

Causal Graphs

– Handout with definitions–

Basic definitions

Node represents a random variable labeled by letter. Observed random variables are marked by solid circle \bullet and unobserved - by hollow circle \circ .

Edge shows dependence between joining variables.

Adjacent variables are connected by an edge.

Adjacent edges meet at a variable.

Directed edges represent the causes by single-headed arrows.

Parent/child is the starting(tail)/ending(head) variable. Therefore, a directed edge represents a direct effect of a parent on a child.

Root is a variable that has no parent, or, in other words, exogenous variable determined only by forces outside of the graph.

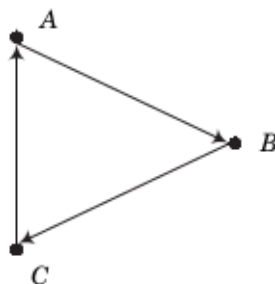
Sink is a variable with no children.

Path is a sequence of adjacent edges.

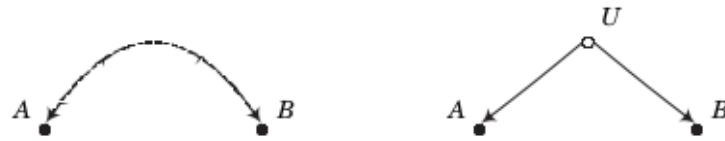
Directed path is a path traced out entirely along arrows tail-to-head.

If there is a directed path from A to B, A is an **ancestor** of B; B is a **descendant** of A.

Directed acyclic graph (DAG) is a graph with only arrows for edges and no feedback loops (i.e. no variable is its own ancestor or its own descendant):



Mutual dependence of two variables on one or more common causes is shown by a curved and dashed bidirected edge:

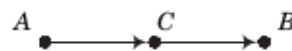


Basic patterns and strategies

Chain of mediation is a relationship when A affects B through A's causal effect on C and C's causal effect on B.

Mutual dependence is a relationship when A and B are both caused by (graph b).

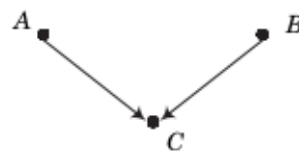
Mutual causation is a relationship when A and B both causes of C.



(a) Mediation



(b) Mutual dependence



(c) Mutual causation

Confounding variable is a variable that affects both the dependent and independent variable (in our example, variable C on the graph b).

Collider is a variable that has two arrows running into it (in our example, variable C on

the graph c).

Conditioning as a modeling strategy means transforming one graph into a simpler set of component graphs where fewer causes are represented.

Back-door path is a path between any causally ordered sequence of two variables that include a directed edge that points to the first variable.

Back-door criterion is a set of conditions used to determine whether or not conditioning on a given set of observed variable will identify the causal effect. The causal effect is identified by conditioning on a set of variables Z if and only if all back-door paths between the causal variable and the outcome variable are blocked after conditioning on Z . All back-door paths are blocked by Z if and only if each back-door path:

1. contains a chain of mediation $A \rightarrow C \rightarrow B$ where the middle variable C is in Z , or
2. contains a fork of mutual dependence $A \leftarrow C \rightarrow B$, where the middle variable C is in Z , or
3. contains an inverted fork of mutual causation $A \rightarrow C \leftarrow B$, where the middle variable C and all of C 's decedents are not in Z .

References

- Morgan, S., & Winship, C. (2014). *Counterfactuals and causal inference*. West Nyack: Cambridge University Press.
- Pearl, J. (2009). *Causality*. Cambridge University Press.