



Causal inference and discovery

Pierre Gentine + Adam Massmann + Jakob Runge

TRANSCENDING DISCIPLINES, TRANSFORMING LIVES

Causal Effects

Introduction

Causal inference:

- Going beyond correlation, but at cause and effects
- Introduction to directed graphs
- Do calculus
- Transportability

Introduction

Causality in many fields

- Economics
- Science
- Politics
- Medicine

Introduction

Causality in economics



There is a gender gap in earnings for the alumni at every top university, although the size of the difference varies greatly. Credit Matt Rourke/Associated Press

Gaps in Earnings Stand Out in Release of College Data

SEPT. 13, 2015
nytimes.com | Sept. 13, 2015

David Card. The causal effect of education on earnings (1999)



Conley and Heerwig. The Long-Term Effects of Military Conscription on Mortality: Estimates From the Vietnam-Era Draft Lottery (2012)

Introduction

Causality in human behavior



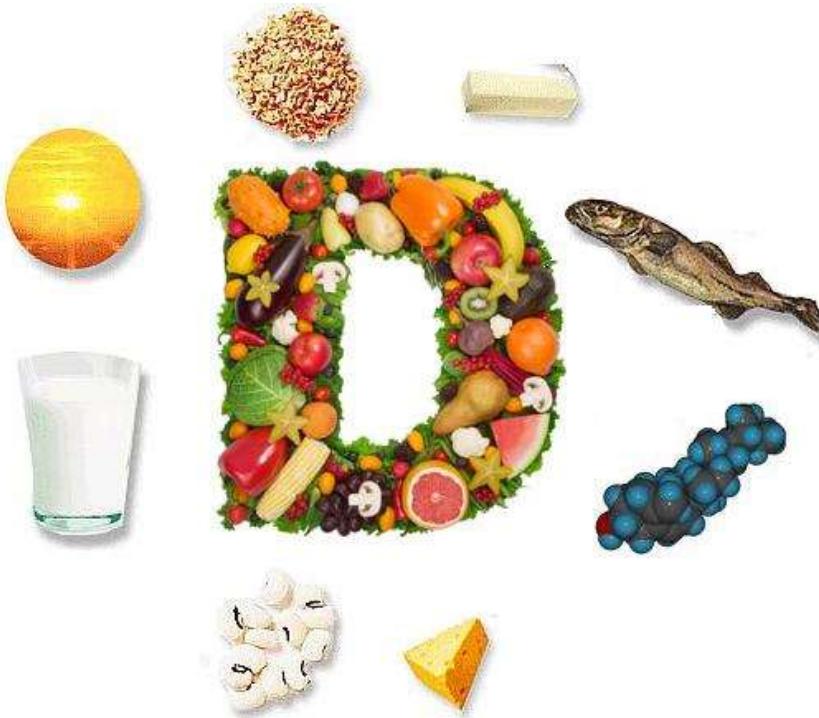
Thistlewaite and Campbell. Effect of public recognition
of scholastic achievement (1960)



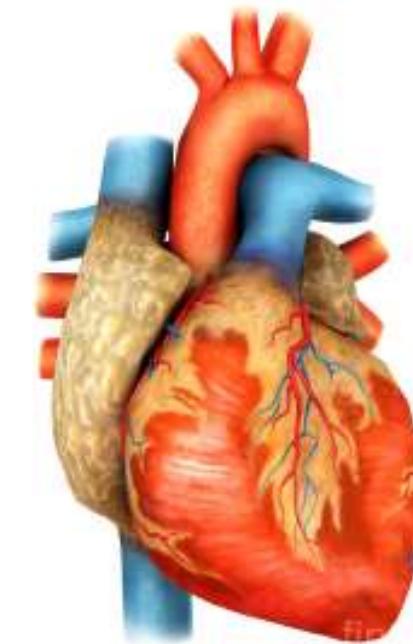
Christakis and Fowler. The collective dynamics
of smoking in a large social network (2008)

Introduction

Causality in human behavior



Effect of Vitamin D deficiency on colon cancer



Effect of heart attack surgery on long-term health of patient

Counterfactual reasoning

Correlation question: How well can X predict Y?

- Machine learning, Statistical estimation.

Interventionist question: If X is changed to X', what will be the value of Y?

- Experiments, Reinforcement learning, Contextual bandits.

Counterfactual question: If X would have been X', what would be the value of Y?

- Today's focus.

Why is causal inference hard?

- Simpson’s paradox

The language of graphical models

- Backdoor criterion
- Frontdoor criterion

Common approaches for causal inference

- Conditioning
- Mechanism-based
- Natural Experiments

Example: Estimating causal impact of recommender systems

Simpson's paradox

Setup:

There is a drug, which we call *drug 1*, that is supposed to lower the risk of having a heart attack. The following table summarizes data comparing a group of subjects that have and that have not taken the drug.¹

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

Question:

Is it a good drug? That is, should one take this drug to reduce the risk of heart attack?

¹Example taken from J. Pearl, D. Mackenzie: *The Book of Why. The New Science of Cause and Effect*, Penguin Books (2018).

Simpson's paradox

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

Risk of having a heart attack among female subjects:

$$\text{Risk in control group: } 0.04 = \frac{1}{1+19}$$

$$\text{Risk in treatment group: } 0.075 = \frac{3}{3+37}$$

⇒ Drug is bad for females

Simpson's paradox

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

Risk of having a heart attack among male subjects:

$$\text{Risk in control group: } 0.30 = \frac{12}{12+28}$$

$$\text{Risk in treatment group: } 0.40 = \frac{8}{8+12}$$

⇒ **Drug is bad for males**

Simpson's paradox

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

Risk of having a heart attack among all subjects:

$$\text{Risk in control group: } 0.22 \simeq \frac{13}{13+47}$$

$$\text{Risk in treatment group: } 0.18 \simeq \frac{11}{11+49}$$

⇒ **Drug is good for people?**

Simpson's paradox

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

Paradoxical finding:

The drug is bad for females, bad for males, but good for people.

This is clearly wrong!

What do you think? Is the drug good or bad?

Answer:

The drug is bad!

Simpson's paradox

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

How the numbers work out:

- Males have a higher risk of heart attack.
- The proportion of males is higher in the control than in the treatment group

But still: Why is the drug bad?

The drug does not influence gender, but gender influences whether the drug is taken. This leads to a spurious, non-causal dependence, such that the effect has to be analyzed separately for female and male subjects.

Simpson's paradox

Setup:

Now consider a second drug, *drug 2*. It is supposed to reduce the risk of heart attacks by lowering blood pressure. High blood pressure is a known cause of heart attacks.

	Control Group (No Drug)		Treatment Group (Took Drug)	
Blood pressure	Heart attack	No heart attack	Heart attack	No heart attack
Low	1	19	3	37
High	12	28	8	12
Total	13	47	11	49

Question:

Is this a good drug, i.e., should one take this drug to reduce the risk of heart attack?

Answer:

The drug is good! Since the drug influences blood pressure we have to assess the drug based on all subjects.

Simpson's paradox

- Questions about *effects* and *influences* cannot be answered from the data alone, hence these are not statistical questions.
- Such questions can only be answered when making additional assumptions about the data generating process, which are of causal nature.

General rule:

- There is a sharp division between statistical and causal questions.
- Answering causal questions always requires some causal assumptions.

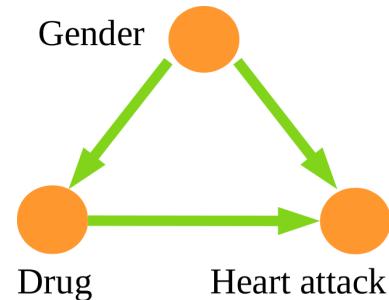
Simpson's paradox

Can we argue more conceptually why the data has to be analyzed differently for the two drugs?

Drug 1: Analyze in subpopulations

	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Female	1	19	3	37
Male	12	28	8	12
Total	13	47	11	49

Direction of causal influence:



Must condition on *Gender* to **control for confounding** of *Drug* and *Heart attack*

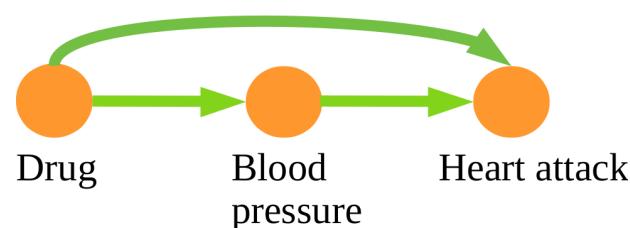
Simpson's paradox

Can we argue more conceptually why the data has to be analyzed differently for the two drugs?

Drug 2: Analyze in full population

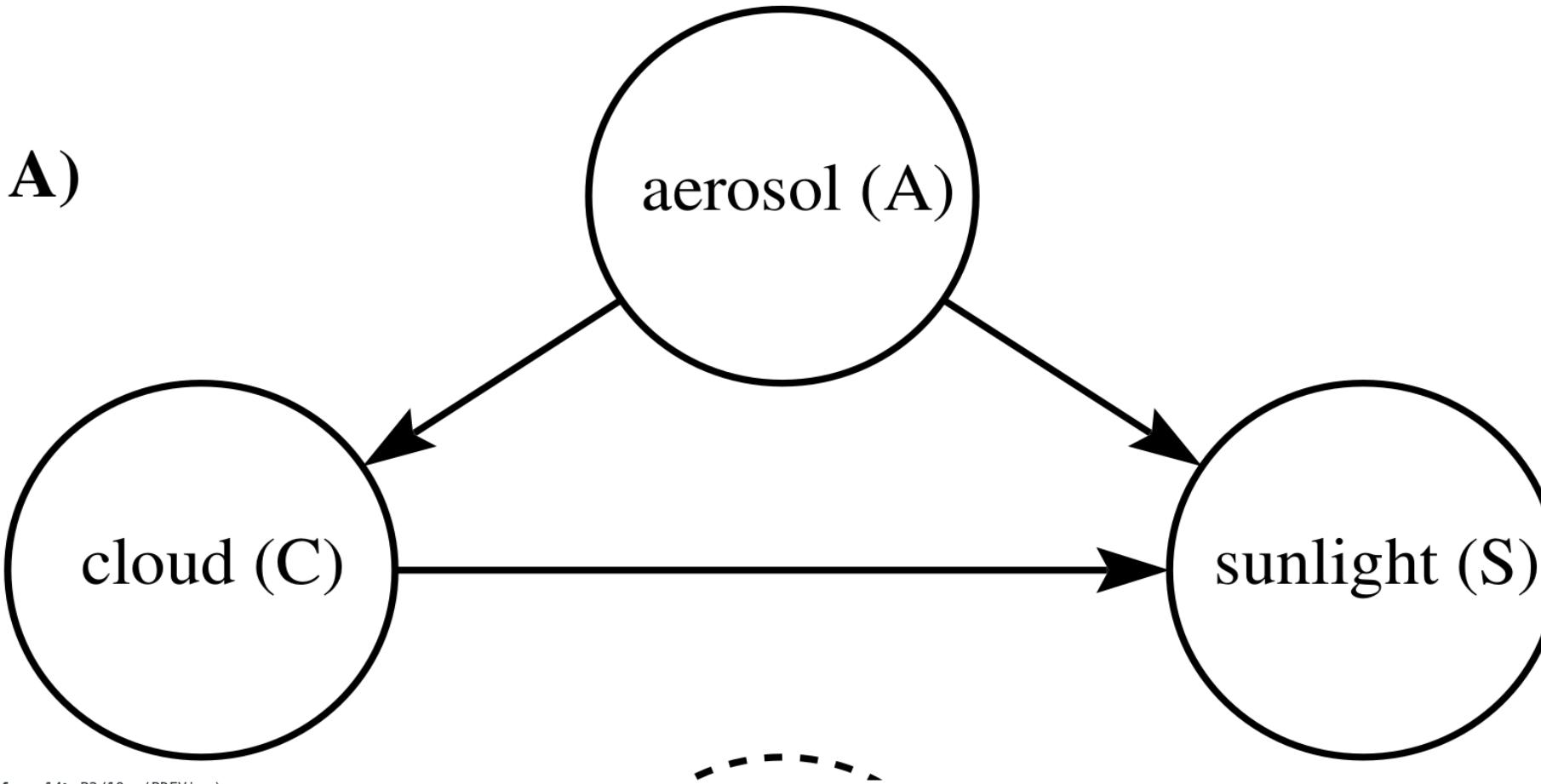
	Control Group (No Drug)		Treatment Group (Took Drug)	
	Heart attack	No heart attack	Heart attack	No heart attack
Blood pressure				
Low	1	19	3	37
High	12	28	8	12
Total	13	47	11	49

Direction of causal influence:



Must not condition on the *Blood pressure*, which **mediates** (part of) the effect of *Drug* on *Heart attack*.

Graphs



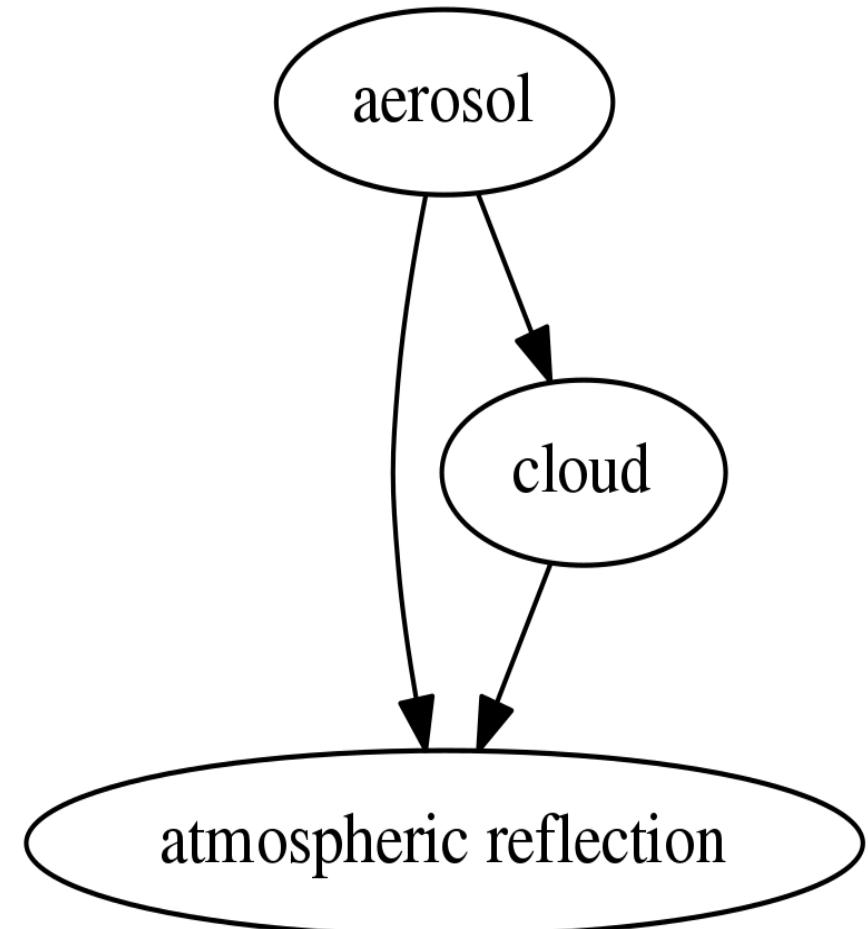
Graphs encode assumptions about causal dependencies in the system

Probabilistic Graphs

Probabilistic graphical models: assumptions and communication

Probabilistic graphical models are **directed acyclic graphs (DAGs)** encoding assumptions about independence and dependence

Graphs are nice because they communicate a general meaning and a technical meaning:

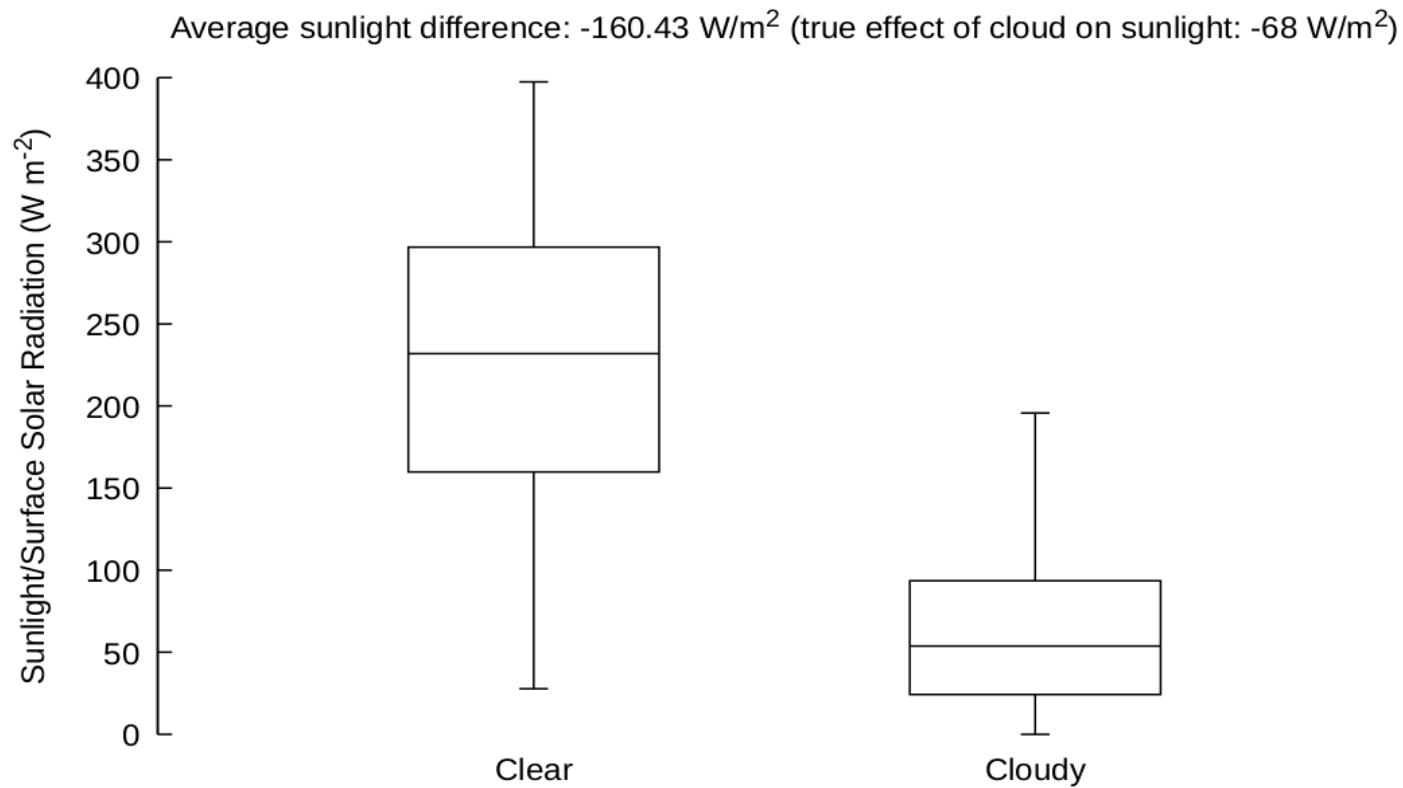


$$p(\text{cloud}, \text{aerosol}, \text{reflection}) = p(\text{reflection} | \text{cloud}, \text{aerosol}) p(\text{cloud} | \text{aerosol}) p(\text{aerosol})$$

How would one estimate a causal effect of clouds on surface downwelling solar radiation?

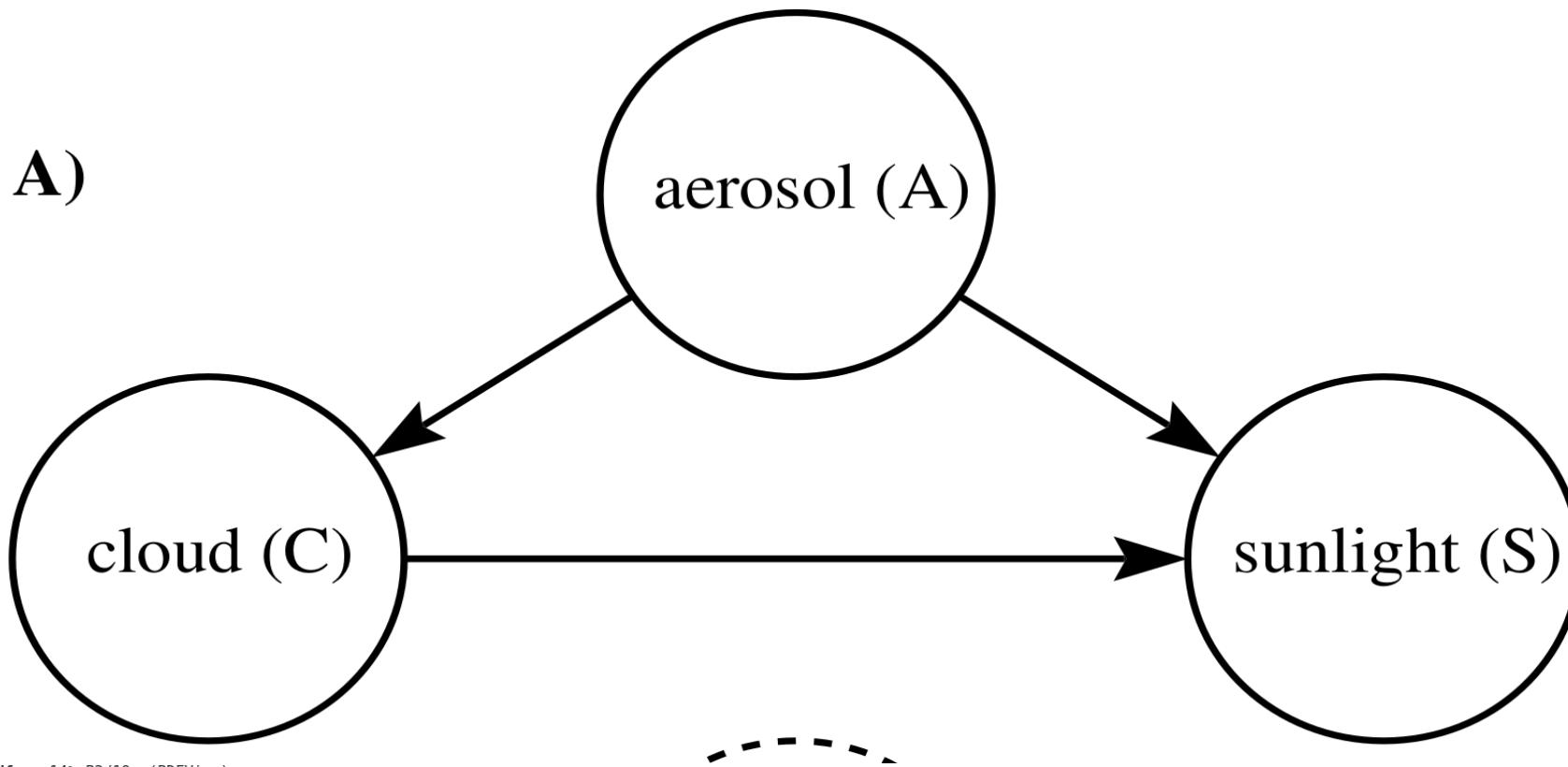
Causal effects

Naive approach (and/or when aerosol data are not available): bin observations by cloudy and clear day and compare the values of sunlight



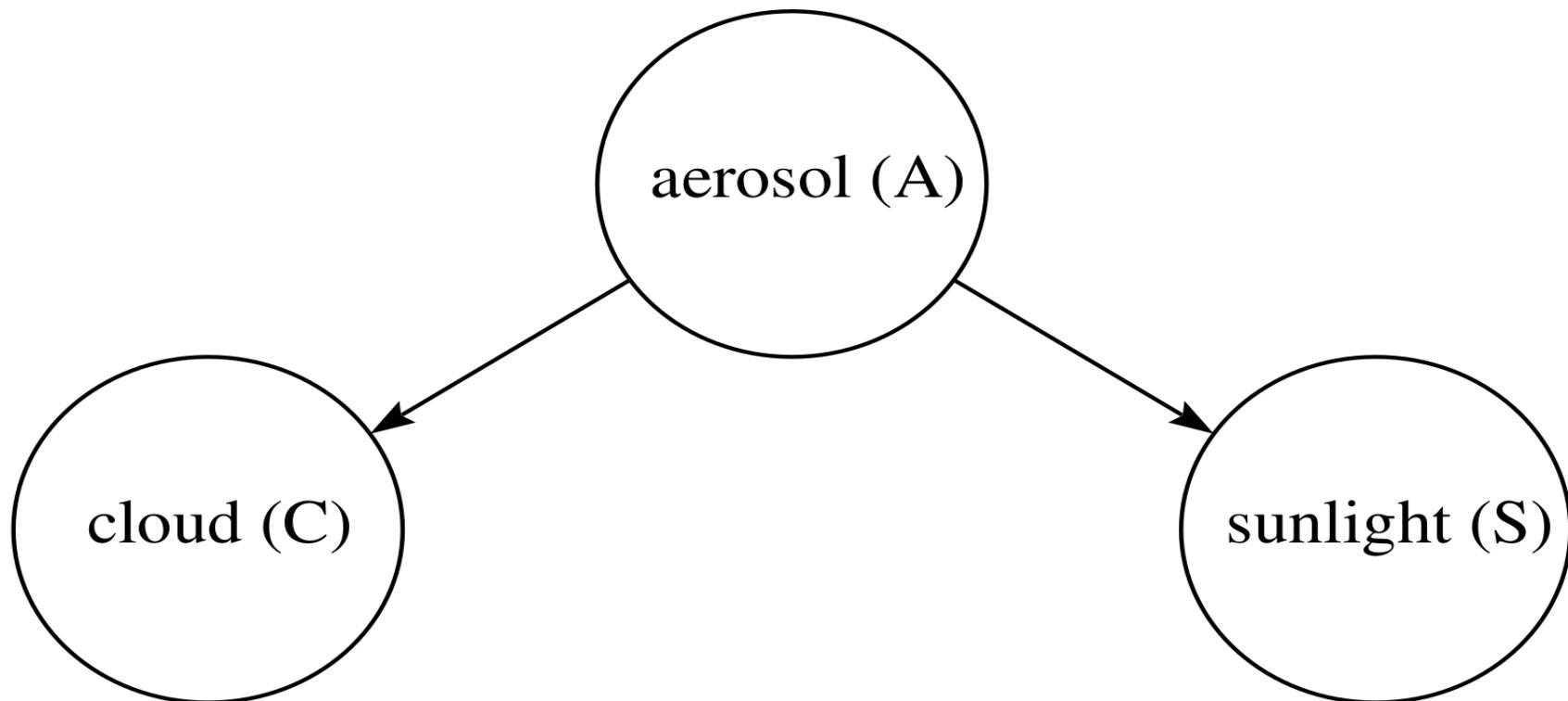
Causal effects

What went wrong?



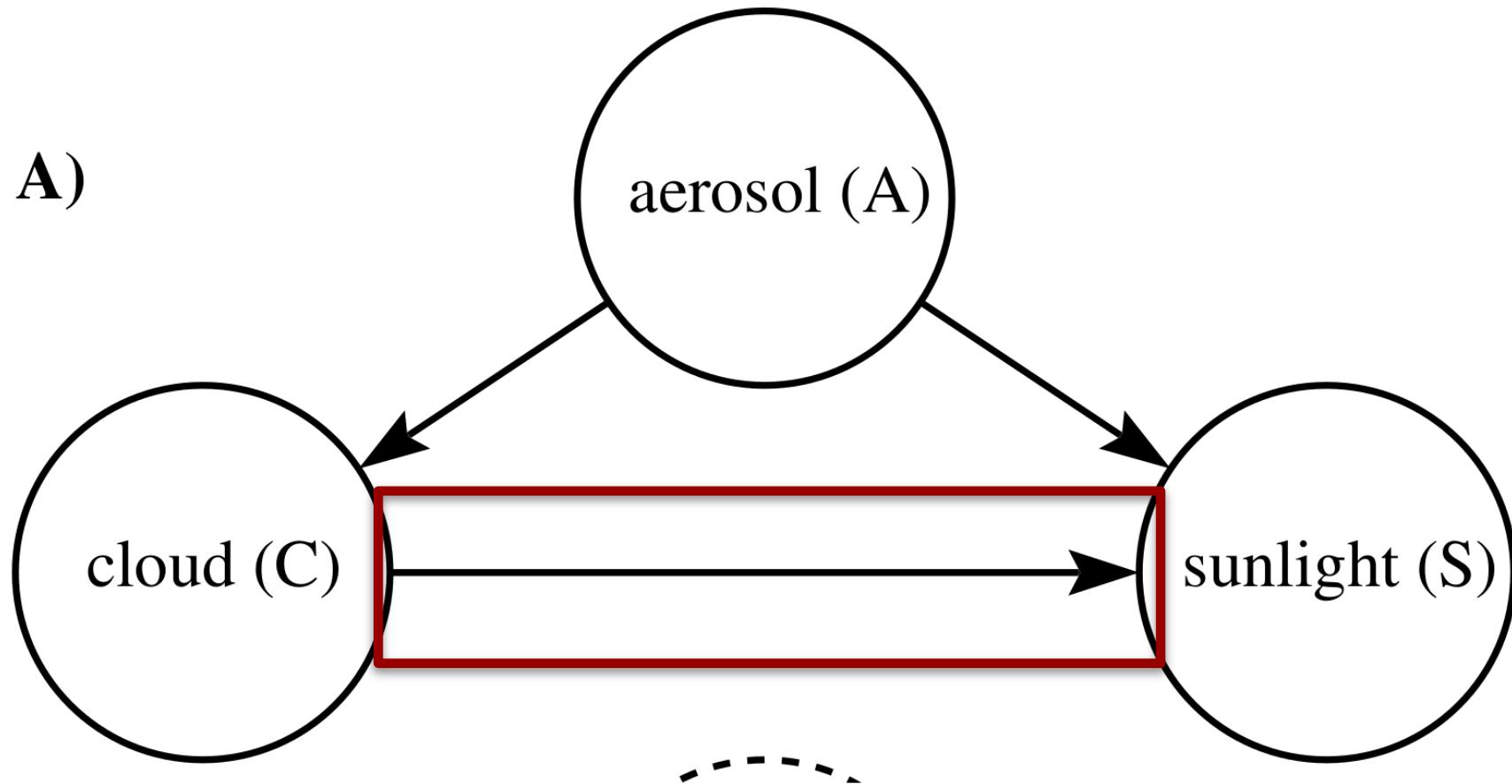
Causal effects

Aerosol induces covariability in cloud and sunlight that has nothing to do with the causal link from cloud to aerosol!



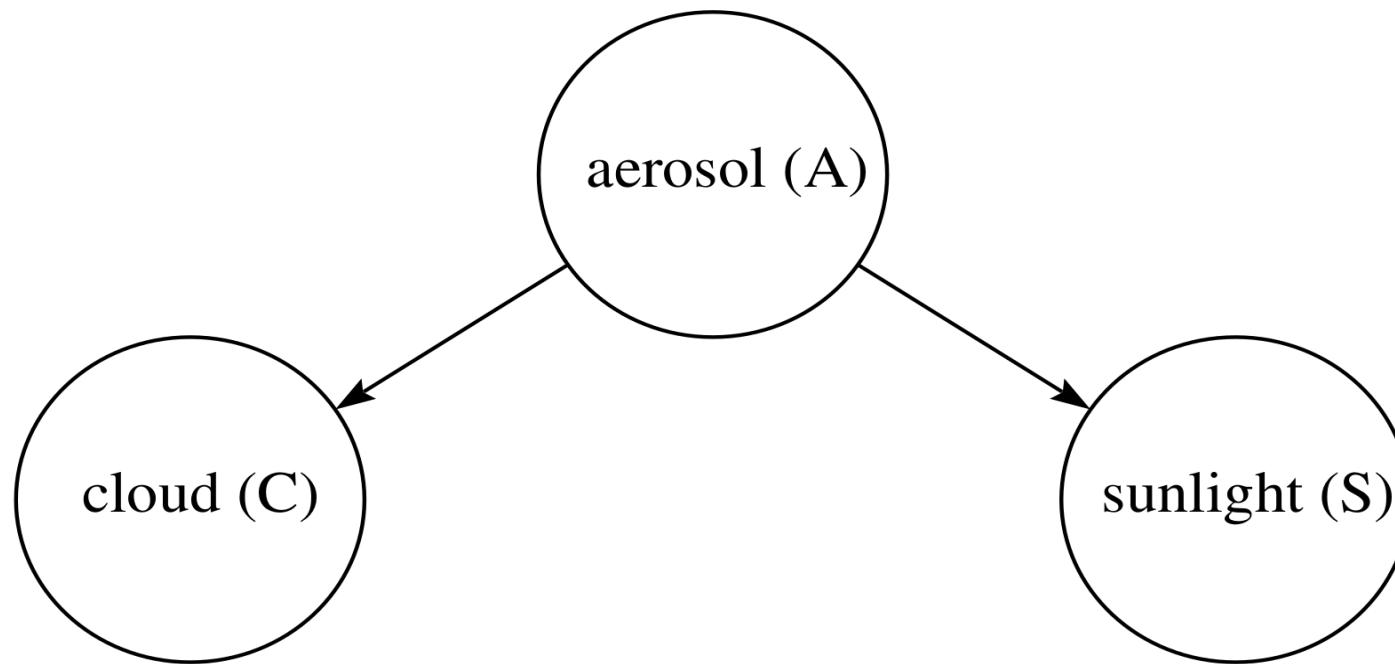
Causal effects

Goal: we want to isolate the variability in aerosol due to the causal link from cloud



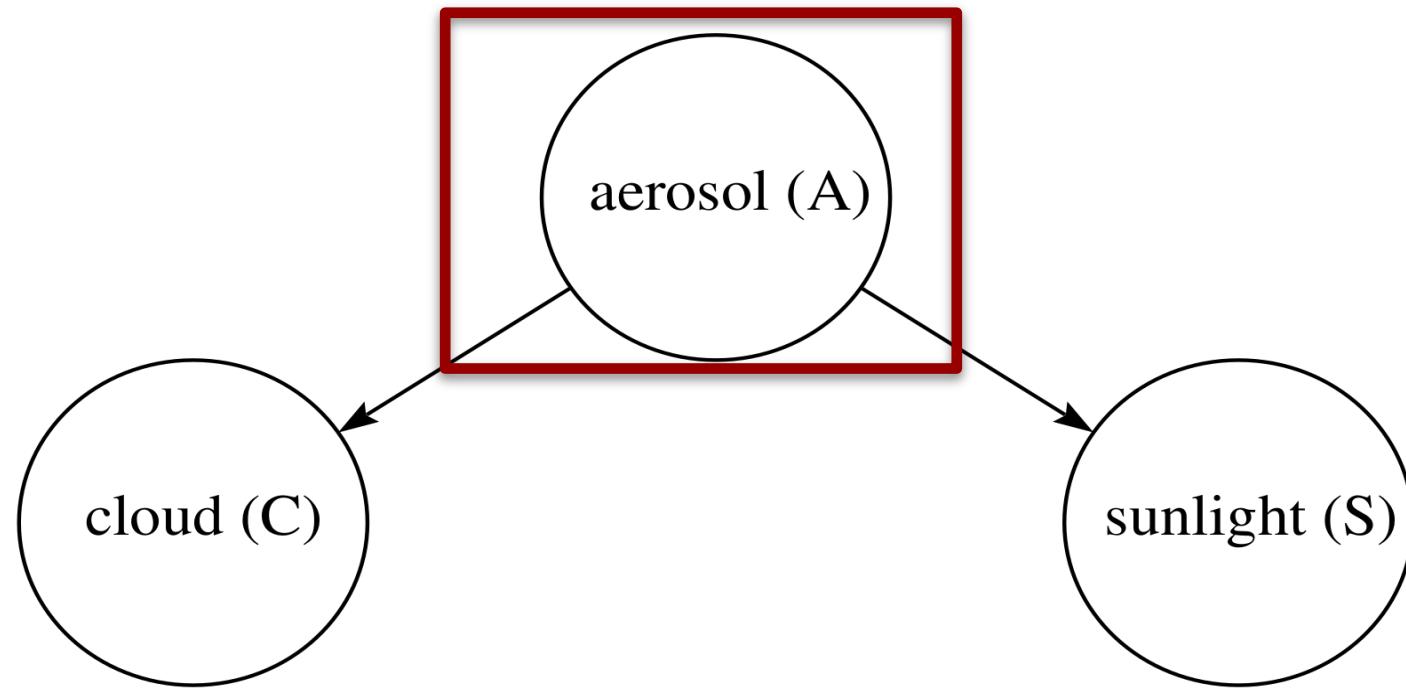
Introduction

Intuitive graphical approach: remove causal paths from the cause to the children of causes and ask, are the cause and effect independent in the “mutilated” graph?



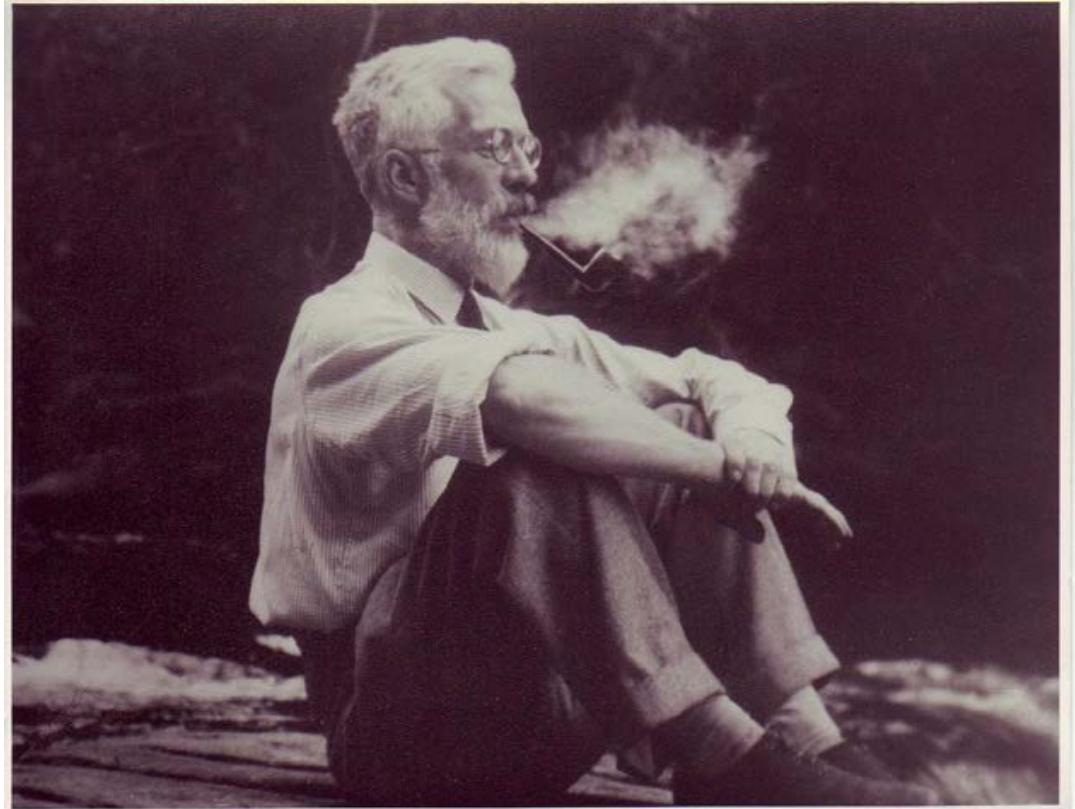
Causal effects

If cause and effect are not independent, observe variables such that they are independent



Causal effects

Does smoking cause lung cancer?



Cancer and Smoking

The curious associations with lung cancer found in relation to smoking habits do not, in the minds of some of us, lend themselves easily to the simple conclusion that the products of combustion reaching the surface of the bronchus induce, though after a long interval, the development of a cancer.

*Ronald A. Fisher
Nature 1958;182(4635):596.*

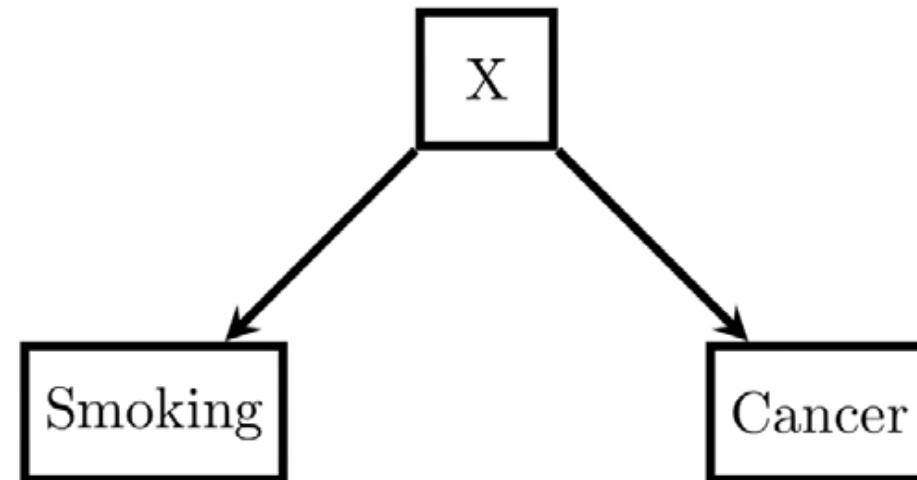
Causal effects

Directed Acyclic Graph (DAG) to express causal relationships

Smoking causes cancer



Patient characteristic X causes both smoking and cancer

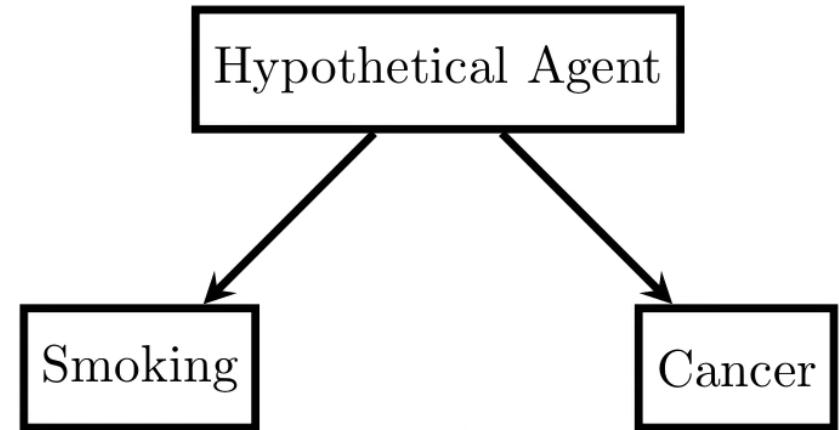


Smoking and lung cancer: recent evidence and a discussion of some questions*

Jerome Cornfield,¹ William Haenszel,² E. Cuyler Hammond,³ Abraham M. Lilienfeld,⁴
Michael B. Shimkin⁵ and Ernst L. Wynder⁶

J. Nat. Cancer Inst. **22**:173–203, 1959

“The magnitude of the excess lung-cancer risk among cigarette smokers is so great that the results can not be interpreted as arising from an indirect association of cigarette smoking with some other agent or characteristic, since this hypothetical agent would have to be at least as strongly associated with lung cancer as cigarette use; no such agent has been found or suggested.“



Smoking and lung cancer: recent evidence and a discussion of some questions*

Jerome Cornfield,¹ William Haenszel,² E. Cuyler Hammond,³ Abraham M. Lilienfeld,⁴
Michael B. Shimkin⁵ and Ernst L. Wynder⁶

J. Nat. Cancer Inst. **22**:173–203, 1959

“The consistency of all the epidemiologic and experimental evidence also supports the conclusion of a causal relationship with cigarette smoking, while there are serious inconsistencies in reconciling the evidence with other hypotheses which have been advanced.“



Causal effects

Association, Prediction, Causality

- Carrying a lighter is strongly associated with lung cancer
- Whether or not somebody carries a lighter is predictive of lung cancer
- But this is not a causal relationship!



In some settings, having a good predictive model may be sufficient.

In others, causality is of main interest

Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

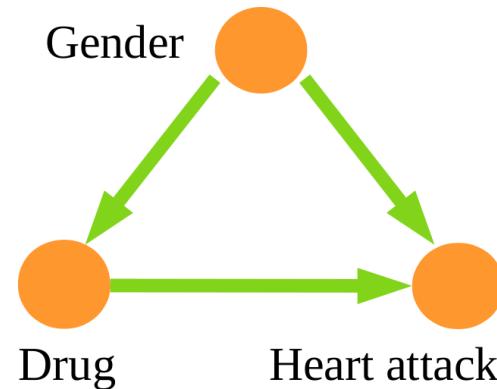
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



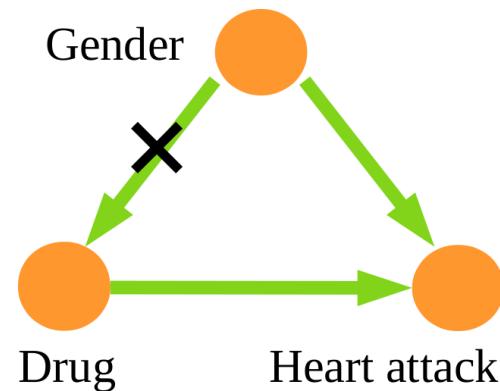
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



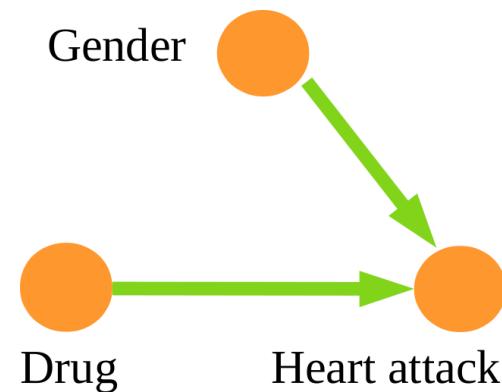
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



The gender of a subject does not longer influence whether a subject belongs to the treatment or control group.

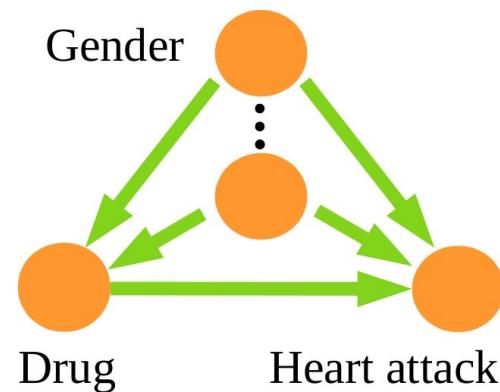
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



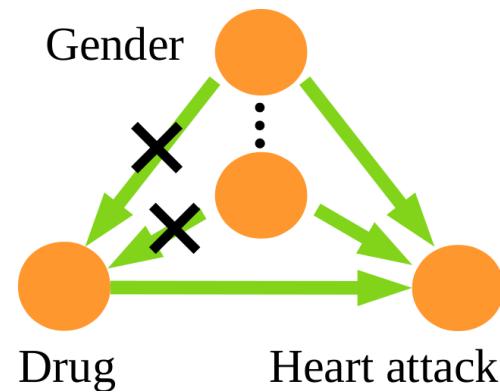
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



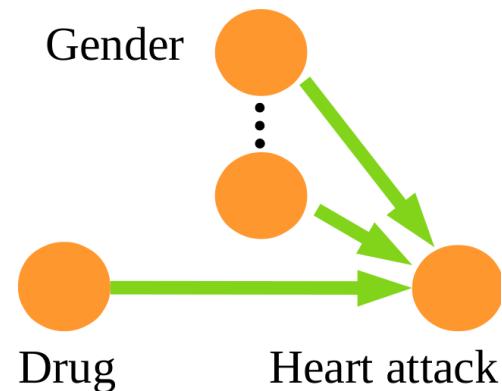
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



In fact, all confounders are eliminated!

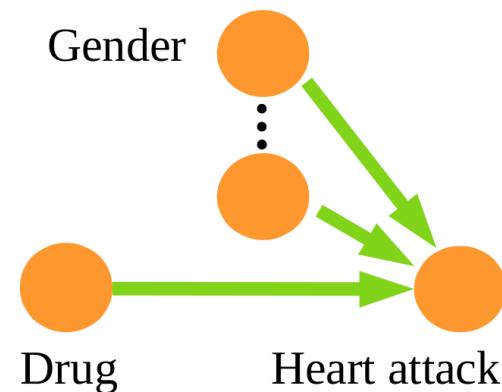
Randomized Control Trials

Question:

How can one, in general, determine whether a certain drug is helpful in reducing the risk of developing a certain disease?

Randomized control trials:

Randomly assign the subjects to the treatment and control group.



In fact, all confounders are eliminated! Therefore, **in RCT data, causal effects can be assessed by statistical proportions.**

Experimentation

Abstraction of the question:

Consider two events or variables X and Y . How can we determine whether X causes Y ?

Experimentation:

Experimentally manipulate X while keeping all other conditions exactly the same. If this results in a change of Y , X causes Y .

Such an idealization of experimental manipulation is referred to as an *intervention on X* .

Experimentation

Abstraction of the question:

Consider two events or variables X and Y . How can we determine whether X causes Y ?

Experimentation:

Experimentally manipulate X while keeping all other conditions exactly the same. If this results in a change of Y , X causes Y .

Such an idealization of experimental manipulation is referred to as an *intervention on X* .

Example: Lawn with sprinkler

- Turning the sprinkler on makes the lawn wet.
- Making the lawn wet does not turn on the sprinkler.

Experimentation

Abstraction of the question:

Consider two events or variables X and Y . How can we determine whether X causes Y ?

Experimentation:

Experimentally manipulate X while keeping all other conditions exactly the same. If this results in a change of Y , X causes Y .

Such an idealization of experimental manipulation is referred to as an *intervention on X* .

Example: Lawn with sprinkler

- Turning the sprinkler on makes the lawn wet.
- Making the lawn wet does not turn on the sprinkler.

We need to imagine the lawn being made wet without changing something else, in particular without turning the sprinkler on

Challenges in Context of Climate Sciences

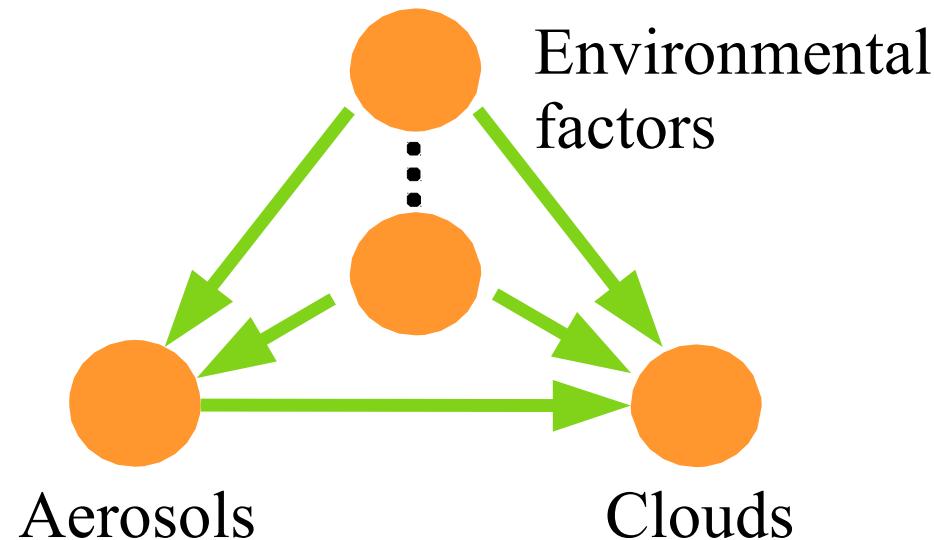
When trying to apply these concepts in the context of environmental sciences, we face two fundamental challenges.

Challenges in Context of Climate Sciences

When trying to apply these concepts in the context of environmental sciences, we face two fundamental challenges.

Challenge: Experimentation in real environment is not possible

For most parts it is impossible, and else arguably ethically questionable, to deliberately intervene on environmental variables.

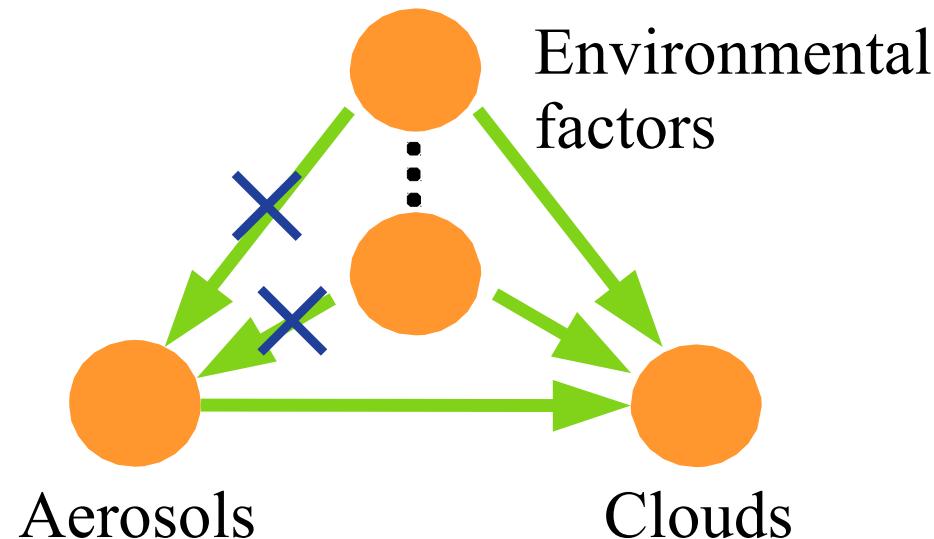


Challenges in Context of Climate Sciences

When trying to apply these concepts in the context of environmental sciences, we face two fundamental challenges.

Challenge: Experimentation in real environment is not possible

For most parts it is impossible, and else arguably ethically questionable, to deliberately intervene on environmental variables.

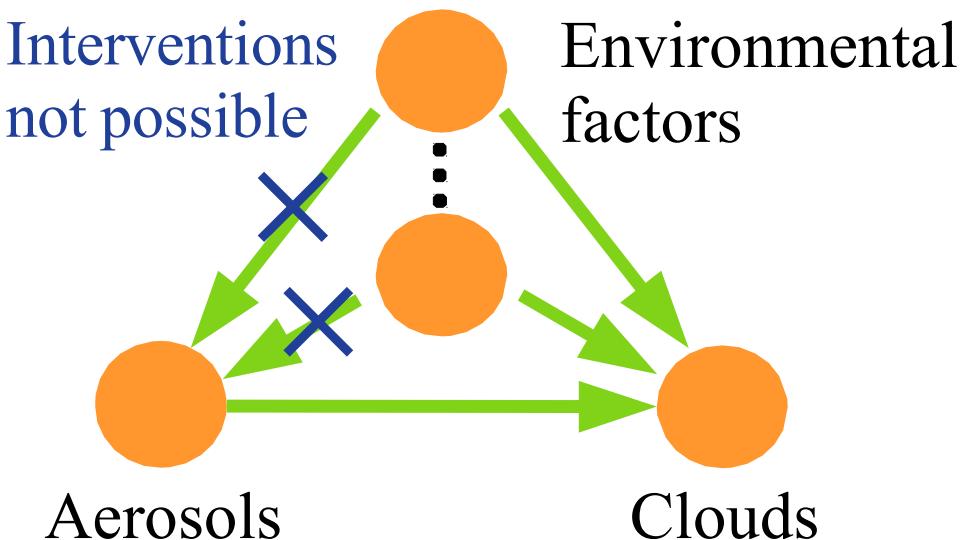


Challenges in Context of Climate Sciences

When trying to apply these concepts in the context of climate sciences, we face two fundamental challenges.

Challenge: Experimentation in real environment is not possible

For most parts it is impossible, and else arguably ethically questionable, to deliberately intervene on climate variables.

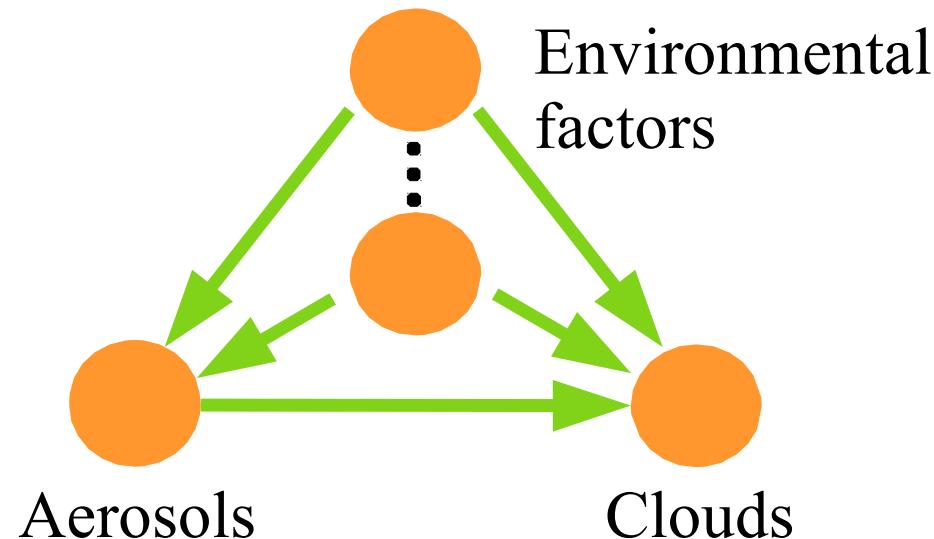


Challenges in Context of Climate Sciences

When trying to apply these concepts in the context of environmental sciences, we face two fundamental challenges.

Challenge: Experimentation in real environment is not possible

For most parts it is impossible, and else arguably ethically questionable, to deliberately intervene on climate variables.



Challenges in Context of Climate Sciences

When trying to apply these concepts in the context of environmental sciences, we face two fundamental challenges.

Challenge: Experimentation in real environment is not possible

For most parts it is impossible, and else arguably ethically questionable, to deliberately intervene on environmental variables.

Challenge: Ground truth not known

Causal relationships are mostly known among microscopic physical variables, but usually less so among macroscopic variables.

Challenges in Context of Climate Sciences

When trying to apply these concepts in the context of environmental sciences, we face two fundamental challenges.

Challenge: Experimentation in real environment is not possible

For most parts it is impossible, and else arguably ethically questionable, to deliberately intervene on environmental variables.

Challenge: Ground truth not known

Causal relationships are mostly known among microscopic physical variables, but usually less so among macroscopic variables.

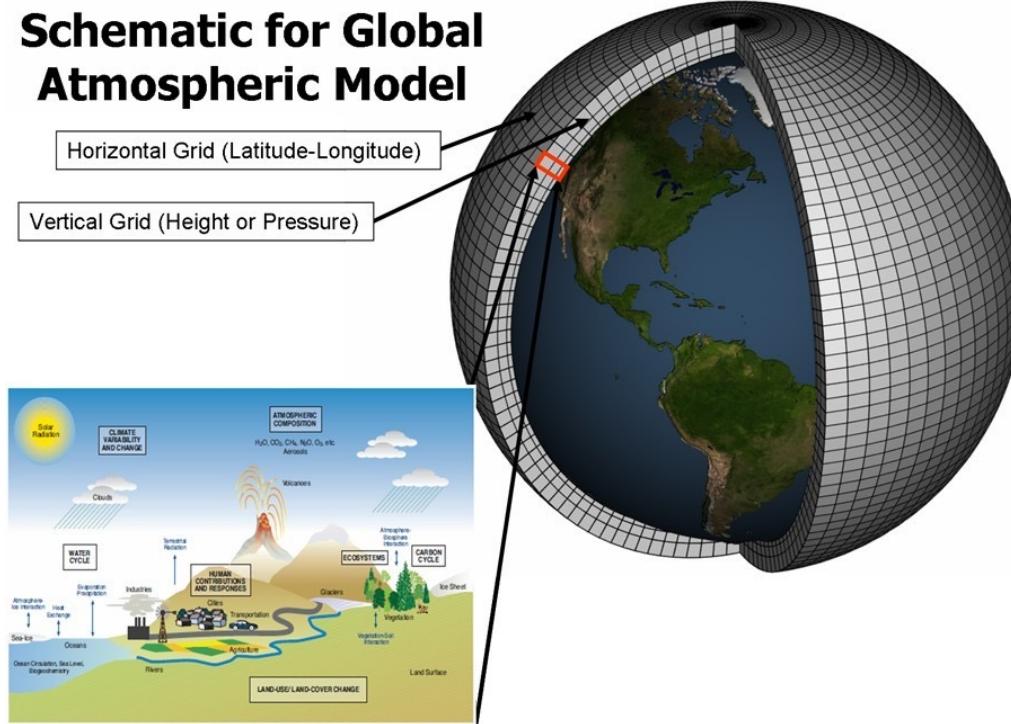
For example:

- Is there a causal relationship between arctic sea ice extend in winter and mid latitude weather in summer?
- Is there a causal relationship between ENSO and a certain extreme weather event somewhere on Earth?

Two Approaches to Address These Challenges

Simulation:

Experimentation in climate models: Fix (intervene on) certain parameters and analyze the resulting effect on other parameters.



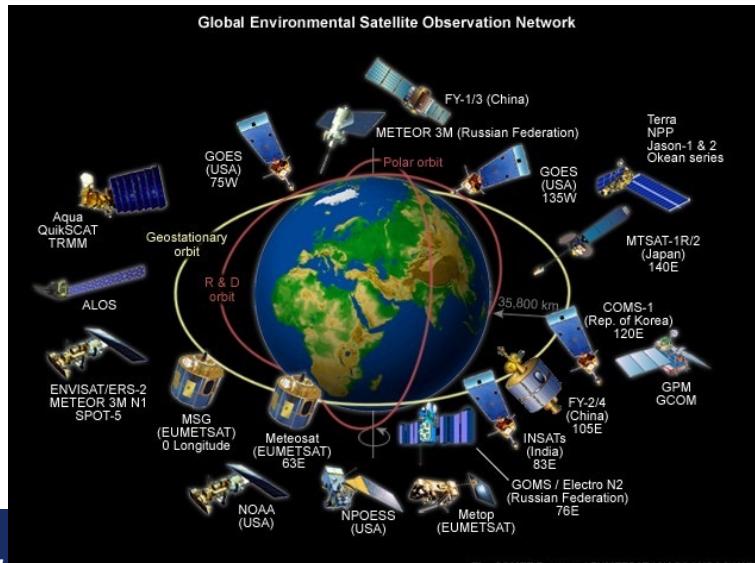
Two Approaches to Address These Challenges

Simulation:

Experimentation in climate models: Fix (intervene on) certain parameters and analyze the resulting effect on other parameters.

Causal inference:

Answer causal questions, for example questions about cause and effect, from observation data.



Two Approaches to Address These Challenges

Simulation:

Experimentation in climate models: Fix (intervene on) certain variables and analyze the resulting effect on other variables.

Causal inference:

Answer causal questions, for example questions about cause and effect, from observation data.

Seminal work by Judea Pearl, Peter Spirtes, Clark Glymour, Richard Scheines.



Two Approaches to Address These Challenges

Simulation:

Experimentation in climate models: Fix (intervene on) certain parameters and analyze the resulting effect on other parameters.

Causal inference:

Answer causal questions, for example questions about cause and effect, from observation data.

To mathematically treat causal inference, we need to formalize the idea of an underlying structural causal model (SCM), define hypothetical interventions, and establish identifiability criteria to decide if and how causal effects can be estimated from data alone (without interventions).

Structural Causal Models

Intuition:

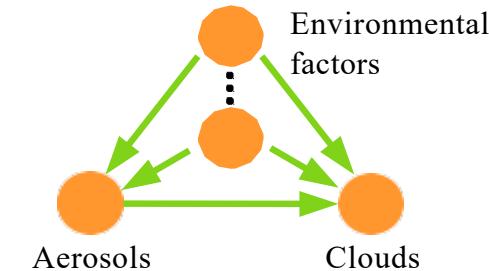
A structural causal model (SCM) specifies the functional causal relationships between a set of random variables.

Example:

$$X_{\text{aerosols}} = f_{\text{aerosols}}(X_{\text{env. facs.}}, \eta_{\text{aerosols}})$$

$$X_{\text{clouds.}} = f_{\text{clouds.}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$



Structural Causal Models

Intuition:

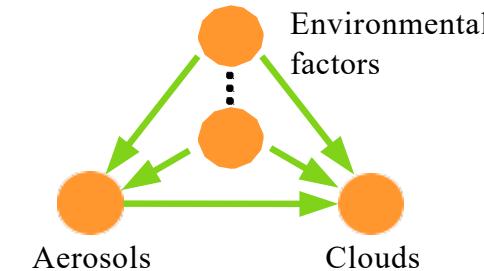
A structural causal model (SCM) specifies the functional causal relationships between a set of random variables.

Example:

$$X_{\text{aerosols}} = f_{\text{aerosols}}(X_{\text{env. facs.}}, \eta_{\text{aerosols}})$$

$$X_{\text{clouds.}} = f_{\text{clouds.}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$



Claims of an SCM:

- Specifies the *direct causes*, also called *parents*, of each variable
- The functions f_i are *independent mechanisms* by which nature determines the values of each variable based on the values of its direct causes plus random fluctuations

Structural Causal Models

Definition:

A structural causal model over the set of random variables

$\mathbf{X} = \{X_1, \dots, X_n\}$ consists of n structural assignments

$$X_1 := f_1(PA_1, \eta_1)$$

⋮

$$X_n := f_n(PA_n, \eta_n)$$

together with a specification of the probability distributions of the jointly independent ‘noise’ variables η_i . Here

$$PA_i \subset \{X_1, \dots, X_{i-1}, X_{i+1}, X_n\} = \mathbf{X} \setminus \{X_i\}$$

are the direct causes of X_i .

Structural Causal Models

Definition:

A structural causal model over the set of random variables

$\mathbf{X} = \{X_1, \dots, X_n\}$ consists of n structural assignments

$$X_1 := f_1(PA_1, \eta_1)$$

⋮

$$X_n := f_n(PA_n, \eta_n)$$

together with a specification of the probability distributions of the jointly independent ‘noise’ variables η_i . Here

$$PA_i \subset \{X_1, \dots, X_{i-1}, X_{i+1}, X_n\} = \mathbf{X} \setminus \{X_i\}$$

are the direct causes of X_i .

Interpretation of noise variables:

- Summarize all background factors outside causal model, i.e., all factors apart from X_1, \dots, X_n
- Their joint independence means the model is sufficient to describe the causal relationship among the variables X_1, \dots, X_n

Causal Graphs

Definition:

Consider an SCM over the variables $\mathbf{x} = \{x_1, \dots, x_n\}$. Its causal graph G is the directed graph with

1. a vertex (node) for each variable x_i
2. for all i a directed edge from each variable $x_j \in PA_i$ to x_i

Causal Graphs

Definition:

Consider an SCM over the variables $\mathbf{x} = \{x_1, \dots, x_n\}$. Its causal graph G is the directed graph with

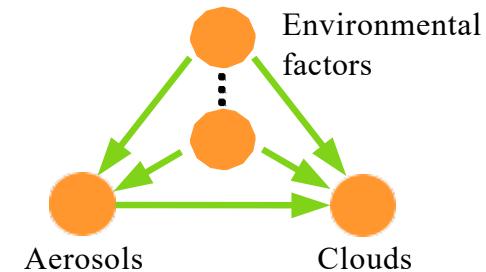
1. a vertex (node) for each variable x_i
2. for all i a directed edge from each variable $x_j \in PA_i$ to x_i

Example:

$$X_{\text{aerosols}} = f_{\text{aerosols}}(X_{\text{env. facs.}}, \eta_{\text{aerosols}})$$

$$X_{\text{clouds.}} = f_{\text{clouds.}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$



Causal Graphs

Definition:

Consider an SCM over the variables $\mathbf{x} = \{x_1, \dots, x_n\}$. Its causal graph G is the directed graph with

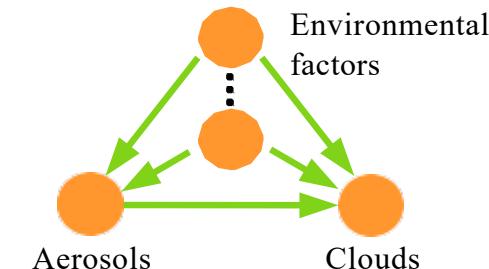
1. a vertex (node) for each variable x_i
2. for all i a directed edge from each variable $x_j \in PA_i$ to x_i

Example:

$$X_{\text{aerosols}} = f_{\text{aerosols}}(X_{\text{env. facs.}}, \eta_{\text{aerosols}})$$

$$X_{\text{clouds}} = f_{\text{clouds}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$



Interpretation:

- Nodes represent variables
- Edges represent causal influences

Interventions

Definition:

Consider SCM over the variables $\mathbf{X} = \{X_1, \dots, X_n\}$.

An *intervention on X_i* , denoted as $do(X_i = x_{i,0})$, defines a modified SCM obtained by replacing the old assignment of X_i with $X_i = x_{i,0}$ where $x_{i,0}$ is some value in the range of X_i :

$$X_1 = f_1(PA_1, \eta_1)$$

⋮

$$X_i = x_{i,0}$$

⋮

$$X_n = f_n(PA_n, \eta_n)$$

Interventions

Definition:

Consider SCM over the variables $\mathbf{X} = \{X_1, \dots, X_n\}$.

An *intervention on X_i* , denoted as $do(X_i = x_{i,0})$, defines a modified SCM obtained by replacing the old assignment of X_i with $X_i = x_{i,0}$ where $x_{i,0}$ is some value in the range of X_i :

$$X_1 = f_1(PA_1, \eta_1)$$

:

$$X_i = x_{i,0}$$

:

$$X_n = f_n(PA_n, \eta_n)$$

Interpretation:

- An intervention $do(X_i := x_{i,0})$ inactivates the natural mechanism $X_i = f_i(PA_i, \eta_i)$ by forcing X_i to take the value $x_{i,0}$
- All other mechanisms remain unmodified
- Interventions can also be *more complex functions* than just $x_{i,0}$

Interventions

Definition:

Consider SCM over the variables $\mathbf{X} = \{X_1, \dots, X_n\}$.

An *intervention on X_i* , denoted as $do(X_i = x_{i,0})$, defines a modified SCM obtained by replacing the old assignment of X_i with $X_i = x_{i,0}$ where $x_{i,0}$ is some value in the range of X_i :

$$X_1 = f_1(PA_1, \eta_1)$$

⋮

$$X_i = x_{i,0}$$

⋮

$$X_n = f_n(PA_n, \eta_n)$$

Interventions model the notion of ‘experimentally manipulate X_i while keeping all other conditions exactly the same’

Effect on causal graphs:

The causal graph of the intervened SCM is obtained from the causal graph of the original SCM by deleting all arrows pointing into X_i .

Interventions

Effect on causal graphs:

The causal graph of the intervened SCM is obtained from the causal graph of the original SCM by deleting all arrows pointing into X_i .

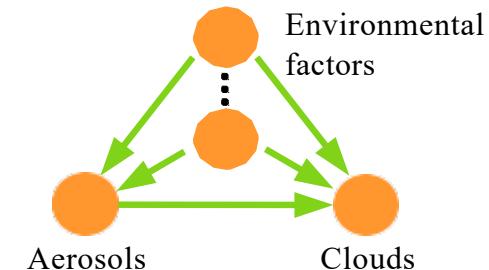
Example:

Original SCM prior to intervention:

$$X_{\text{aerosols}} = f_{\text{aerosols}}(X_{\text{env. facs.}}, \eta_{\text{aerosols}})$$

$$X_{\text{clouds}} = f_{\text{clouds}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$



Interventions

Effect on causal graphs:

The causal graph of the intervened SCM is obtained from the causal graph of the original SCM by deleting all arrows pointing into X_i .

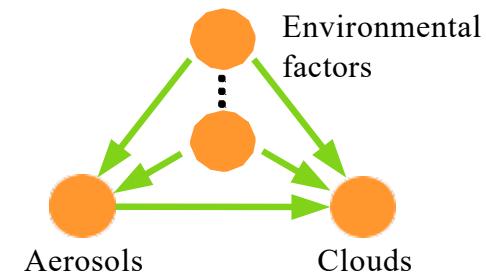
Example:

Original SCM prior to intervention:

$$X_{\text{aerosols}} = f_{\text{aerosols}}(X_{\text{env. facs.}}, \eta_{\text{aerosols}})$$

$$X_{\text{clouds.}} = f_{\text{clouds.}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$

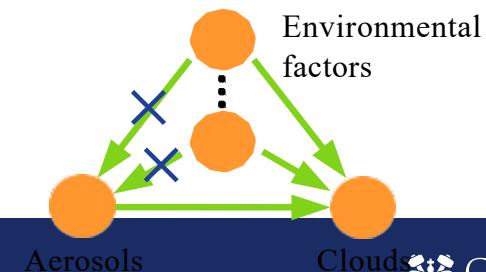


After intervention $do(X_{\text{aerosols}}) := \text{a certain aerosol concentration}$:

$$X_{\text{aerosols}} = \text{a certain aerosol concentration}$$

$$X_{\text{clouds}} = f_{\text{clouds}}(X_{\text{aerosols}}, X_{\text{env. facs.}}, \eta_{\text{clouds}})$$

$$X_{\text{env. facs.}} = f_{\text{env. facs.}}(\eta_{\text{env. facs.}})$$



Identification of Causal Effects

General formalism:

- Formally define the notion of *causal effect*.
- Based on the causal graph, decide whether the causal effect of X on Y can be identified from the observational data.
- If the effect can be identified, express causal effect in terms of observational data

Identification of Causal Effects

Example 1: No confounding

Consider the linear SCM:

$$Z = \eta_Z,$$

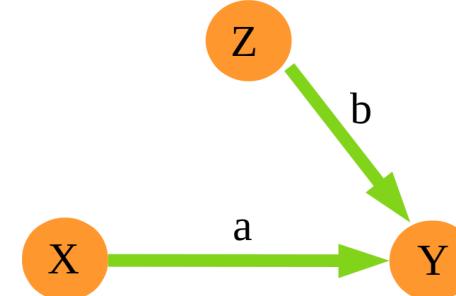
$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$X = \eta_X,$$

$$\eta_X \sim N(0, \sigma_X^2)$$

$$Y = a \cdot X + b \cdot Z + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Identification of Causal Effects

Example 1: No confounding

Consider the linear SCM:

$$Z = \eta_Z,$$

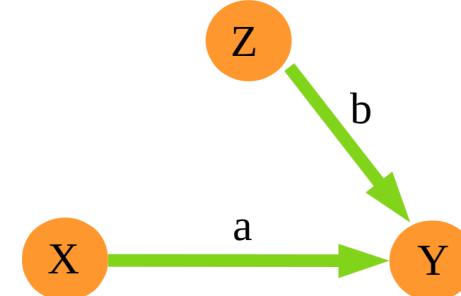
$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$X = \eta_X,$$

$$\eta_X \sim N(0, \sigma_X^2)$$

$$Y = a \cdot X + b \cdot Z + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Goal: Determine a , the causal effect of X on Y

Identification of Causal Effects

Example 1: No confounding

Consider the linear SCM:

$$Z = \eta_Z,$$

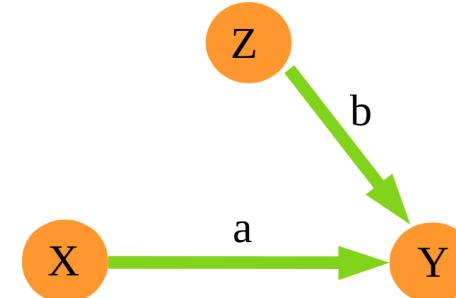
$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$X = \eta_X,$$

$$\eta_X \sim N(0, \sigma_X^2)$$

$$Y = a \cdot X + b \cdot Z + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Goal: Determine a , the causal effect of X on Y

Idea: Use covariance of Y on X

$$\begin{aligned} \text{Cov}(Y, X) &= \text{Cov}(a \cdot \eta_X + b \cdot \eta_Z + \eta_Y, \eta_X,) \\ &= a \cdot \text{Cov}(\eta_X, \eta_X) + b \cdot \text{Cov}(\eta_Z, \eta_X) + \text{Cov}(\eta_Y, \eta_Z) \\ &= a \cdot \text{Var}(X) + b \cdot 0 + 0 \\ \Rightarrow a &= \frac{\text{Cov}(Y, X)}{\text{Var}(X)} = \rho(Y, X) \cdot \sqrt{\frac{\text{Var}(Y)}{\text{Var}(X)}} \end{aligned}$$

Identification of Causal Effects

Example 2: Observed confounding

Consider the linear SCM:

$$Z = \eta_Z,$$

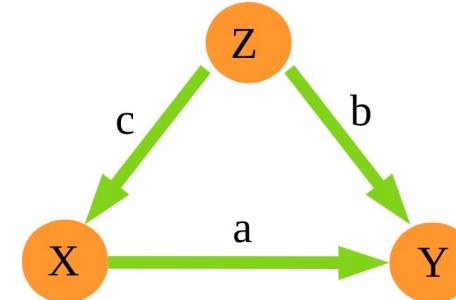
$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$X = c \cdot Z + \eta_X,$$

$$\eta_X \sim N(0, \sigma_X^2)$$

$$Y = a \cdot X + b \cdot Z + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Identification of Causal Effects

Example 2: Observed confounding

Consider the linear SCM:

$$Z = \eta_Z,$$

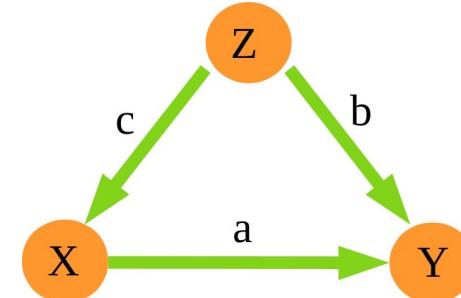
$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$X = c \cdot Z + \eta_X,$$

$$\eta_X \sim N(0, \sigma_X^2)$$

$$Y = a \cdot X + b \cdot Z + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Goal: Determine a , the causal effect of X on Y

Identification of Causal Effects

Example 2: Observed confounding

Consider the linear SCM:

$$Z = \eta_Z,$$

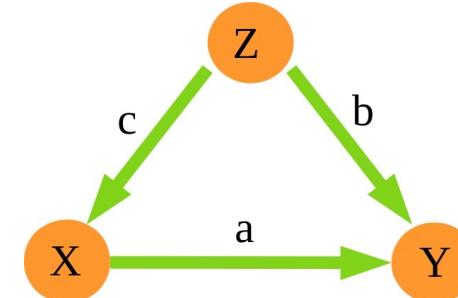
$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$X = c \cdot Z + \eta_X,$$

$$\eta_X \sim N(0, \sigma_X^2)$$

$$Y = a \cdot X + b \cdot Z + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Goal: Determine a , the causal effect of X on Y

First attempt: Linearly regress of Y on X

$$\begin{aligned} \text{Cov}(Y, X) &= \text{Cov}((a \cdot c + b) \cdot \eta_Z + a \cdot \eta_X + \eta_Y, c \cdot \eta_Z + \eta_X) \\ &= c \cdot (a \cdot c + b) \cdot \text{Cov}(\eta_Z, \eta_Z) + a \cdot \text{Cov}(\eta_X, \eta_X) \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= \text{Cov}(c \cdot \eta_Z + \eta_X, c \cdot \eta_Z + \eta_X) \\ &= c^2 \cdot \text{Cov}(\eta_Z, \eta_Z) + \text{Cov}(\eta_X, \eta_X) \end{aligned}$$

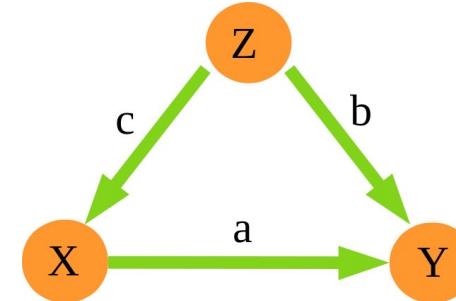
$$\Rightarrow a \neq \frac{\text{Cov}(Y, X)}{\text{Var}(X)}$$

Identification of Causal Effects

Example 2: Observed confounding

Second attempt:

- Control for Z : Regress out influence of Z on X and Y to obtain residuals ΔX and ΔY
- Linearly regress ΔY on ΔX

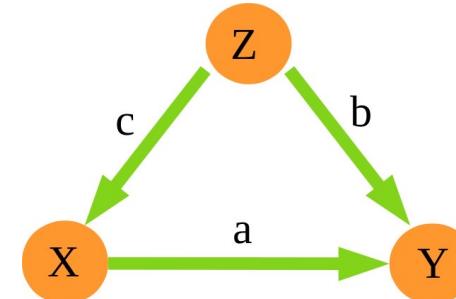


Identification of Causal Effects

Example 2: Observed confounding

Second attempt:

- Control for Z : Regress out influence of Z on X and Y to obtain residuals ΔX and ΔY
- Linearly regress ΔY on ΔX



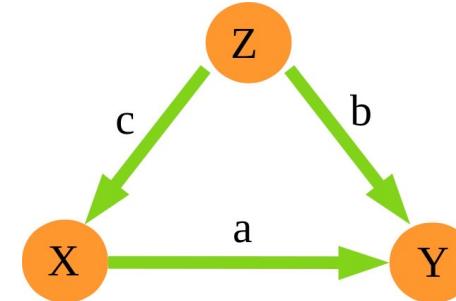
$$\begin{aligned}\Delta X &= X - \frac{\text{Cov}(X, Z)}{\text{Var}(Z)} \cdot Z \\ &= X - \frac{c \cdot \text{Cov}(\eta_Z, \eta_Z)}{\text{Cov}(\eta_Z, \eta_Z)} \cdot Z \\ &= X - c \cdot Z \\ &= \eta_X\end{aligned}$$

Identification of Causal Effects

Example 2: Observed confounding

Second attempt:

- Control for Z : Regress out influence of Z on X and Y to obtain residuals ΔX and ΔY
- Linearly regress ΔY on ΔX



$$\begin{aligned}\Delta X &= X - \frac{\text{Cov}(X, Z)}{\text{Var}(Z)} \cdot Z \\&= X - \frac{c \cdot \text{Cov}(\eta_Z, \eta_Z)}{\text{Cov}(\eta_Z, \eta_Z)} \cdot Z \\&= X - c \cdot Z \\&= \eta_X\end{aligned}$$

$$\begin{aligned}\Delta Y &= Y - \frac{\text{Cov}(Y, Z)}{\text{Var}(Z)} \cdot Z \\&= Y - \frac{(a \cdot c + b) \text{Cov}(\eta_Z, \eta_Z)}{\text{Cov}(\eta_Z, \eta_Z)} \cdot Z \\&= X - (a \cdot c + b) \cdot Z \\&= a \cdot \eta_X + \eta_Y\end{aligned}$$

Identification of Causal Effects

Example 3: Partial mediation

Consider the linear SCM:

$$X = \eta_X,$$

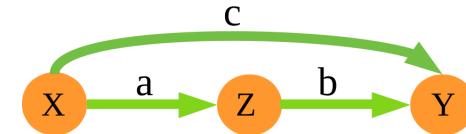
$$\eta_X \sim N(0, \sigma_X^2)$$

$$Z = a \cdot X + \eta_Z,$$

$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$Y = b \cdot Z + c \cdot X + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Identification of Causal Effects

Example 3: Partial mediation

Consider the linear SCM:

$$X = \eta_X,$$

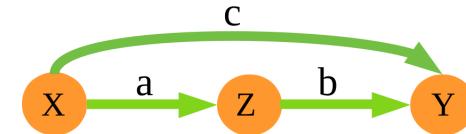
$$\eta_X \sim N(0, \sigma_X^2)$$

$$Z = a \cdot X + \eta_Z,$$

$$\eta_Z \sim N(0, \sigma_Z^2)$$

$$Y = b \cdot Z + c \cdot X + \eta_Y,$$

$$\eta_Y \sim N(0, \sigma_Y^2)$$



Total causal effect of X on Y : $(a \cdot b + c)$

- Linearly regress Y on X

Direct causal effect of X on Y : c

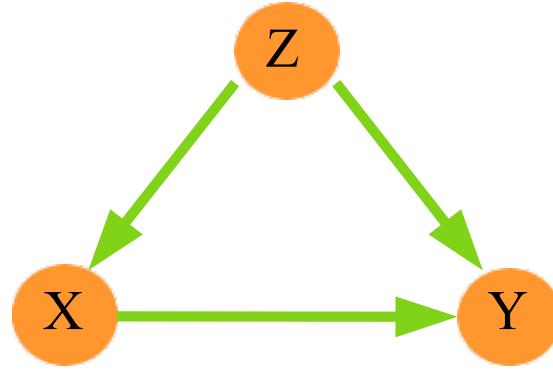
- Control for Z by regressing out influence of Z on X and Y to obtain residuals ΔX and ΔY
- Linearly regress ΔY on ΔX
- Same as *multivariate regression* of Y on its parents (X, Z)

Conceptual Findings

- Interventions model the notion of ‘experimentally manipulate X_i while keeping all other conditions exactly the same’.
- The causal graph allows us to decide whether a certain causal effect is identifiable.
- If the causal effect is identifiable, the causal graph further allows us to estimate it from observational data.

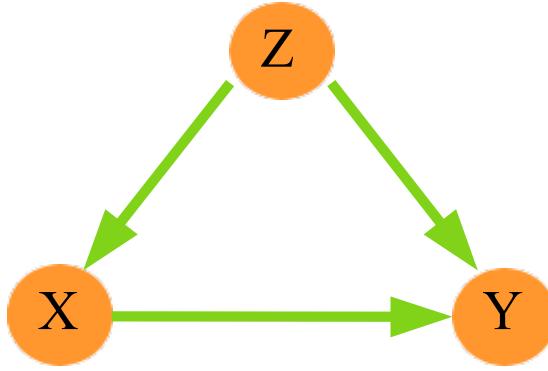
Overview: Identification of Causal Effect

Example: Observed confounder



Overview: Identification of Causal Effect

Example: Observed confounder



Control for confounder Z

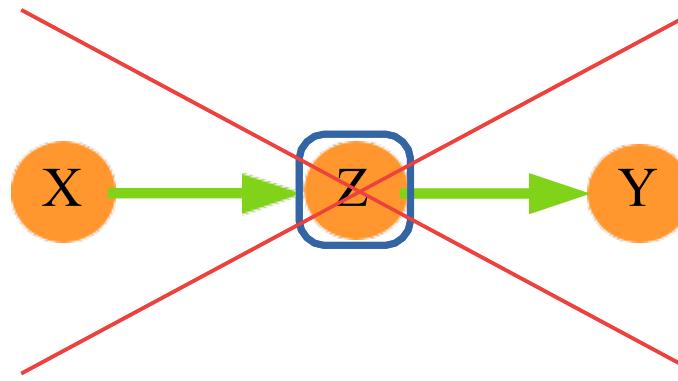
Overview: Identification of Causal Effect

Example: Mediation



Overview: Identification of Causal Effect

Example: Mediation



Overview: Identification of Causal Effect

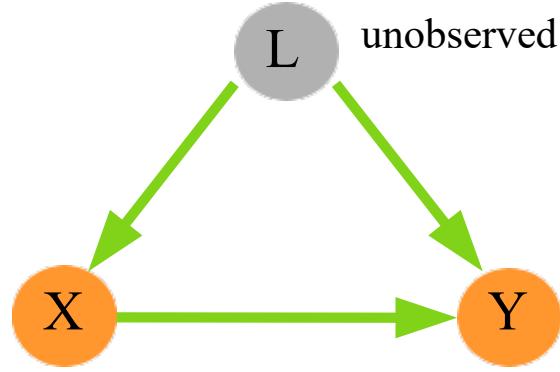
Example: Mediation



Do not control for mediator Z

Overview: Identification of Causal Effect

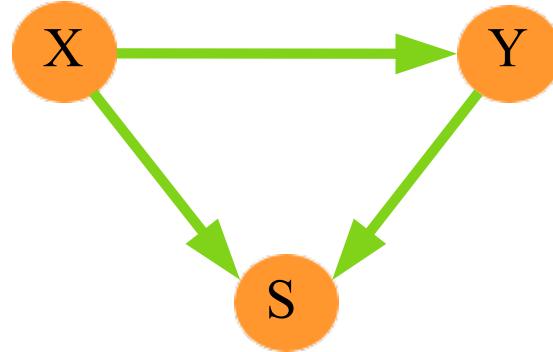
Example: Unobserved confounder



Causal effect cannot be identified

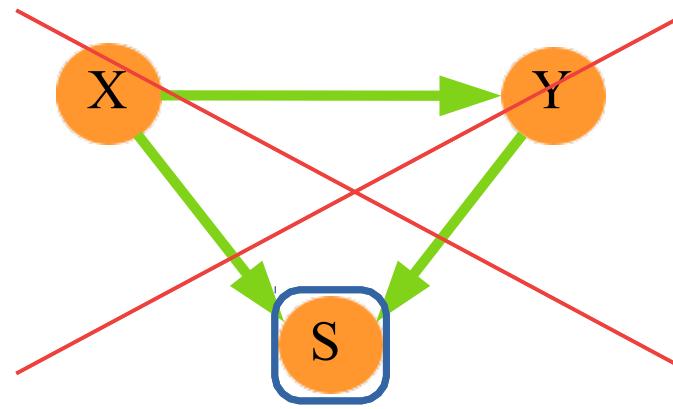
Overview: Identification of Causal Effect

Example: Common effect



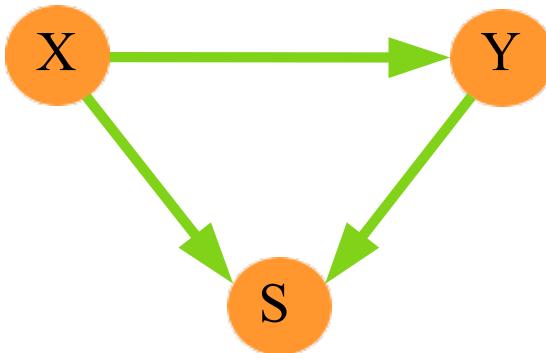
Overview: Identification of Causal Effect

Example: Common effect



Overview: Identification of Causal Effect

Example: Common effect

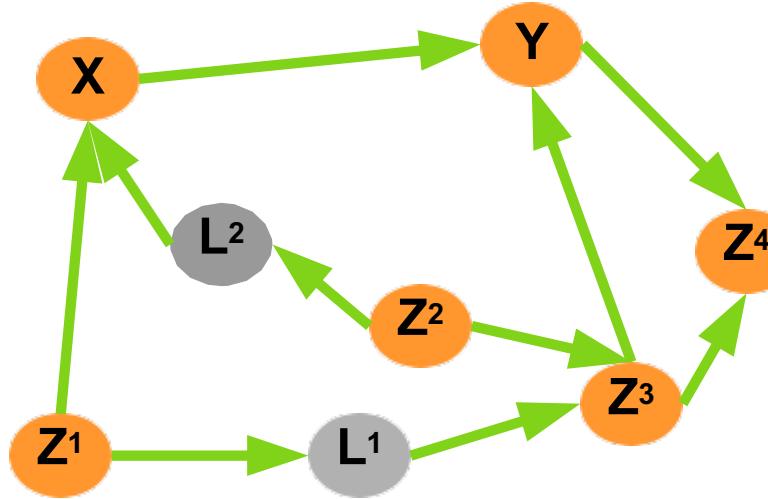


Do not control for common effects of X and Y , this would introduce bias.

However, sometimes the available dataset is already biased, e.g., if samples are missing (selection bias).

Overview: Identification of Causal Effect

Example: Complicated case

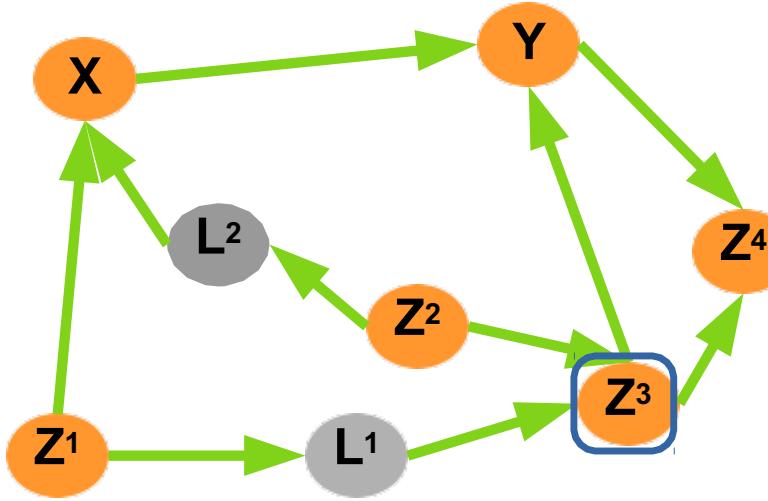


Back-door criterion:

Block all non-causal dependence by controlling for an appropriate set of variables

Overview: Identification of Causal Effect

Example: Complicated case

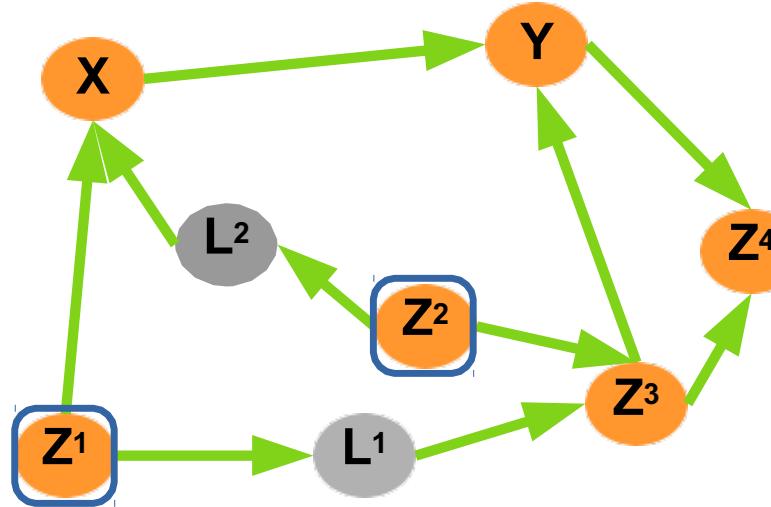


Back-door criterion:

Block all non-causal dependence by controlling for an appropriate set of variables

Overview: Identification of Causal Effect

Example: Complicated case

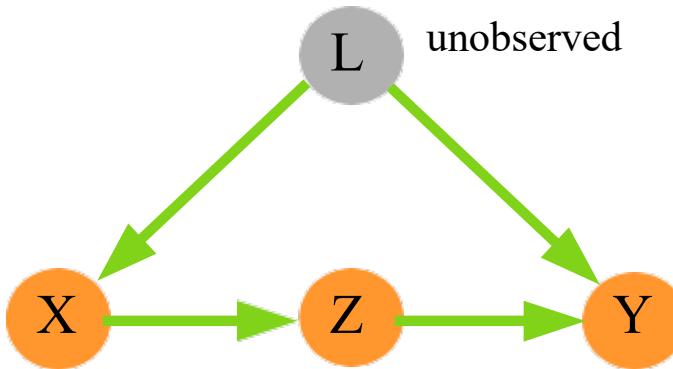


Back-door criterion:

Block all non-causal dependence by controlling for an appropriate set of variables

Overview: Identification of Causal Effect

Example: Unobserved confounder plus mediation



Front-door criterion:

Allows to identify causal effect, but not as simple as controlling for a certain set of variables

Observational Causal Discovery

How to Get Causal Graphs

Relevance: Result of previous section

Knowing the causal graph of the data generating process, we can determine whether and how a causal effect can be identified from observational data

How to Get Causal Graphs

Relevance: Result of previous section

Knowing the causal graph of the data generating process, we can determine whether and how a causal effect can be identified from observational data

How to get the causal graph?

How to Get Causal Graphs

Relevance: Result of previous section

Knowing the causal graph of the data generating process, we can determine whether and how a causal effect can be identified from observational data

How to get the causal graph?

Option 1: Scientific knowledge

Talk to domain experts, general reasoning

Option 2: Observational causal discovery

Learn from observational data

Causal Graphs and (Conditional) Independencies

Fact:

The structure of the causal graph often has observable implications in terms of (conditional) independencies in the observed data.

Intuition:

- Statistical dependencies derive from causal relationships
- Conditioning can block and open the ‘flow of information’

Causal Graphs and (Conditional) Independencies

Fact:

The structure of the causal graph often has observable implications in terms of (conditional) independencies in the observed data.

Intuition:

- Statistical dependencies derive from causal relationships
- Conditioning can block and open the ‘flow of information’

Example 1: Chain



- X_1 influences X_2 : $X_1 \not\perp\!\!\!\perp X_2$
- X_2 influences X_3 : $X_2 \not\perp\!\!\!\perp X_3$
- X_1 influences X_3 through X_2 : $X_1 \not\perp\!\!\!\perp X_3$
- Knowing X_2 , X_1 does not say more about X_3 : $X_1 \perp\!\!\!\perp X_3 | X_2$

Causal Graphs and (Conditional) Independencies

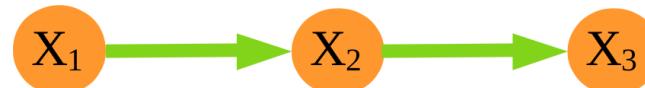
Fact:

The structure of the causal graph often has observable implications in terms of (conditional) independencies in the observed data.

Intuition:

- Statistical dependencies derive from causal relationships
- Conditioning can block and open the ‘flow of information’

Example 1: Chain



• X_1 influences X_2 :

$$X_1 \not\perp\!\!\!\perp X_2$$

• X_2 influences X_3 :

$$X_2 \not\perp\!\!\!\perp X_3$$

• X_1 influences X_3 through X_2 :

$$X_1 \not\perp\!\!\!\perp X_3$$

• Knowing X_2 , X_1 does not say more about X_3 :

$$X_1 \perp\!\!\!\perp X_3 | X_2$$

Causal Graphs and (Conditional) Independencies

Example 2: Fork



- X_2 influences X_1 :
- X_2 influences X_3 :
- Observing X_1 says something about X_2 and hence about X_3 :
- Knowing X_2 , X_1 does not say more about X_3 :

$$X_1 \perp\!\!\!\perp X_2$$

$$X_2 \perp\!\!\!\perp X_3$$

$$X_1 \perp\!\!\!\perp X_3$$

$$X_1 \perp\!\!\!\perp X_3 | X_2$$

Causal Graphs and (Conditional) Independencies

Example 2: Fork



- X_2 influences X_1 :
- X_2 influences X_3 :
- Observing X_1 says something about X_2 and hence about X_3 :
- Knowing X_2 , X_1 does not say more about X_3 :

$$X_1 \perp\!\!\!\perp X_2$$

$$X_2 \perp\!\!\!\perp X_3$$

$$X_1 \perp\!\!\!\perp X_3$$

$$X_1 \perp\!\!\!\perp X_3 | X_2$$

Causal Graphs and (Conditional) Independencies

Example 3: Collider



- X_1 influences X_2 : $X_1 \not\perp\!\!\!\perp X_2$
- X_3 influences X_2 : $X_2 \not\perp\!\!\!\perp X_3$
- No influence between X_1 and X_3 : $X_1 \perp\!\!\!\perp X_3$
- Observing X_2 introduces selection bias between X_1 and X_3 : $X_1 \not\perp\!\!\!\perp X_3 | X_2$

Causal Graphs and (Conditional) Independencies

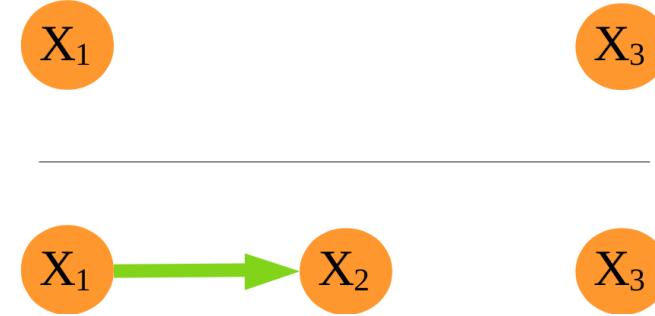
Example 3: Collider



- X_1 influences X_2 : $X_1 \not\perp\!\!\!\perp X_2$
- X_3 influences X_2 : $X_2 \not\perp\!\!\!\perp X_3$
- No influence between X_1 and X_3 : $X_1 \perp\!\!\!\perp X_3$
- Observing X_2 introduces selection bias between X_1 and X_3 : $X_1 \not\perp\!\!\!\perp X_3 | X_2$

Causal Graphs and (Conditional) Independencies

Example 4: Disconnected variables



- No influence between X_1 and X_3 : $X_1 \perp\!\!\!\perp X_3$

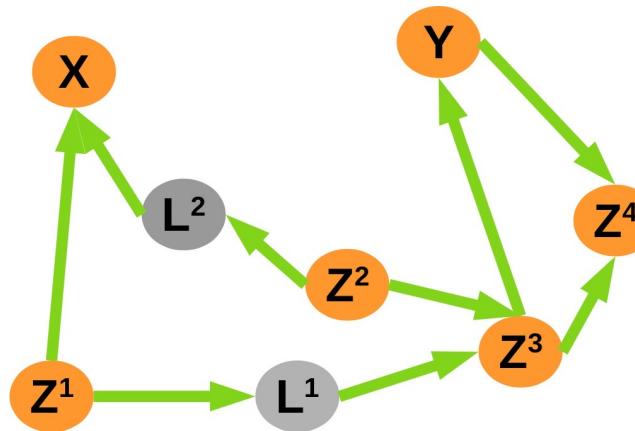
Causal Graphs and (Conditional) Independencies

General rule: d-separation

The graphical criterion of *d-separation* allows to read off all (conditional) independencies implied by the structure of a particular causal graph.

d-separation: Whether a set X of variables is independent of another set Y, given a third set Z. The idea is to associate "dependence" with "connectedness" (i.e., the existence of a connecting path) and "independence" with "unconnected-ness" or "separation".

Example 5: Complicated case



Some independencies implied by d-separation

$$X \perp\!\!\!\perp Y | Z^1, Z^2$$

$$X \perp\!\!\!\perp Y | Z^3$$

Constraint Based Causal Discovery

Idea:

- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Constraint Based Causal Discovery

Idea:

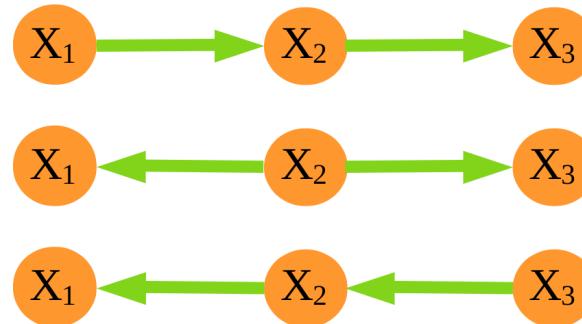
- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Example 1:

Test decisions:

- $X_1 \perp\!\!\!\perp X_2$
- $X_2 \perp\!\!\!\perp X_3$
- $X_1 \perp\!\!\!\perp X_3$
- $X_1 \perp\!\!\!\perp X_3 | X_2$

Possible causal graphs:



Constraint Based Causal Discovery

Idea:

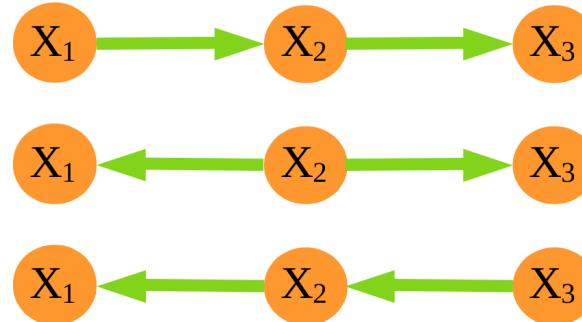
- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Example 1:

Test decisions:

- $X_1 \perp\!\!\!\perp X_2$
- $X_2 \perp\!\!\!\perp X_3$
- $X_1 \perp\!\!\!\perp X_3$
- $X_1 \perp\!\!\!\perp X_3 | X_2$

Possible causal graphs:



observationally equivalent graphs

Constraint Based Causal Discovery

Idea:

- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Example 2:

Test decisions:

- $X_1 \perp\!\!\!\perp X_2$
- $X_2 \perp\!\!\!\perp X_3$
- $X_1 \perp\!\!\!\perp X_3$

Possible causal graphs:



Constraint Based Causal Discovery

Idea:

- Perform statistical tests of (conditional) independence in observational data
- Use test results to constrain the structure of the causal graph

Example 2:

Test decisions:

- $X_1 \perp\!\!\!\perp X_2$
- $X_2 \perp\!\!\!\perp X_3$
- $X_1 \perp\!\!\!\perp X_3$

Possible causal graphs:



observationally equivalent graphs

Required Assumptions

Recall:

Causal questions cannot be answered from observational data without making any additional assumptions.

Required Assumptions

Recall:

Causal questions cannot be answered from observational data without making any additional assumptions.

Assumption 1: Existence of SCM

The data generating process can be described as an SCM. This implies the so called *causal Markov condition*:

d-separation in causal graph \Rightarrow (conditional) independence

Assumption 2: Faithfulness

(conditional) independence \Rightarrow d-separation in causal graph

Assumption 3: Causal sufficiency

There are no unobserved confounders and there are no selection variables.

Required Assumptions

Recall:

Causal questions cannot be answered from observational data without making any additional assumptions.

Assumption 1: Existence of SCM

The data generating process can be described as an SCM. This implies the so called *causal Markov condition*:

d-separation in causal graph \Rightarrow (conditional) independence

Assumption 2: Faithfulness

(conditional) independence \Rightarrow d-separation in causal graph

Assumption 3: Causal sufficiency

There are no unobserved confounders and there are no selection variables.

Assumption 3 can be dropped (but much more complicated)

Assumption 2 can be relaxed

Some Famous Algorithms

SGS-Algorithm:

Tests all possible (conditional) independence statements, use results to constrain the causal graph as discussed in the examples.

SGS stands for *S*pirites, *G*lymour, and *S*cheines.

PC-Algorithm:

Based on SGS-Algorithm, but using much fewer conditional independence tests.

PC stands for *P*eter *S*pirtes and *C*lark *G*lymour.

FCI-Algorithm:

Similar to the PC-Algorithm, but without requiring the assumption of causal sufficiency, i.e., allowing for hidden confounders and selection variables.

FCI stands for *f*ast *c*ausal *i*nference.

Two Main Approaches of Causal Discovery

Constraint-based causal discovery:

Constrain causal graph by using results of conditional independence tests in observational data.

⇒ discussed

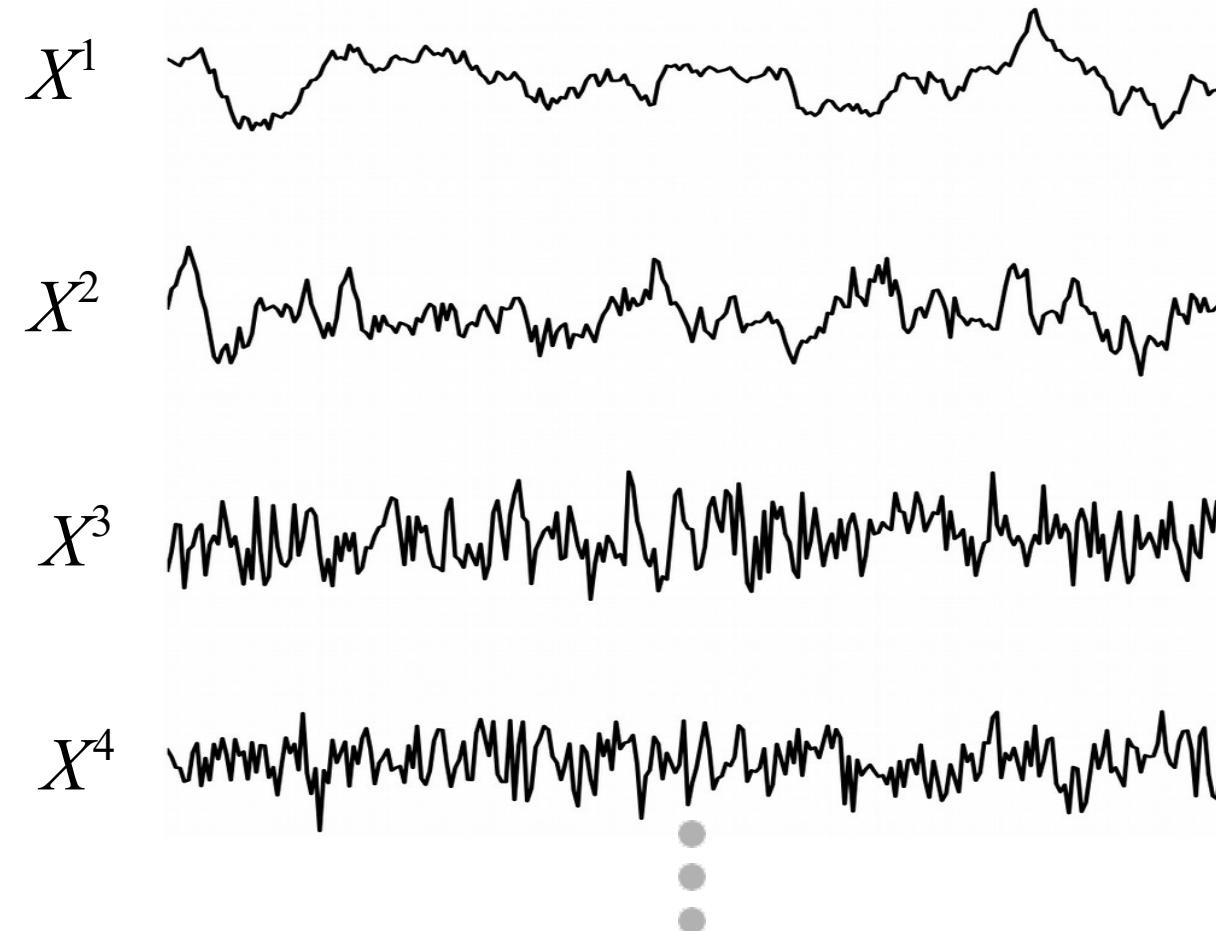
SCM-based causal discovery:

Make assumption on functional causal relationships (e.g., linear or non-linear) and noise distributions (e.g., Gaussian or non-Gaussian) of data generating SCM.

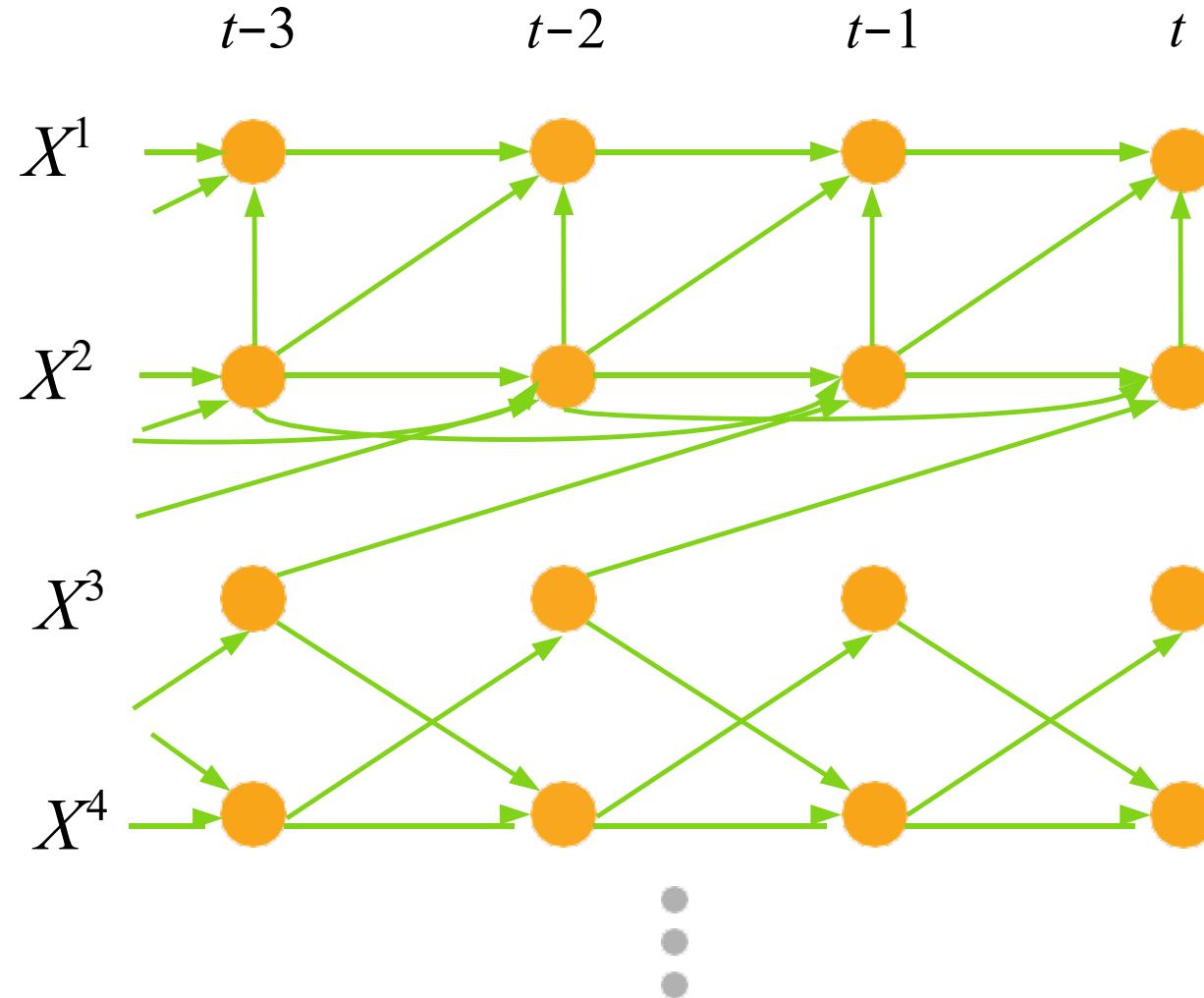
Generically, model can fit in one direction only. This allows to identify direction of causal influence.

⇒ not discussed

Causal Discovery in Time Series



Causal Discovery in Time Series



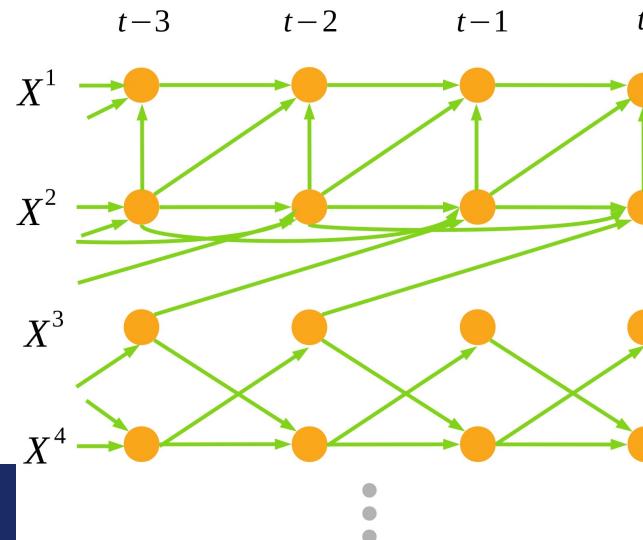
Causal Discovery in Time Series

Observations in time make things easier:

- Additional constraint: Causation cannot go back in time.

Observations in time make things harder:

- High dimensionality: Resolving in time increases the number of variables
- Statistical issues: Autocorrelation makes conditional independence tests statistically harder



Causal Discovery in Time Series

Observations in time make things harder:

- High dimensionality: Resolving in time increases the number of variables
- Statistical issues: Autocorrelation makes conditional independence tests statistically harder

PCMCI-Algorithm:

Adaption of the PC (Peters/Clark)-algorithm to time series in which these challenges are addressed.

Alternative Concept: GrangerCausality

Idea:

X causes Y if the past of X helps in predicting the Y from its own past.

Limitations:

- Requires causal sufficiency
- Does not allow contemporaneous interactions

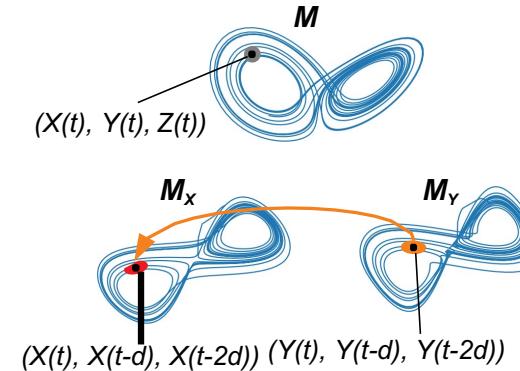
State of the art: see Runge et al. NatComm 2019

a Granger causality

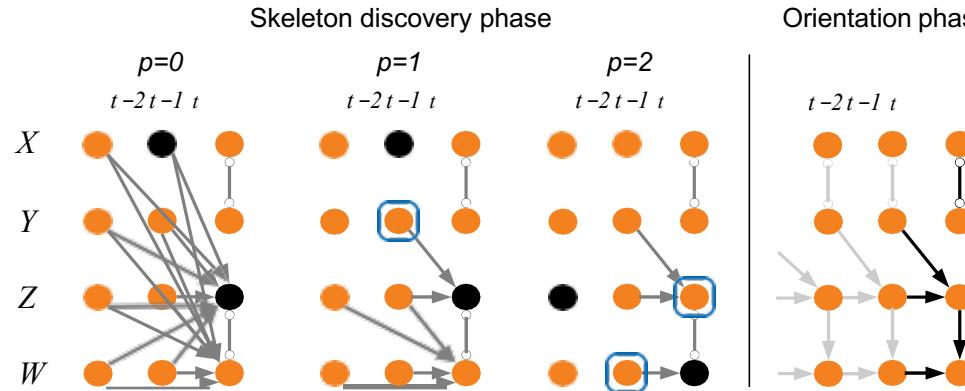
$$Y_t = \sum_{\tau=1}^p \beta_\tau Y_{t-\tau} + \alpha_\tau X_{t-\tau} + E_t^Y \quad (1)$$

$$Y_t = \sum_{\tau=1}^p \beta'_\tau Y_{t-\tau} + E_t'^Y \quad (2)$$

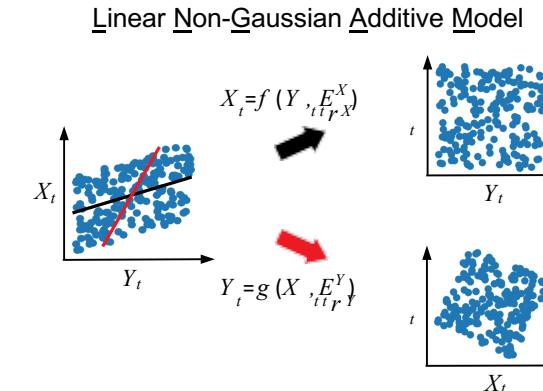
b Nonlinear state-space methods



c Causal network learning algorithms



d Structural causal models



Real world challenges: see Runge et al. NatComm 2019

Challenges

Process:

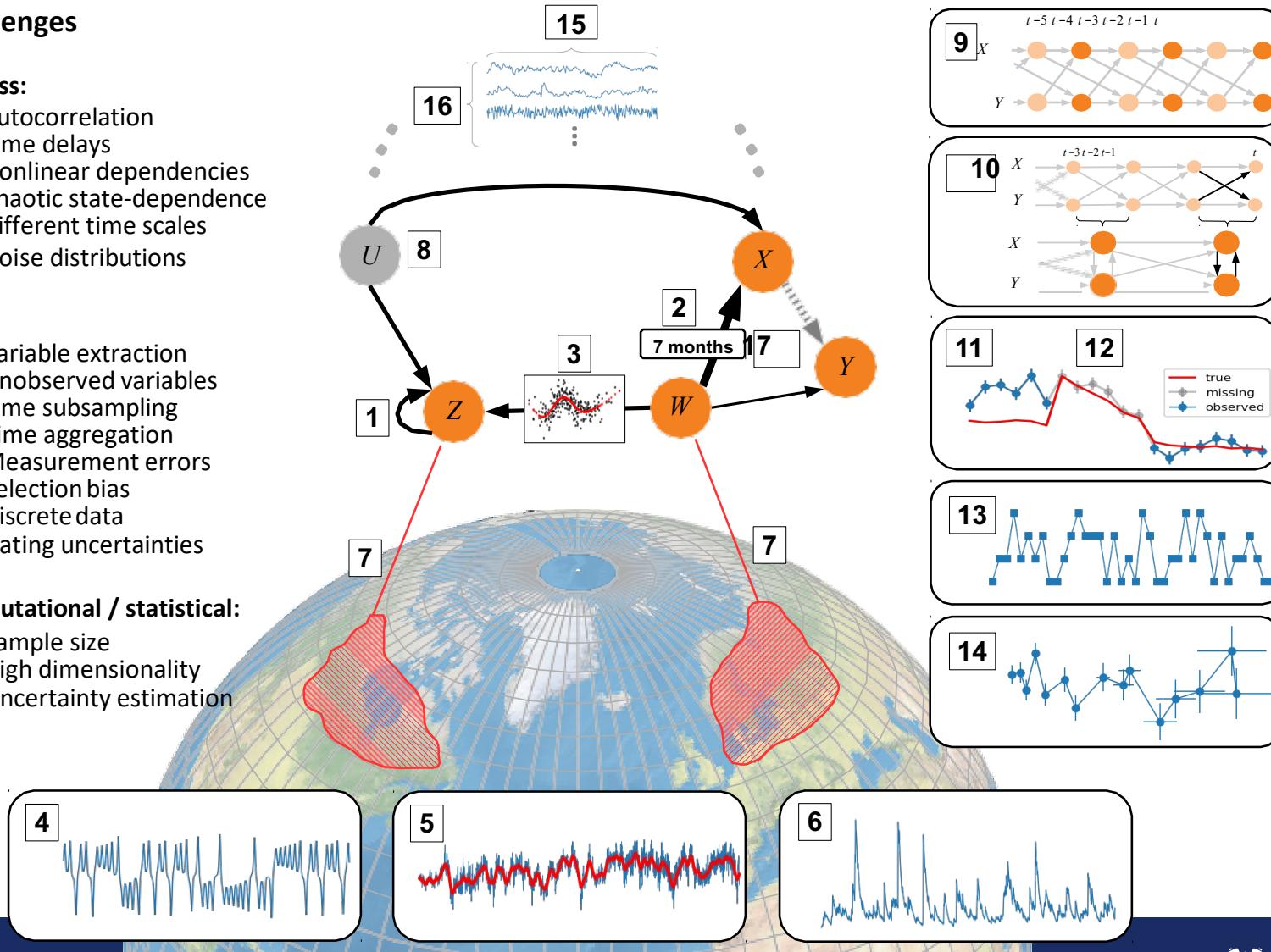
- 1 Autocorrelation
- 2 Time delays
- 3 Nonlinear dependencies
- 4 Chaotic state-dependence
- 5 Different time scales
- 6 Noise distributions

Data:

- 7 Variable extraction
- 8 Unobserved variables
- 9 Time subsampling
- 10 Time aggregation
- 11 Measurement errors
- 12 Selection bias
- 13 Discrete data
- 14 Dating uncertainties

Computational / statistical:

- 15 Sample size
- 16 High dimensionality
- 17 Uncertainty estimation



Summary

- Causal inference:
Answering causal questions from empirical data

Summary

- Causal inference:
Answering causal questions from empirical data
- Two settings:

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)
 - Potentially causal sufficiency

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)
 - Potentially causal sufficiency
 - Assumptions on dependency types (linearity, etc) and distributions

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)
 - Potentially causal sufficiency
 - Assumptions on dependency types (linearity, etc) and distributions
 - Time series data: time order, stationarity,...

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)
 - Potentially causal sufficiency
 - Assumptions on dependency types (linearity, etc) and distributions
 - Time series data: time order, stationarity,...
- These assumptions can sometimes not be **tested** from the same data or even any empirical data

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)
 - Potentially causal sufficiency
 - Assumptions on dependency types (linearity, etc) and distributions
 - Time series data: time order, stationarity,...
- These assumptions can sometimes not be **tested** from the same data or even any empirical data
- Causal conclusions require to

Summary

- **Causal inference:**
Answering causal questions from empirical data
- Two settings:
 1. Causal graph assumed known → **estimate causal effects**
 2. Causal graph unknown → **causal discovery** (→ then causal effects)
- **Causal inference requires assumptions**
 - Underlying SCM / Causal Markov Condition
 - Faithfulness (for causal discovery)
 - Potentially causal sufficiency
 - Assumptions on dependency types (linearity, etc) and distributions
 - Time series data: time order, stationarity,...
- These assumptions can sometimes not be **tested** from the same data or even any empirical data
- Causal conclusions require to
 - **state assumptions** and explain reasons for believing them