Causal Inference 2: Directed Acyclic Graphs

Ian Lundberg[1] & Kristin Liao[2]

SICSS UCLA
25 June 2024

[1]Assistant Professor, Information Science, Cornell, ilundberg@cornell.edu
[2]PhD Student, Sociology, UCLA, ktliao@g.ucla.edu

Learning goals for today

- fork structures
- collider structures
- causal reasoning and statistical independence

A hypothetical experiment in two population subgroups

A hypothetical experiment in two population subgroups

People who like exercise	People who don't like exercise

A hypothetical experiment in two population subgroups

People who like exercise	People who don't like exercise
Treatment 75% assigned an exercise coach for 1 month	**Treatment** 25% assigned an exercise coach for 1 month

A hypothetical experiment in two population subgroups

People who like exercise

Treatment
75% assigned an exercise coach for 1 month

People who don't like exercise

Treatment
25% assigned an exercise coach for 1 month

Outcome: How many pull-ups can they do?

A hypothetical experiment in two population subgroups

People who like exercise	People who don't like exercise
Treatment 75% assigned an exercise coach for 1 month	**Treatment** 25% assigned an exercise coach for 1 month

Outcome: How many pull-ups can they do?

Question for you:
Give 2 reasons why those assigned a coach can do more pull-ups

A hypothetical experiment in two population subgroups

$$\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

A hypothetical experiment in two population subgroups

$$\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

Nodes are random variables. **Edges** ($\rightarrow$) are causal relations

A hypothetical experiment in two population subgroups

$$
\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}
$$

Nodes are random variables. **Edges** ($\rightarrow$) are causal relations

The graph links causal assumptions to statistical dependence

A hypothetical experiment in two population subgroups

$$\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

Nodes are random variables. **Edges** ($\rightarrow$) are causal relations

The graph links causal assumptions to statistical dependence

In this graph, (Assigned a Coach) and (Pull-Ups) are statistically dependent because of two open paths:

A hypothetical experiment in two population subgroups

$$\begin{tikzcd} \text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\ \text{Likes Exercise} \arrow[u] \arrow[ur] & \end{tikzcd}$$

Nodes are random variables. **Edges** ($\rightarrow$) are causal relations

The graph links causal assumptions to statistical dependence

In this graph, (Assigned a Coach) and (Pull-Ups) are statistically dependent because of two open paths:

- (Assigned a Coach) $\rightarrow$ (Pull-Ups)
    - a causal path: all arrows go one direction

A hypothetical experiment in two population subgroups

$$\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

Nodes are random variables. **Edges** ($\to$) are causal relations

The graph links causal assumptions to statistical dependence

In this graph, (Assigned a Coach) and (Pull-Ups) are statistically dependent because of two open paths:

- (Assigned a Coach) $\to$ (Pull-Ups)
  - a causal path: all arrows go one direction
- (Assigned a Coach) $\leftarrow$ (Likes Exercise) $\to$ (Pull-Ups)
  - a backdoor path containing a fork

A hypothetical experiment in two population subgroups

$$\begin{tikzcd} \text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\ \text{Likes Exercise} \arrow[u] \arrow[ur] & \end{tikzcd}$$

How to study the causal effect (Assigned a Coach) $\rightarrow$ (Pull-Ups)?

A hypothetical experiment in two population subgroups

$$
\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}
$$

How to study the causal effect (Assigned a Coach) $\rightarrow$ (Pull-Ups)?

- split into two subgroups: likes exercise and don't

A hypothetical experiment in two population subgroups

$$\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

How to study the causal effect (Assigned a Coach) $\to$ (Pull-Ups)?

- split into two subgroups: likes exercise and don't
- we have a simple randomized experiment within each subgroup

A hypothetical experiment in two population subgroups

$$
\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}
$$

How to study the causal effect (Assigned a Coach) $\rightarrow$ (Pull-Ups)?

- split into two subgroups: likes exercise and don't
- we have a simple randomized experiment within each subgroup
- estimate within each subgroup

A hypothetical experiment in two population subgroups

$$\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}$$

How to study the causal effect (Assigned a Coach) $\rightarrow$ (Pull-Ups)?

- split into two subgroups: likes exercise and don't
- we have a simple randomized experiment within each subgroup
- estimate within each subgroup
- pool the two estimates for an average causal effect estimate

A hypothetical experiment in two population subgroups

$$
\begin{tikzcd}
\text{Assigned a Coach} \arrow[r] & \text{Pull-Ups} \\
\text{Likes Exercise} \arrow[u] \arrow[ur] &
\end{tikzcd}
$$

How to study the causal effect (Assigned a Coach) $\rightarrow$ (Pull-Ups)?

- split into two subgroups: likes exercise and don't
- we have a simple randomized experiment within each subgroup
- estimate within each subgroup
- pool the two estimates for an average causal effect estimate

Terminology: Identify by **conditioning** on (Likes Exercise)

Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

- I set my sprinklers to turn on at random times

Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

- I set my sprinklers to turn on at random times
- It rains at random times

Colliders: The sprinkler example

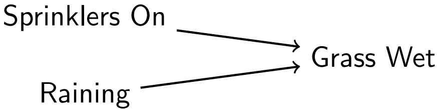
Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

- I set my sprinklers to turn on at random times
- It rains at random times
- (Sprinklers) or (Rain) can make the grass wet

Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

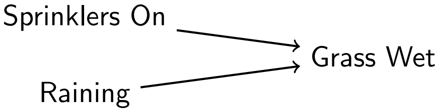
- I set my sprinklers to turn on at random times
- It rains at random times
- (Sprinklers) or (Rain) can make the grass wet



Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

- I set my sprinklers to turn on at random times
- It rains at random times
- (Sprinklers) or (Rain) can make the grass wet

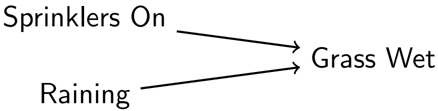


Questions for you:

- Are (Sprinklers On) and (Raining) statistically dependent?
- Are (Sprinklers On) and (Raining) statistically dependent once I restrict to times when the (Grass Wet = `TRUE`)?

Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.



Colliders: The sprinkler example

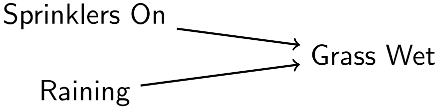
Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.

$$\begin{tikzcd} \text{Sprinklers On} \arrow[dr] & \\ & \text{Grass Wet} \\ \text{Raining} \arrow[ur] & \end{tikzcd}$$

- (Grass Wet) is a **collider** (arrows collide $\rightarrow\leftarrow$)

Colliders: The sprinkler example

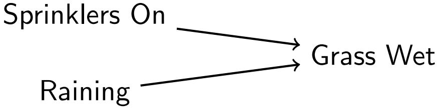
Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.



- (Grass Wet) is a **collider** (arrows collide $\rightarrow\leftarrow$)
- A collider blocks a path
  - marginal independence of (Sprinklers On) and (Raining)

Colliders: The sprinkler example

Example from Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.



- (Grass Wet) is a **collider** (arrows collide $\rightarrow\leftarrow$)
- A collider blocks a path
  - marginal independence of (Sprinklers On) and (Raining)
- Conditioning on a collider opens the path
  - conditional dependence of (Sprinklers On) and (Raining) when restricting to times when (Grass Wet $=$ `True`)

Using DAGs to identify causal effects: Game plan

Using DAGs to identify causal effects: Game plan

1. Draw a DAG

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes
     - anything causally related to any two nodes in the graph

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes
     - anything causally related to any two nodes in the graph
     - also add the causal edges

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes
     - anything causally related to any two nodes in the graph
     - also add the causal edges
2. List all paths between treatment and outcome

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes
     - anything causally related to any two nodes in the graph
     - also add the causal edges
2. List all paths between treatment and outcome
3. Choose a sufficient adjustment set:
   variables that jointly block all non-causal paths

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes
     - anything causally related to any two nodes in the graph
     - also add the causal edges
2. List all paths between treatment and outcome
3. Choose a sufficient adjustment set:
   variables that jointly block all non-causal paths
   - a path is blocked if it contains an adjusted non-collider

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
    - Create nodes for treatment and outcome
    - Add other nodes
        - anything causally related to any two nodes in the graph
        - also add the causal edges
2. List all paths between treatment and outcome
3. Choose a sufficient adjustment set:
   variables that jointly block all non-causal paths
    - a path is blocked if it contains an adjusted non-collider
    - a path is blocked if it contains an unadjusted collider
      (and no descendant of that collider is adjusted)

Using DAGs to identify causal effects: Game plan

1. Draw a DAG
   - Create nodes for treatment and outcome
   - Add other nodes
     - anything causally related to any two nodes in the graph
     - also add the causal edges
2. List all paths between treatment and outcome
3. Choose a sufficient adjustment set:
   variables that jointly block all non-causal paths
   - a path is blocked if it contains an adjusted non-collider
   - a path is blocked if it contains an unadjusted collider
     (and no descendant of that collider is adjusted)
   - otherwise unblocked

Practice

To what extent does completing a 4-year college degree affect a person's future earnings?

Effect of a 4-year degree on future earnings

1) Draw a DAG

Effect of a 4-year degree on future earnings

1) Draw a DAG

degree earnings

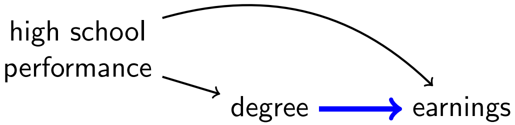
Effect of a 4-year degree on future earnings

1) Draw a DAG

degree $\longrightarrow$ earnings

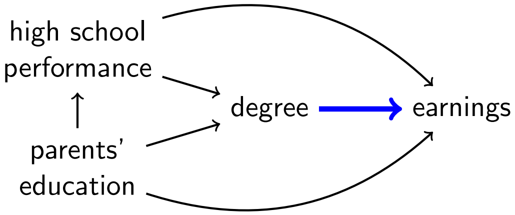
Effect of a 4-year degree on future earnings

1) Draw a DAG



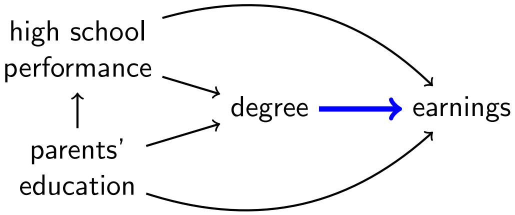
Effect of a 4-year degree on future earnings

1) Draw a DAG



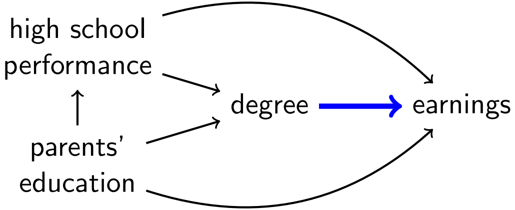
Effect of a 4-year degree on future earnings

2) List all paths between the treatment and outcome



Effect of a 4-year degree on future earnings

2) List all paths between the treatment and outcome

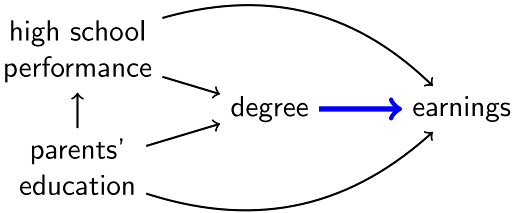


Causal paths

(degree) $\rightarrow$ (earnings)

Effect of a 4-year degree on future earnings

2) List all paths between the treatment and outcome



Causal paths

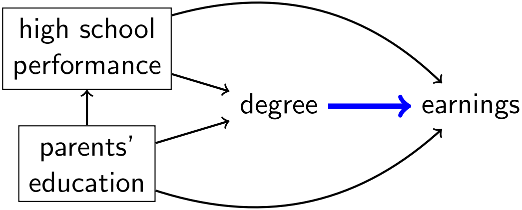
(degree) $\rightarrow$ (earnings)

Backdoor paths

(degree) $\leftarrow$ (high school performance) $\rightarrow$ (earnings)
(degree) $\leftarrow$ (parents' education) $\rightarrow$ (earnings)
(degree) $\leftarrow$ (high school performance) $\leftarrow$ (parents' education) $\rightarrow$ (earnings)

Effect of a 4-year degree on future earnings

3) Choose a sufficient adjustment set
{high school performance, parents' education}



Causal paths

(degree) → (earnings)

Backdoor paths

(degree) ← $\boxed{\text{high school performance}}$ → (earnings)

(degree) ← $\boxed{\text{parents' education}}$ → (earnings)

(degree) ← $\boxed{\text{high school performance}}$ ← $\boxed{\text{parents' education}}$ → (earnings)

DAGs: A promising path

- DAGs connect causal theories to statistical dependence
- Statistical dependence arises through causal paths
- Paths may contain two key structures
  - forks: $A \leftarrow B \rightarrow C$
    (A and C dependent if B unadjusted)
  - colliders: $A \rightarrow B \leftarrow C$
    (A and C dependent if B adjusted)
- Causal identification goal:
  choose a sufficient adjustment set so only the causal path of interest remains open
- Experimental analog:
  Among units who are identical on the sufficient adjustment set, we have a simple randomized experiment

DAGs: Words of warning

DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

- Your claim:
  If this is the DAG,
  then adjusting for $\vec{X}$ identifies the effect

DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

- Your claim:
  If this is the DAG,
  then adjusting for $\vec{X}$ identifies the effect

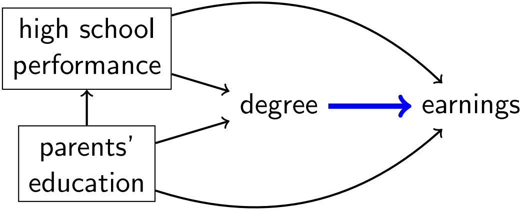
It is important to reason about when the DAG may not hold

DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

- Your claim:
  If this is the DAG,
  then adjusting for $\vec{X}$ identifies the effect

It is important to reason about when the DAG may not hold

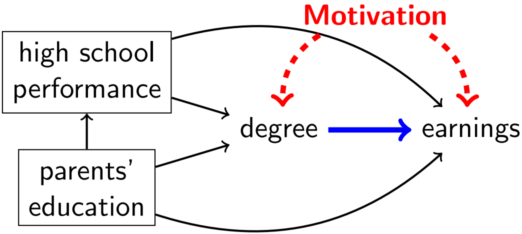


DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

- Your claim:
  If this is the DAG,
  then adjusting for $\vec{X}$ identifies the effect

It is important to reason about when the DAG may not hold

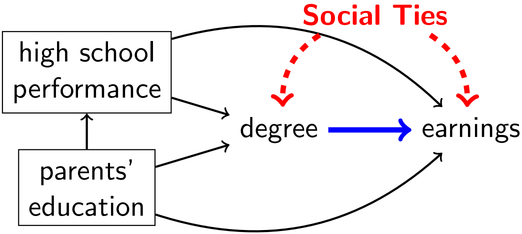


DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

- Your claim:
  If this is the DAG,
  then adjusting for $\vec{X}$ identifies the effect

It is important to reason about when the DAG may not hold

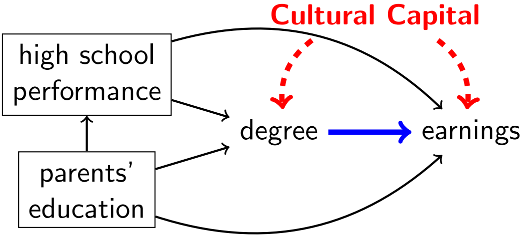


DAGs: Words of warning

Inference is only valid to the degree that the DAG holds

- Your claim:
  If this is the DAG,
  then adjusting for $\vec{X}$ identifies the effect

It is important to reason about when the DAG may not hold



Resources to learn more

- Hernán, M.A., & J.M. Robins. 2020.
  Causal Inference: What If?
  Boca Raton: Chapman & Hall / CRC.
- Pearl, J., & Mackenzie, D. (2018).
  The Book of Why: The New Science of Cause and Effect.
  Basic Books.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016).
  Causal Inference in Statistics: A Primer.
  John Wiley & Sons.
- Pearl, J. (2000).
  Causality.
  Cambridge University Press.