical Models

Jose M. Peña
STIMA, IDA, LiU

Lecture 2: Causal Effect Identification and *do*-Calculus

Contents

Literature

- Main source
  - Pearl, J. *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press, 2009. Chapter 3.
- Additional sources
  - Pearl, J. Causal Diagrams for Empirical Research. *Biometrika*, 82:669-710, 1995.
  - Pearl, J. *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press, 2009. Epilogue chapter.
  - Pearl, J. *Causality: Models, Reasoning, and Inference* (1st ed.). Cambridge University Press, 2000. Chapter 3.

Representations of Interventions

- Intervening on a variable $X_i \in V$ aims to modify the **natural** causal mechanism of X_i. For simplicity, we only consider interventions that set X_i to a fixed value x_i', and denote it as $do(x_i')$ or $\hat{x}_i'$.
- Assume that the causal model at hand consists of a DAG G over V and a set of structural equations $x_i = f_i(pa_i, u_i)$ for all $X_i \in V$ together with a set of distributions $p(u_i)$ or, alternatively, a set of conditional distributions $p(x_i|pa_i)$ for all $X_i \in V$.
- The result of an intervention can be represented by modifying the given causal model:
  - **Delete** the equation corresponding to X_i.
  - Replace x_i with x_i' in the remaining equations.
  - **Delete** from G the directed edges into X_i.
- Or, alternatively, by conditioning in an augmented causal model of the given one:
  - Augment G with the edge $F_i \rightarrow X_i$.
  - Let $Pa_i' = Pa \cup F_i$. Redefine

$$p(x_i|pa_i') = \begin{cases} p(x_i|pa_i) & \text{if } F_i = idle \\ 1 & \text{if } F_i = do(x_i') \text{ and } x_i = x_i' \\ 0 & \text{otherwise.} \end{cases}$$

  - Let p' be the distribution corresponding to the augmented G with an arbitrary prior on F_i. **Condition** on $F_i = do(x_i')$, i.e.

$$p(v|do(x_i')) = p'(v|F_i = do(x_i')).$$

- The above can be extended to multiple interventions.

Truncated Factorization

- Either representation of an intervention results in a truncated factorization

$$p(v|do(x_i')) = \begin{cases} \prod_{j\neq i} p(x_j|pa_j) & \text{if } x_i = x_i' \\ 0 & \text{otherwise.} \end{cases}$$

- Note that

$$\prod_{j\neq i} p(x_j|pa_j) = p(v)/p(x_i'|pa_i) = p(v)p(pa_i)/p(x_i', pa_i)$$
$$= p(v \setminus \{x_i'\} \setminus pa_i|x_i', pa_i)p(pa_i).$$

- **Adjustment for direct causes:** Let $X_i, Y \in V$ st $Y \notin Pa_i$. Then,

$$p(y|do(x_i')) = \sum_{pa_i} p(y|x_i', pa_i)p(pa_i).$$

- The goal of the above is to eliminate spurious (i.e., non-causal) correlations between cause and effect.
- Note that if Y is not a descendant of X_i, then $Y \perp_G X_i|Pa_i$ and thus, as expected,

$$p(y|do(x_i')) = \sum_{pa_i} p(y|x_i', pa_i)p(pa_i) = \sum_{pa_i} p(y|pa_i)p(pa_i) = p(y).$$

- Things get more complicated when some variables in Pa_i are unobserved, since it prevents estimation of $p(y|x_i', pa_i)$ and $p(pa_i)$. This requires dropping the assumption that the error terms U_i are independent.

Truncated Factorization

$$\begin{tikzcd} & Gender \arrow[dl] \arrow[dr] & \\ Drug \arrow[rr] & & Recovery \end{tikzcd}$$

	Drug=1	Drug=0
Gender=1	81/87 recovered	234/270 recovered
Gender=0	192/263 recovered	55/80 recovered

- Average causal effect:

  $$E[R|do(D=1)] - E[R|do(D=0)] = p(R=1|do(D=1)) - p(R=1|do(D=0))$$

  which can also be interpreted as the fraction of the population that recovers if everyone takes the drug compared to when no one takes the drug. Moreover, adjusting for the direct causes gives

$$\begin{aligned} p(R=1|do(D=1)) &= p(R=1|D=1,G=1)p(G=1) + p(R=1|D=1,G=0)p(G=0) \\ &= (81/87)(87+270)/700 + (192/263)(263+80)/700 = 0.832 \end{aligned}$$

$$\begin{aligned} p(R=1|do(D=0)) &= p(R=1|D=0,G=1)p(G=1) + p(R=1|D=0,G=0)p(G=0) \\ &= (234/270)(87+270)/700 + (55/80)(263+80)/700 = 0.7818 \end{aligned}$$

Truncated Factorization

- Plan evaluation is relevant in fields such as health management, economic policy making or robot motion planning. Say Z_k represents the process state at time t_k (e.g., temperature), X_k stands for some control variables (e.g., chemicals), and Y is the process outcome (e.g., product quality).

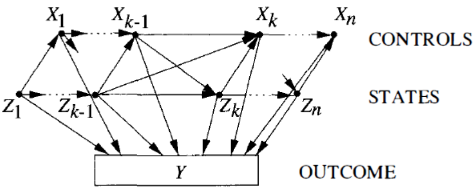


Figure 3.3 Dynamic causal diagram illustrating typical dependencies among the control variables $X_1, \ldots, X_n$, the state variables $Z_1, \ldots, Z_n$, and the outcome variable Y of a sequential process.

- Factorization:

$$
\begin{aligned}
& p(y, z_1, \ldots, z_n, x_1, \ldots, x_n) \\
= {} & p(y|z_1, \ldots, z_n, x_1, \ldots, x_n) \prod_k p(x_k|x_{k-1}, z_k, z_{k-1}) \prod_k p(z_k|z_{k-1}, x_{k-1}).
\end{aligned}
$$

- Plan evaluation, i.e. the plan is a set of actions $do(x'_k)$:

$$
\begin{aligned}
p^*(y) &= p(y|do(x'_1), \ldots, do(x'_n)) \\
&= \sum_{z_1, \ldots, z_n} p(y|z_1, \ldots, z_n, x'_1, \ldots, x'_n) \prod_k p(z_k|z_{k-1}, x'_{k-1}).
\end{aligned}
$$

Causal Effect Identifiability

- Given a causal structure which may include unobserved variables, the causal effect $p(y|do(x_i'))$ is **identifiable** if it can be computed uniquely from any positive probability distribution over the observed variables.
- Positivity ensures that the effect is well defined.
- Therefore, $p(y|do(x_i'))$ is identifiable if Y, X_i, and Pa_i are observed, i.e. measured. The effect is computed by adjusting for the parents.
- $p(y|do(x_i'))$ is not identifiable in the **bow graph**:

$$\begin{tikzcd} X \arrow[r] \arrow[r, bend left, leftrightarrow] & Y \end{tikzcd} \quad \equiv \quad \begin{tikzcd} & U \arrow[dl] \arrow[dr] & \\ X \arrow[rr] & & Y \end{tikzcd}$$

- Proof: We construct two causal models M_1 and M_2 st $p_1(x, y) = p_2(x, y)$ but $p_1(y|do(x')) \neq p_2(y|do(x'))$. Specifically, let X, Y and U be binary, and take

M_1	M_2
$u = Uniform(0, 1)$	$u = Uniform(0, 1)$
$x = u$	$x = u$
$y = XOR(x, u)$	$y = 0.$

Back-Door Criterion

- A set of variables Z satisfies the **back-door criterion** wrt an ordered pair of sets of variables (X, Y) in a causal structure G which may include unobserved variables if
  - Z contains no descendants of X, and
  - Z blocks every path between X and Y that contains an arrow into X.

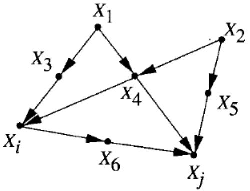


Figure 3.4 A diagram representing the back-door criterion; adjusting for variables $\{X_3, X_4\}$ (or $\{X_4, X_5\}$) yields a consistent estimate of $P(x_j \mid \hat{x}_i)$.

- If Z satisfies the back-door criterion wrt (X, Y), then

$$p(y|do(x)) = \sum_z p(y|x,z)p(z).$$

- The role of Z is to block only the paths entering X through the back-door. Several such sets Z may exist but some may be preferred, e.g. due to their size. Note that $Z = Pa_X$ always satisfies the criterion (what we called adjustment for the direct causes) but it may include latent variables.

Back-Door Criterion

- Proof: Augment G with the edges $F_i \to X_i$ for all $X_i \in X$. Call the result G'. Note that $Y \perp_{G'} F_X | X \cup Z$ and, thus, the intervention $F_X = do(x)$ cannot be distinguished from the observation $X = x$, i.e.

$$p(y|z, x, F_X = do(x)) = p(y|z, x, F_X = idle) = p(y|z, x).$$

Moreover, note that $Z \perp_{G'} F_X | \varnothing$ since Z are non-descendants of F_X.

Therefore,

$$\begin{aligned} p(y|do(x)) = p'(y|F_x) = \sum_z p'(y|z, F_x)p'(z|F_x) &= \sum_z p'(y|z, x, F_x)p'(z|F_x) \\ &= \sum_z p'(y|z, x)p'(z) \end{aligned}$$

where adding x in the third equality is licensed by the fact that F_x ($\equiv F_X = do(x)$) implies $X = x$.

Front-Door Criterion

- A set of variables Z satisfies the **front-door criterion** wrt an ordered pair of sets of variables (X, Y) in a causal structure G which may include unobserved variables if
  - Z blocks all the directed paths from X to Y,
  - there is no unblocked back-door path from X to Z, and
  - all the back-door paths from Z to Y are blocked by X.

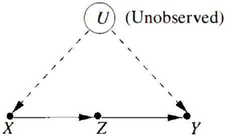


Figure 3.5 A diagram representing the front-door criterion. A two-step adjustment for Z yields a consistent estimate of $P(y \mid \hat{x})$.

- Note that Figure 3.5 is Figure 3.4 with $U = \{X_1, \ldots, X_5\}$. Note that the back-door criterion does not help here.
- If Z satisfies the front-door criterion wrt (X, Y), then

$$p(y|do(x)) = \sum_z p(z|x) \sum_{x'} p(y|x', z)p(x').$$

Front-Door Criterion

- Proof: Since $\varnothing$ satisfies the back-door criterion wrt (X, Z), then $p(z|do(x)) = p(z|x)$. Since X satisfies the back-door criterion wrt (Z, Y), then

  $$p(y|do(z)) = \sum_{x'} p(y|x', z)p(x')$$

  and $Y \perp_{G'} F_Z | X \cup Z$.

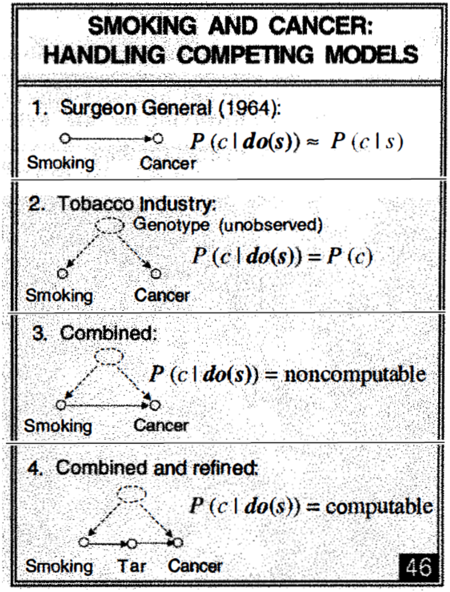
  Finally, combine the previous results via

  $$\begin{aligned} p(y|do(x)) = \sum_z p(y|do(x), z)p(z|do(x)) &= \sum_z p(y|do(x), do(z))p(z|do(x)) \\ &= \sum_z p(y|do(z))p(z|do(x)) \end{aligned}$$

  where the last equality follows from the fact that Z blocks all the directed paths from X to Y.

Front-Door Criterion

- The effect of smoking on lung cancer: Non-identifiable vs identifiable.



do-Calculus

- Three rules whose repeated application together with standard probability manipulations, aims to transform a causal effect into an expression that only involves observational quantities:
  - Rule 1 (**insertion/deletion of observations**)

    $$p(y|do(x), \boldsymbol{z}, w) = p(y|do(x), w) \text{ if } Y \perp_{G'} Z|X \cup W||X$$

    where the antecedent is satisfied if $Y \perp Z|X \cup W$ holds in the graph resulting from intervening on X in G', i.e. delete all the (bi)directed edges into X from the original causal structure augmented with the edges $F_V \to V$.
  - Rule 2 (**intervention/observation exchange**)

    $$p(y|do(x), \boldsymbol{do(z)}, w) = p(y|do(x), \boldsymbol{z}, w) \text{ if } Y \perp_{G'} F_Z|X \cup W \cup Z||X.$$

  - Rule 3 (**insertion/deletion of interventions**)

    $$p(y|do(x), \boldsymbol{do(z)}, w) = p(y|do(x), w) \text{ if } Y \perp_{G'} F_Z|X \cup W||X.$$

- The rules are sound and **complete**.
- There is a sound and complete algorithm to apply the rules.

do-Calculus

- Proof:
  - In rule 2, simply note that $Y \perp_{G'} F_Z | X \cup W \cup Z || X$ implies that the intervention $F_Z = do(z)$ cannot be distinguished from the observation $Z = z$.

    Alternatively, note that no $(X \cup W)$-open path from Y to F_Z can reach Z through one of its parents. So, the unblocked paths from F_Z to Y leave Z through its children. For these paths, intervening or observing is the same.
  - In rule 3, simply note that $Y \perp_{G'} F_Z | X \cup W || X$ implies that the intervention $F_Z = do(z)$ is irrelevant for the causal effect at hand.

    Alternatively, note that no $(X \cup W)$-open path from Z to Y can leave Z through one its children. For the rest of unblocked paths, intervening on Z is irrelevant.
  - Rule 1 follows from rules 2 and 3.

do-Calculus

- Alternative formulation:
  - Rule 1 (**insertion/deletion of observations**)

    $$p(y|do(x), \mathbf{z}, w) = p(y|do(x), w) \text{ if } Y \perp_{G_{\overline{X}}} Z | X \cup W$$

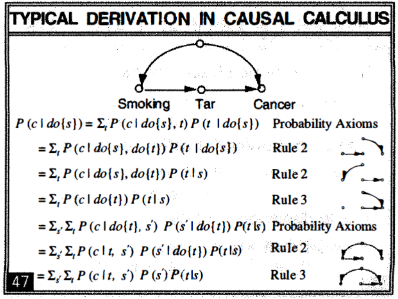
    $G_{\overline{X}}$ is G after deleting all the (bi)directed edges into X, i.e. simulate $do(x)$.
  - Rule 2 (**intervention/observation exchange**)

    $$p(y|do(x), \mathbf{do(z)}, w) = p(y|do(x), \mathbf{z}, w) \text{ if } Y \perp_{G_{\overline{X}\underline{Z}}} Z | X \cup W$$

    $G_{\overline{X}\underline{Z}}$ is G after deleting all the (bi)directed edges into X and all the directed edges out of Z.
  - Rule 3 (**insertion/deletion of interventions**)

    $$p(y|do(x), \mathbf{do(z)}, w) = p(y|do(x), w) \text{ if } Y \perp_{G_{\overline{X}\,\overline{Z(W)}}} Z | X \cup W$$

    where $Z(W)$ are the nodes in Z that are not ancestors of W in $G_{\overline{X}}$.



Causal Effect Identifiability

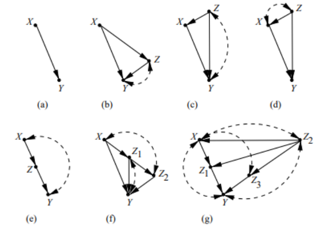


Figure 3.8: Typical models in which the effect of X on Y is identifiable. Dashed arcs represent confounding paths, and Z represents observed covariates.

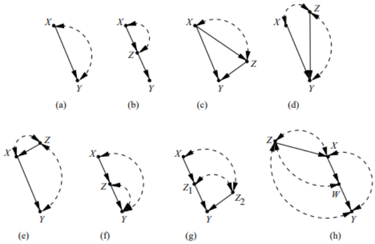


Figure 3.9: Typical models in which $P(y|\hat{x})$ is not identifiable.

- Given $X \in V$ and $Y \subseteq V$, $p(y|do(x))$ is identifiable if there is no bidirected path between X and any of its children on the directed paths from X to Y.

Surrogate Experiments

- Assume that we cannot identify $p(y|do(x))$ by the previous methods or by randomizing X.
- Assume that we can **randomize** a surrogate variable Z, e.g. Y = heart condition, X = cholesterol levels, and Z = diet.
- Then, $p(y|do(x))$ is identifiable whenever
  - X blocks all the directed paths from Z to Y, and
  - $p(y|do(x))$ is identifiable in $G_{\overline{Z}}$.
- In the cholesterol example, the conditions above require that diet (Z) has no direct effect on heart condition (Y), and cholesterol level (X) and heart condition are not confounded unless we can neutralize the confounding.
- See Figures 3.9 (e) and (h) for some examples. In Figure 3.9 (e), for instance, we have that

  $$p(y|do(x)) = p(y|x, do(z)).$$

  So if Figure 3.9 (e) is the causal graph in our cholesterol example, then we should hold the diet constant for the individuals in the population $(do(z))$ and, afterwards, estimate the distribution of heart disease (Y) given cholesterol level (X).
- Note that there is no need to randomize Z, just set it to an arbitrary constant.
- Proof: $p(y|do(x)) = p(y|do(x), do(z))$ by rule 3, because $Y \perp_{G_{\overline{X}\,\overline{Z}}} Z|X$. Moreover, $p(y|do(x), do(z))$ is the causal effect of X on Y in a model governed by $G_{\overline{Z}}$, which is identifiable.

Summary

- Representations of Interventions
- Truncated Factorization
- Causal Effect Identifiability
- Back-Door Criterion
- Front-Door Criterion
- *do*-Calculus
- Surrogate Experiments

Thank you