Topics in Causal Inference

STAT41530

Jingshu Wang

Lecture 3

Topic: causal directed acyclic graph (DAG)

- Markov factorization
- Connection with the potential outcome framework
- D-separation and conditional independence

Directed Acyclic Graphs (DAGs)

- DAG: a directed graph with no cycles

$$G = (V, E)$$

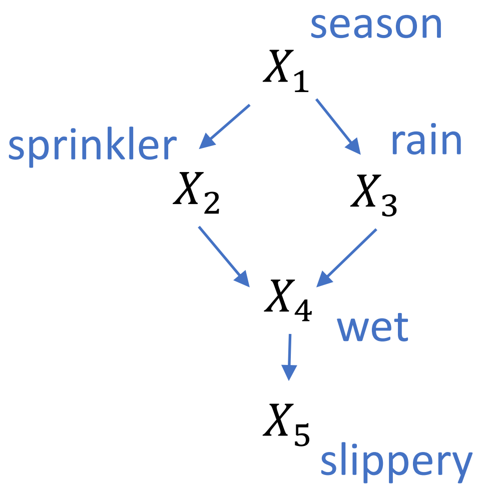
nodes

Directed edges

$$A \rightarrow Y$$

parent

descendant



- Bayesian network: nodes are random variables

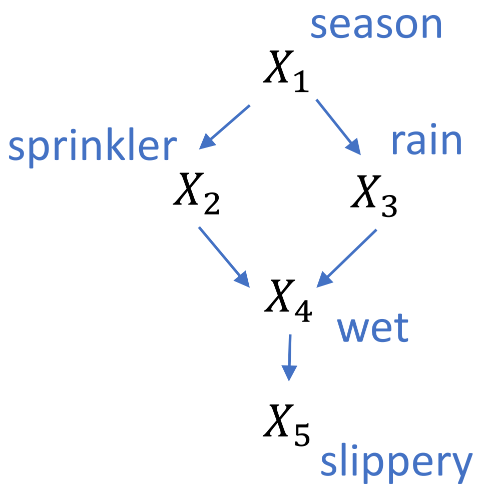
$$\mathbb{P}[V] = \mathbb{P}[X_1, X_2, \cdots, X_J] = \prod_{j=1}^{J} \mathbb{P}[X_j \mid X_1, \cdots, X_{j-1}]$$

Markov Factorization

- Markov assumption:

$$\mathbb{P}[V] = \prod_{j=1}^{J} \mathbb{P}[X_j \mid PA_j]$$

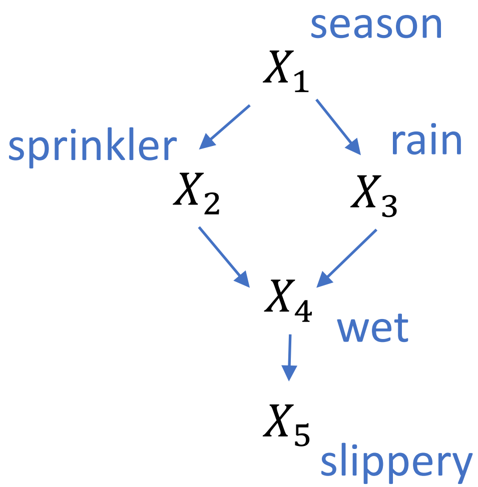
PA_j: parents of X_j (nodes that have direct arrows to X_j)



- In the example, the Markov assumption assume that

$$\mathbb{P}[X_1, X_2, X_3, X_4, X_5] = \mathbb{P}[X_1]\,\mathbb{P}[X_2 \mid X_1]\,\mathbb{P}[X_3 \mid X_1]\,\mathbb{P}[X_4 \mid X_2, X_3]\,\mathbb{P}[X_5 \mid X_4]$$

Causal DAGs (Pearl 1995, Biometrika)



A DAG is a causal DAG if

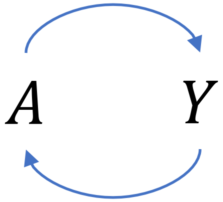
- X_j not a parent of X_m $\Leftrightarrow$ absense of a direct causal effect
- The DAG satisfies Markov assumption

$$\mathbb{P}[V] = \prod_{j=1}^{J} \mathbb{P}[X_j \mid PA_j]$$

Conditional on its direct causes, any variable on the DAG is independent of any other variable for which it is not a cause

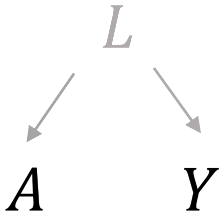
What are assumed?

- No directed circle:



A can not cause itself, $P(A,Y) \neq P(A|Y)P(Y|A)$

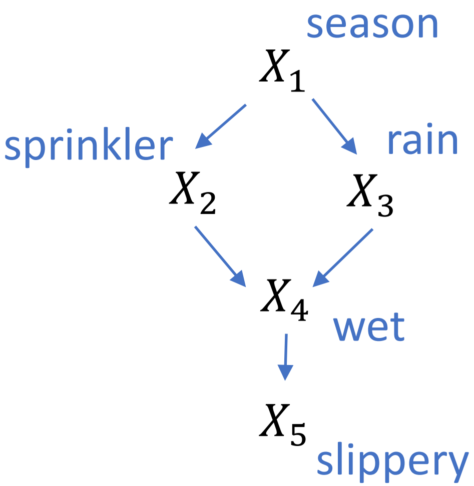
- The common causes of any pair of variables in the graph must be in the graph



$$P(A,Y) \neq P(A)P(Y)$$

- SUTVA: consistency and no interference
- Intervention can be done on any node that has an arrow out

Connection with the potential outcome framework

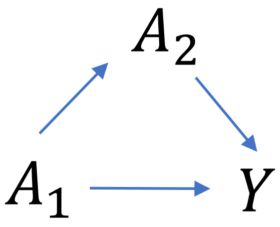


For each X_j, assume existence of random error $E_{X_j} \perp PA_j$ and deterministic unknow function f_j for each node j

$$\boxed{X_j(pa_j) = f_j(pa_j, E_{X_j})}$$

Potential outcomes for joint intervention on parents PA_j

Non-parametric structure equation



Potential outcomes:

$$A_2(a_1) = f_2(a_1, E_{A_2})$$

$$Y(a_1, a_2) = f(a_1, a_2, E_Y)$$

$$Y(a_1) = f(a_1, A_2(a_1), E_Y)$$

Observed outcome:

$$Y = f(A_1, A_2, E_Y)$$

Structural equations and the Markov assumption

Structural equations $X_j\left(pa_j\right) = f_j(pa_j, E_{X_j})$

+ $\left\{E_{X_j}, j = 1, 2, \cdots, J\right\}$ mutually independent

$$\Longrightarrow$$

$$P(V) = \prod_{j=1}^{J} P\left[X_j \mid PA_j\right]$$

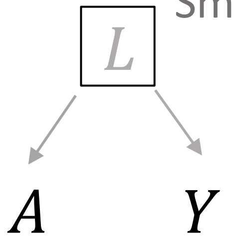
Proof idea:

1) (topological ordering) There exists an order so that for each j, $PA_j \subseteq \{X_1, X_2, \cdots, X_{j-1}\}$ (proof by induction on J)
2) As $\left\{E_{X_j}, j = 1, 2, \cdots, J\right\}$ are mutually independent, $E_{X_j} \perp S$ if set S do not contain X_j's descendent
3) $X_j = f_j\left(PA_j, E_{X_j}\right) \perp \{X_1, \cdots, X_{j-1}\} / PA_j \mid PA_j$

$$\mathbb{P}[V] = \mathbb{P}[X_1, X_2, \cdots, X_J] = \prod_{j=1}^{J} \mathbb{P}[X_j \mid X_1, \cdots, X_{j-1}] = \prod_{j=1}^{J} \mathbb{P}[X_j \mid PA_j]$$

Causal DAG and marginal / conditional independence

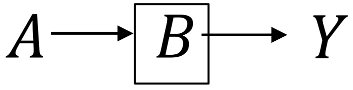
A few examples:


L Smoker/non-smoker

A Carrying a lighter

Y Lung cancer

$A \not\perp Y \quad A \perp Y \mid L$


Platelet aggregation

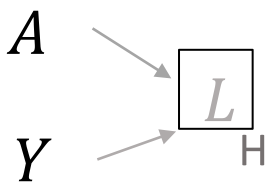
$A \rightarrow \boxed{B} \rightarrow Y$

aspirin

Heart disease

$A \not\perp Y \quad A \perp Y \mid B$

B is a mediator


Genotype

A

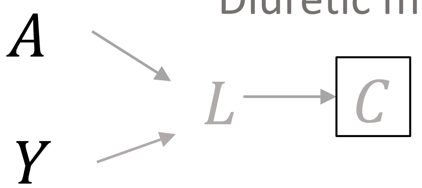
Y

Smoking

L Heart disease

$A \perp Y \quad A \not\perp Y \mid L$

L is a Collider


Genotype

A

Y

Smoking

Diuretic medication

$L \rightarrow \boxed{C}$

Heart disease

$A \perp Y \quad A \not\perp Y \mid C$

C is a consequence of the Collider

D-separation

Path: A path between two nodes X_1 and X_2 is a route that connects X_1 and X_2 by a sequence of edges such that the route visits no variable more than once

A path is blocked or open according to the following rules: (fine point 6.1)

- If there are no variables being conditioned on, a path is blocked if and only if two arrowheads on the path collide at some node on the path.
  $L \longrightarrow A \longrightarrow Y$: open $A \longrightarrow L \longleftarrow Y$: blocked

- Any path that contains a non-collider that has been conditioned on is blocked
  $L \longrightarrow A \longrightarrow Y$ blocked conditioning on A

- A collider that has been conditioned on does not block a path
  $A \longrightarrow L \longleftarrow Y$ open conditioning on L

- A collider that has a descendant that has been conditioned on does not block a path

Two variables (nodes) are d-separated if all path between them are blocked

D-separation

Let X, Y, Z be disjoint sets in a DAG G.
X and Y are d-separated by Z (Z can be empty) if and only if
Z blocks every path from a node in X to a node in Y

$$(X \perp\!\!\!\perp Y \mid Z)_G$$

- D-separation is a purely graphical concept
- Conditional independence is a probabilistic property

D-separation and conditional independence

D-separation $\Longrightarrow$ conditional independence (Pearl 1988)

The Markov assumption implies

$$(X \perp\!\!\!\perp Y \mid Z)_G \implies X \perp Y \mid Z$$

$$X \perp Y \mid Z \;\not\Rightarrow\; (X \perp\!\!\!\perp Y \mid Z)_G$$

Faithfulness assumption:

$$X \perp Y \mid Z \implies (X \perp\!\!\!\perp Y \mid Z)_G$$

Examples of unfaithfulness

Example 1:

$$\begin{pmatrix} Y(0) \\ Y(1) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

Assume $A \perp Y(a)$, then

$$Y \mid A = 0 \sim \mathcal{N}(0,1), \quad Y \mid A = 1 \sim \mathcal{N}(0,1)$$

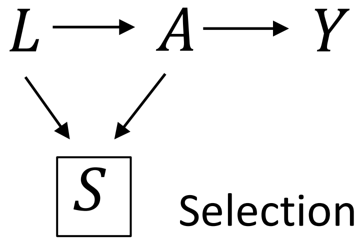
So $Y \perp A$

On the other hand, A has a causal effect on Y

so $(A \not\perp\!\!\!\perp Y)_G$

Individual causal effects can cancel out

Example 2:



Selection (matching)

- After matching $A \perp L \mid S$ by design
- $(A \perp\!\!\!\perp L \mid S)_G$ is not true as shown on DAG

Two paths cancel out