Causal Inference with Graphical Models

Jose M. Peña
STIMA, IDA, LiU

Lecture 0: Introduction

Causal Reasoning

- We want to compute the causal effect of an **intervention**, e.g.

$$p(cholesterol|\boldsymbol{do}(exercise)).$$

- **Intervention**: Fixing the value of a variable (for the whole population) so that it is no longer governed by its natural causes.
- **Observation**: Focus on the subpopulation that attains a particular value for a variable, e.g.

$$p(cholesterol|exercise).$$

- **Randomized controlled trials**: Gold standard for assessing causal effects, but they are not always feasible, e.g. the treatment/intervention may be too costly or prohibited due to ethical considerations.
- Can we compute causal effects from **observational** data and, thus, **without** performing interventions ? Yes, but not always.

No Causation without Manipulation ?

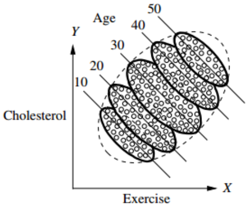


Figure 1.1: Results of the exercise-cholesterol study, segregated by age

- $p(cholesterol|\mathbf{do}(exercise)) = f(p(cholesterol, exercise, age))$?
- $p(cholesterol|\mathbf{do}(exercise)) = p(cholesterol|exercise)$?

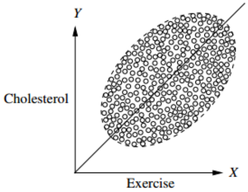


Figure 1.2: Results of the exercise-cholesterol study, unsegregated. The data points are identical to those of Figure 1.1, except the boundaries between the various age groups are not shown

No Causation without Manipulation ?

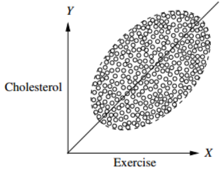


Figure 1.2: Results of the exercise-cholesterol study, unsegregated. The data points are identical to those of Figure 1.1, except the boundaries between the various age groups are not shown

- Due to the **confounder** Age,

$$p(cholesterol|\mathbf{do}(exercise)) \neq p(cholesterol|exercise).$$

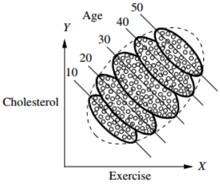
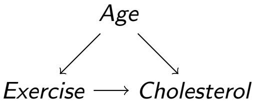


Figure 1.1: Results of the exercise-cholesterol study, segregated by age



- Instead,

$$p(cholesterol|\mathbf{do}(exercise)) = \sum_{age} p(cholesterol|exercise, age)p(age).$$

No Causation without Manipulation ?

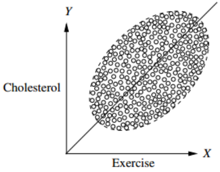


Figure 1.2: Results of the exercise-cholesterol study, unsegregated. The data points are identical to those of Figure 1.1, except the boundaries between the various age groups are not shown

- Due to the **confounder** Age,

$$p(cholesterol|\mathbf{do}(exercise)) \neq p(cholesterol|exercise).$$

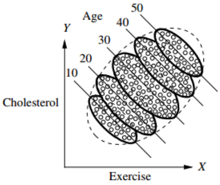
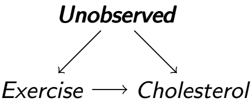


Figure 1.1: Results of the exercise-cholesterol study, segregated by age



- Now,

$$p(cholesterol|\mathbf{do}(exercise)) \neq f(p(cholesterol, exercise)).$$

No Causation without Manipulation ?

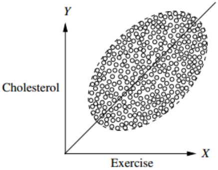


Figure 1.2: Results of the exercise-cholesterol study, unsegregated. The data points are identical to those of Figure 1.1, except the boundaries between the various age groups are not shown

- Due to the **confounder** Age,

$$p(cholesterol|\mathbf{do}(exercise)) \neq p(cholesterol|exercise).$$

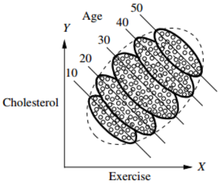
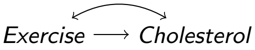


Figure 1.1: Results of the exercise-cholesterol study, segregated by age

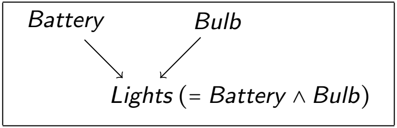


- Now,

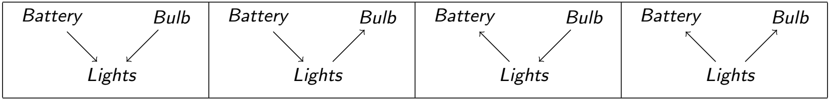
$$p(cholesterol|\mathbf{do}(exercise)) \neq f(p(cholesterol, exercise)).$$

Correlation is not Causation ?

- Can we learn a causal model of the domain at hand from **observational** data ? Yes, but not always.



- $Battery \perp Bulb | \emptyset \Rightarrow$ No edge between *Battery* and *Bulb*.
- $Battery \not\perp Lights | Bulb \Rightarrow$ Edge between *Battery* and *Lights*.
- $Bulb \not\perp Lights | Battery \Rightarrow$ Edge between *Bulb* and *Lights*.



- $Battery \not\perp Bulb | Lights \Rightarrow$ Rule out the last three models.

Correlation is not Causation ?

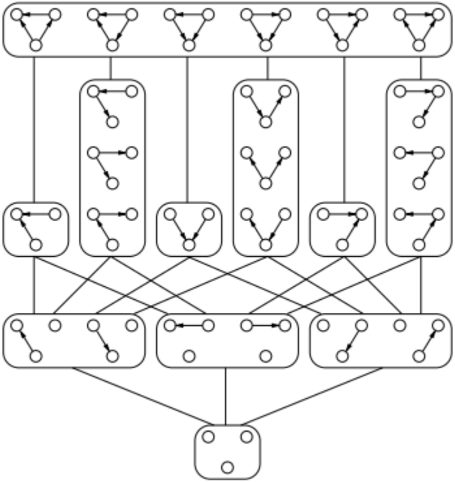
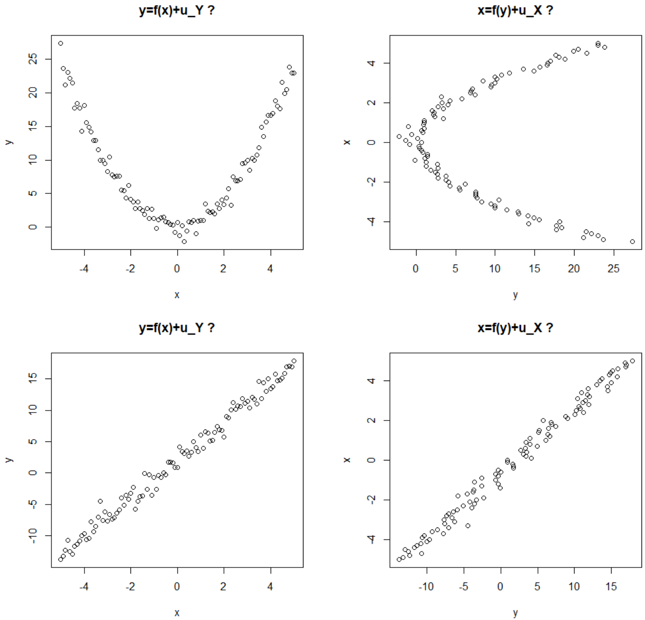


Figure 2: Hasse diagram of the space of Markov equivalence classes of Bayesian network structures over three variables.

Correlation is not Causation ?

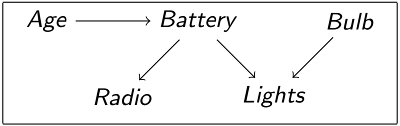
- **Assume** additive noise, i.e. $b = f(a) + u_B$ instead of $b = f(a, u_B)$.



No Causation without Manipulation ? Correlation is not Causation ?

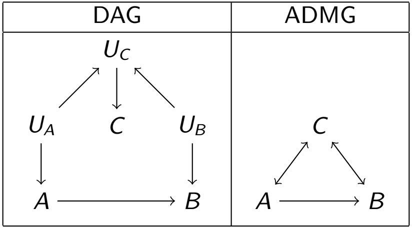
- True in general, but not always.
- The causal effect of interest may be identifiable.
- The causal model of the domain at hand may be learnable.
- How ? By using graphical models, non-linearities in the data, etc.
- The rest of this introductory lecture:
  - Directed acyclic graphs.
  - Acyclic directed mixed graphs.
  - *do*-calculus.
  - Counterfactuals.

Directed Acyclic Graphs



- Natural representation of **causal** models.
- DAGs also represent **independence** models.
- **Chain**: $Age \rightarrow Battery \rightarrow Radio$
  - $Age \not\perp Radio | \emptyset$
  - $Age \perp Radio | Battery$
- **Fork**: $Radio \leftarrow Battery \rightarrow Lights$
  - $Radio \not\perp Lights | \emptyset$
  - $Radio \perp Lights | Battery$
- **Collider**: $Battery \rightarrow Lights \leftarrow Bulb$
  - $Battery \perp Bulb | \emptyset$
  - $Battery \not\perp Bulb | Lights$
- **Chain + collider**: $Age \rightarrow Battery \rightarrow Lights \leftarrow Bulb$
  - $Age \perp Bulb | \emptyset$
  - $Age \not\perp Bulb | Lights$
  - $Age \perp Bulb | Lights, Battery$
- **Separation**: A DAG G represents the **independence** $X \perp Y | Z$ iff every path between X and Y is such that
  - some non-collider node in the path is in Z, or
  - some collider node in the path is not in Z and has no descendant in Z.

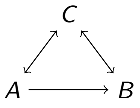
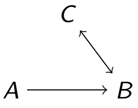
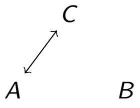
Acyclic Directed Mixed Graphs

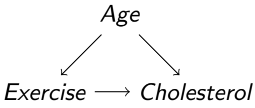
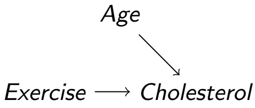


- $X \leftrightarrow Y$ is in the ADMG iff X and Y have a **confounder** in the DAG, i.e.
  - $X \leftarrow \boldsymbol{U_X} \rightarrow U_Y \rightarrow Y$ or $X \leftarrow U_X \leftarrow \cdots \leftarrow \boldsymbol{U_Z} \rightarrow \cdots \rightarrow U_Y \rightarrow Y$ is in the DAG,
  - or equivalently $U_X \not\perp U_Y | \varnothing$ in the DAG.

Interventions

- **Intervention**: Fixing the value of a variable (for the whole population) so that it is no longer governed by its natural causes.

Observational	Post-interventional $do(a)$	Post-interventional $do(b)$
		

Observational	Post-interventional $do(exercise)$
	

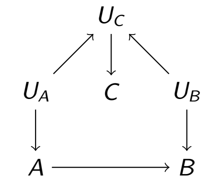
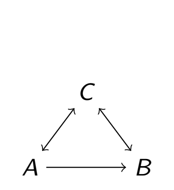
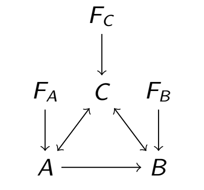
- Then,

$$p(cholesterol, age|do(exercise)) = p(c|a, do(e))p(a|do(e)) = p(c|a, e)p(a)$$

and thus

$$p(c|do(e)) = \sum_a p(c, a|do(e)) = \sum_a p(c|a, e)p(a).$$

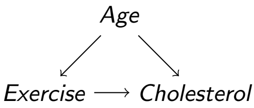
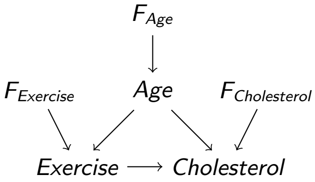
do-Calculus

DAG	ADMG	Extended ADMG
		

- Set of rules to be applied repeatedly for causal effect identification:
  - Rule 1 (**insertion/deletion of observations**):
    $p(y|do(x), \mathbf{z}, w) = p(y|do(x), w)$ if $Y \perp Z | X \cup W || X$.
  - Rule 2 (**intervention/observation exchange**):
    $p(y|do(x), \mathbf{do(z)}, w) = p(y|do(x), \mathbf{z}, w)$ if $Y \perp F_Z | X \cup W \cup Z || X$.
  - Rule 3 (**insertion/deletion of interventions**):
    $p(y|do(x), \mathbf{do(z)}, w) = p(y|do(x), w)$ if $Y \perp F_Z | X \cup W || X$.

  Where $\cdot \perp \cdot | \cdot || X$ denotes independence in the extended ADMG after intervention on X:
  - Delete all the directed and bidirected edges into X.
  - Apply the separation criterion.
- There is a sound and complete algorithm to apply the rules.

do-Calculus

DAG	Extended ADMG
	

- Then,

$$p(cholesterol, age|do(exercise)) = p(c|a, do(e))p(a|do(e)) = p(c|a, e)p(a)$$

because

$$p(c|a, do(e)) = p(c|a, e) \text{ by } \textbf{rule 2} \text{ and } C \perp F_E | A \cup E$$

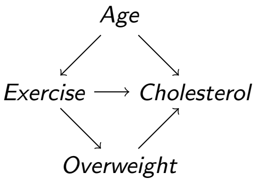
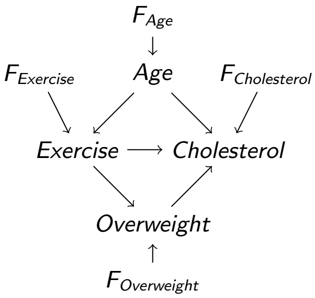
and

$$p(a|do(e)) = p(a) \text{ by } \textbf{rule 3} \text{ and } A \perp F_E | \varnothing.$$

- Thus,

$$p(c|do(e)) = \sum_a p(c, a|do(e)) = \sum_a p(c|a, e)p(a).$$

do-Calculus

DAG	Extended ADMG
	

- Then,

$$p(c,a,o|do(e)) = p(c|a,o,do(e))p(o|a,do(e))p(a|do(e))$$
$$= p(c|a,o,e)p(o|a,do(e))p(a) = p(c|a,o,e)p(o|a,e)p(a)$$

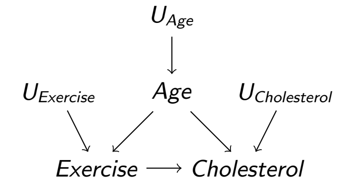
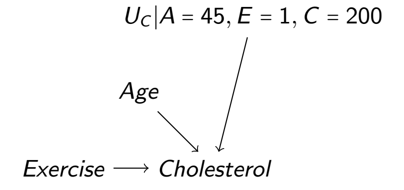
by factorization, rule 2, rule 3 and rule 2 and, thus,

$$p(c|do(e)) = \sum_a \sum_o p(c,a,o|do(e)) = \sum_a p(c|a,e)p(a).$$

- The same expression as before. That is, we adjust for A but not for O. How do we know when to adjust and when not ? **The DAG guides us**. Therefore, using/learning the correct DAG is crucial.

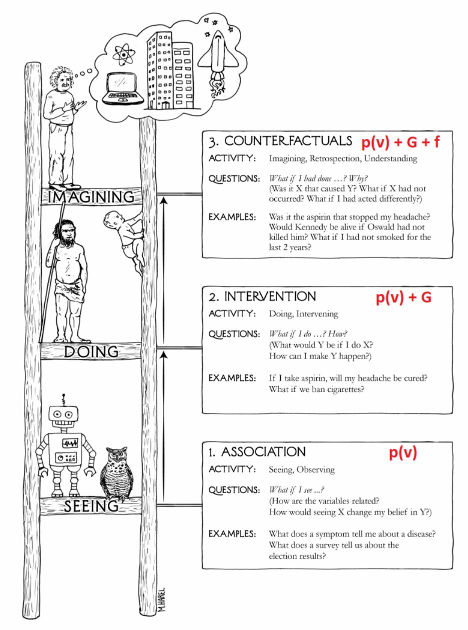
Counterfactuals

- *do*-calculus = population-level effect ≠ **individual-level** = counterfactuals.
- Individual counterfactual: What would **my** cholesterol level have been, had **I** doubled the exercise time ?

Magnified DAG	Updated DAG
	

- Subpopulation counterfactual: What would the expected cholesterol level of **individuals with high cholesterol** have been, had **they** doubled the exercise time ?
- Probability of necessity: What is the probability that **I** would have had high cholesterol level without doing exercise, given that **I** have low cholesterol level and do exercise ?
- Probability of sufficiency: What is the probability that **I** would have had low cholesterol level by doing exercise, given that **I** have high cholesterol level and do not do exercise ?

The Ladder of Causation



Summary

- No causation without manipulation ? Correlation is not causation ? True in general, but not always.
- Can we compute causal effects from **observational** data and, thus, **without** performing interventions ? Yes, but not always.
- Can we learn a causal model of the domain at hand from **observational** data ? Yes, but not always.
- How ? By using graphical models, non-linearities in the data, etc.
- In this introductory lecture, we have briefly covered the following topics:
  - Directed acyclic graphs.
  - Acyclic directed mixed graphs.
  - *do*-calculus.
  - Counterfactuals.
- Recommended readings:

  - Darwiche, A. Human-Level Intelligence or Animal-Like Abilities ? *Communications of the ACM*, 61:56-67, 2018.
  - Pearl, J. The Seven Tools of Causal Inference with Reflections on Machine Learning. *Communications of the ACM*, 62:54-60, 2019.
  - Pearl, J. and Mackenzie, D. *The Book of Why: The New Science of Cause and Effect*. Basic Books, 2018.

Thank you