A short note on the proximal causal inference

Yen-Chi Chen
University of Washington
February 17, 2022

References:

[TYCSM2020 ] Tchetgen, E. J. T., Ying, A., Cui, Y., Shi, X., & Miao, W. (2020). An introduction to proximal causal learning. arXiv preprint arXiv:2009.10982.

[SWT2021 ] Shpitser, I., Wood-Doughty, Z., & Tchetgen, E. J. T. (2021). The proximal id algorithm. arXiv preprint arXiv:2108.06818.

# 1 Problem setup

In this note, I will briefly review the idea of proximal causal inference approach. This will be a short summary/introduction to the above two recent papers.

We consider a standard causal inference problem under observational data that $A \in \{0,1\}$ is the binary treatment variable and $Y \in \mathbb{R}$ is the outcome/response of interest. Let $X$ be observed covariates that may include part of the confounders. Under the potential outcome model, the outcome $Y$ admits two potential outcomes $Y(0), Y(1)$. The parameter of interest is the average treatment effect (ATE) $\tau = \mathbb{E}(Y(1) - Y(0))$.

Suppose $X$ contains all confounders, we have

$$(Y(0), Y(1)) \perp A | X$$

and it is well-known that the ATE is identifiable.

However, if $X$ does not contain all confounders, i.e., there is another unobserved random vector $U$, the ATE is in general non-identifiable. Note that in this case, we still have

$$(\mathbf{U}) \qquad (Y(0), Y(1)) \perp A | X, U. \tag{1}$$

To identify the ATE, we need additional information. The idea of proximal causal inference is based on the introduction of two proxy variables $(Z, W)$:

- *The treatment proxy* $Z$. $Z$ is a variable that is caused by the treatment $A$ so it admits two potential outcomes $Z(0), Z(1)$ and it may be dependent with the confounders $(X, U)$. However, the proxy control $Z$ is not associated with $Y$ or $W$ conditioned on the confounders.
- *The outcome proxy (negative control outcome)* $W$. $W$ is a variable that is known to be unaffected by the treatment $A$ but it may be dependent with confounders $(X, U)$ and $Y$.

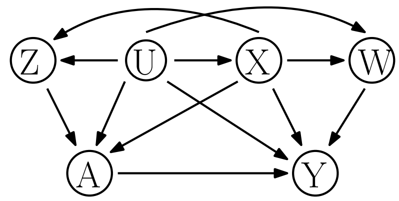


Figure 1: A causal DAG for the proximal inference.

Figure 1 provides a graphical representation under the proximal inference. Formally, the two proxies $(Z,W)$ has to satisfy the following proximal assumptions:

$$(\mathbf{P1}) \qquad Z(a) \perp Y(a)|X,U \tag{2}$$

$$(\mathbf{P2}) \qquad W \perp (A,Z(a))|X,U. \tag{3}$$

Note that the above assumptions imply that

$$Y \perp Z|U,X,A, \quad W \perp Z|U,X,A. \tag{4}$$

## 2 Bridge function

Based on the above notations, our observed data consists of variables $(Y,A,X,Z,W)$. And there is a bridge function $b^*(y,w,a,x,z)$ such that

$$p(y|a,x,z) = \int b^*(y,w,a,x,z)p(w|a,x,z)dw.$$

Namely,

$$p(y|a,x,z) = \mathbb{E}[b^*(y,W,A,X,Z)|A=a,X=x,Z=z].$$

To identify the ATE, we need two additional assumptions related to the bridge function

$$(\mathbf{A1}) \qquad b^*(y,w,a,x,z) = b^*(y,w,a,x) \tag{5}$$

$$(\mathbf{A2}) \qquad \mathbb{E}(f(U)|A,X,Z) = 0 \quad \forall A,X,Z \Leftrightarrow f(u) = 0. \tag{6}$$

Assumption (A1) is the main assumption on the bridge function. Under assumption (A1), [TYCSM2020] has provided an approach to construct $b^*$ from the data. So in what follows, we assume that the bridge function $b^*(y,w,a,x)$ is known.

Assumption (A2) is called the *completeness assumption.* It requires that most information/randomness of unmeasured confounders $U$ is contained in the variables $(A,X,Z)$.

## 3 Identification of potential outcome

Now we will show that under (U), (P1-2), and (A1-2), the distribution of potential outcome $Y(a)$ is identifiable.

**Theorem 1** *Assume (U), (P1-2), and (A1-2). Then the distribution of potential outcome $Y(a)$, $p(y(a))$, is identified from the following formula:*

$$p(y(a)) = \int b^*(y,w,a,x)p(w,x)dwdx.$$

The above formula is also known as *proximal g-formula*.

**Proof.**

First, note that

$$\begin{aligned}
p(y|a,x,z) &= \int p(y,u|a,x,z) \\
&= \int p(y|u,a,x,z)p(u|a,x,z)du \\
&\overset{(4)}{=} \int p(y|a,u,x)p(u|a,x,z)du.
\end{aligned}$$

Similarly,

$$\begin{aligned}
p(w|a,x,z) &= \int p(w,u|a,x,z) \\
&= \int p(w|u,a,x,z)p(u|a,x,z)du \\
&\overset{(4)}{=} \int p(w|u,a,x)p(u|a,x,z).
\end{aligned}$$

Recall from assumption (A1), we have

$$p(y|a,x,z) = \int b^*(y,w,a,x)p(w|a,x,z)dw.$$

So the above two new representations of $p(y|a,x,z)$ and $p(w|a,x,z)$ will imply:

$$\int p(y|a,u,x)p(u|a,x,z)du = \int\int b^*(y,w,a,x)p(w|u,a,x)p(u|a,x,z)dwdu,$$

which is equivalent to

$$\begin{aligned}
0 &= \int \left[p(y|a,u,x) - \int b^*(y,w,a,x)p(w|u,a,x)dw\right] p(u|a,x,z)du \\
&= \mathbb{E}[p(y|A,U,X) - \int b^*(y,w,A,X)p(w|U,A,X)dw|A=a,X=x,Z=z]
\end{aligned}$$

for all $a,x,z$. By assumption (A2), this implies that

$$p(y|a,u,x) = \int b^*(y,w,a,x)p(w|a,u,x)dw. \tag{7}$$

By assumption (U),

$$p(y|a,u,x) = p(y(a)|a,u,x) \overset{(U)}{=} p(y(a)|u,x).$$

Thus,

$$\begin{aligned}
p(y(a)) &= \int p(y(a)|u,x)p(u,x)dudx \\
&= \int p(y|a,u,x)p(u,x)dudx \\
&\overset{(7)}{=} \int\int b^*(y,w,a,x) \underbrace{p(w|a,u,x)}_{=p(w|u,x) \text{ by } (3)} p(u,x)dwdudx \\
&= \int b^*(y,w,a,x) \int p(w,u,x)dudwdx \\
&= \int b^*(y,w,a,x)dwdx,
\end{aligned}$$

which completes the proof.

□