Causal Directed Acyclic Graphs

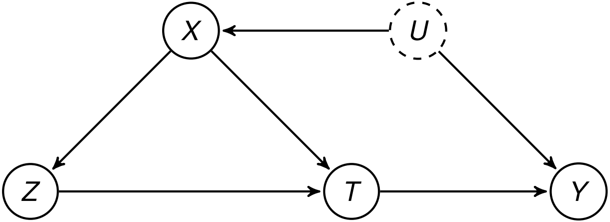
Kosuke Imai

Harvard University

Spring 2021

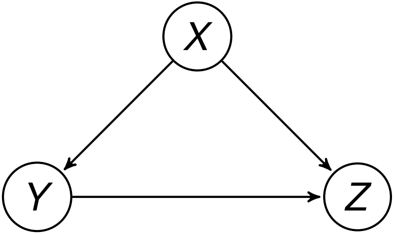
Elements of DAGs (Pearl. 2000. *Causality*. Cambridge UP)

- $\mathcal{G} = (E, V)$
  1. V: nodes or vertices $\rightsquigarrow$ variables (observed and onobserved)
  2. E: directed arrows $\rightsquigarrow$ *possibly non-zero* direct causal effects



- Acyclic: no cycle, no simultaneity
- Encoded assumptions
  - Absence of variables: no other common causes (observed and unobserved) of any pair of variables
  - Absence of arrows: zero causal effect

DAG Terminology

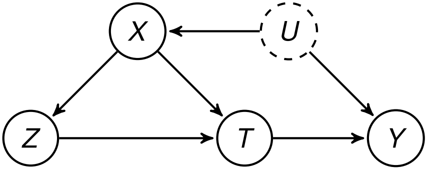


- chain: $X \rightarrow Y \rightarrow Z$
- fork: $Y \leftarrow X \rightarrow Z$
- inverted fork: $X \rightarrow Z \leftarrow Y$

- Parents (Children): directly causing (caused by) a vertex $i \rightarrow j$
- Ancestors (Descendents): directly or indirectly causing (caused by) a vertex $i \rightarrow \cdots \rightarrow j$
- Path: an acyclic sequence of adjacent nodes
  - causal path: all arrows pointing out of i and into j
  - non-causal path: some arrows going against causal order
- Collider: a vertex on a path with two incoming arrows

Nonparametric Structural Equation Models (NPSEM)

- Equivalence to the nonparametric structural equation models:



$$
\begin{aligned}
Y &= f_1(T, U, \epsilon_1) \\
T &= f_2(X, Z, \epsilon_2) \\
Z &= f_3(X, \epsilon_3) \\
X &= f_4(U, \epsilon_4)
\end{aligned}
$$

- NPSEM allows for
  1. causal interpretation
  2. any functional form
  3. any form of heterogenous and interaction effects
  4. LSEM as a special case

- Factorized likelihood function:

$$P(X_1, X_2, \ldots, X_J) = \prod_{j=1}^{J} P(X_j \mid \text{pa}(X_j))$$

$$P(T, U, X, Y, Z) \; = \; P(Y \mid T, U) P(T \mid X, Z) P(Z \mid X) P(X \mid U) P(U)$$

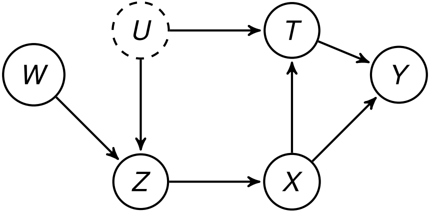
D-separation

- Does the conditional independence relation, $A \perp\!\!\!\perp B \mid C$, hold where A, B, C are sets of vertices?
  1. Identify all paths from any vertex in A to any vertex in B
  2. Check if each path is blocked
  3. If all paths are blocked, then A is *d*-separated from B by C

- Path is blocked if at least one of the following conditions holds
  1. it includes a noncollider vertex that is in C
  2. it includes a collider that is not in C and no descendant of any collider is in C

- If A and B are *d*-separated, $A \perp\!\!\!\perp B \mid C$ holds
- If A and B are *d*-connected (i.e., not *d*-separated), $A \not\!\perp\!\!\!\perp B \mid C$ in at least one distribution compatible with the DAG

D-separation Example



1. Are W and Y marginally independent of each other?
   - blocked: $W \rightarrow Z \leftarrow U \rightarrow T \rightarrow Y$ $W \rightarrow Z \leftarrow U \rightarrow T \leftarrow X \rightarrow Y$
   - unblocked: $W \rightarrow Z \rightarrow X \rightarrow Y$ $W \rightarrow Z \rightarrow X \rightarrow T \rightarrow Y$
2. Can we make W and Y independent by conditioning?
   - block the unblocked paths without unblocking the blocked paths
   - conditioning on X would block the first path because X is a descendant of collider Z
   - conditioning on T and/or Z would unblock some of the blocked paths because they are colliders

Backdoor Criterion

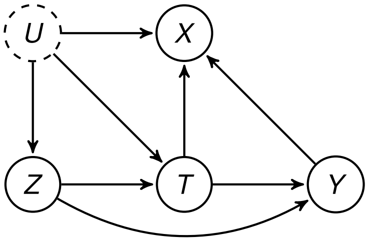
- Can we nonparametrically identify the causal effect of T on Y given a set of variables X?

- Backdoor criterion for X:
  1. No vertex in X is a decendent of T (no post-treatment bias), and
  2. X blocks all paths between T and Y with an incoming arrow into T (backdoor paths)
- Idea: block all non-causal paths
- Estimation:

$$P(Y(t)) = \sum_{x} P(Y \mid T = t, X = x) P(X = x)$$

- Confounder selection criterion (VanderWeele and Shpitser. 2011. *Biometrics*)
  If there exist a set of observed covariates that meet the backdoor criterion, it is sufficient to condition on all observed pretreatment covariates that either cause treatment, outcome, or both.

Example of Backdoor Criterion



- Can we identify the causal effect of T on Y?
- Backdoor paths:
  1. $T \leftarrow Z \rightarrow Y$ and $T \leftarrow U \rightarrow Z \rightarrow Y$
  2. $T \leftarrow Z \leftarrow U \rightarrow X \leftarrow Y$
  3. $T \leftarrow U \rightarrow X \leftarrow Y$
- Block (1) by conditioning on Z, which will not unblock (2) and (3)

$$P(Y(t)) = \sum_z P(Y \mid T = t, Z = z)P(Z = z)$$

DAGitty (`http://dagitty.net/`)