Causality II

Marlene Kretschmer

University of Reading

m.j.a.kretschmer@reading.ac.uk
@Marlene_Climate
RECAP

- Scientific data analysis requires causal reasoning
- Causal knowledge/hypotheses are best expressed using causal networks
- To extract causal effects from data, one needs to control for all confounding factors
- Causal inference gives the formal rules how to achieve this
IN THIS LECTURE

1) Conditioning on a common effect
2) Controlling for confounders
3) An example from climate science
4) Non-linear dependencies
5) Conclusions
1) Conditioning on a common effect
 BASIC CAUSAL STRUCTURES AND THEIR IMPLICATIONS 

- **Z is a common driver of X and Y**
  - Implication for data:
    - X and Y correlated
    - X and Y are uncorrelated conditional on Z

- **Z is a mediator of X and Y**
  - Implication for data:
    - X and Y uncorrelated
    - X and Y are correlated conditional on Z

- **Z is a common effect of X and Y**
CREATING INTUITION

Age

COVID-19 infection

App usage

Griffith et al., Nature Communications (2020)
**The Stratospheric Polar Vortex (SPV)**

Data from the seasonal forecasting model SEAS5

THE STRATOSPHERIC POLAR VORTEX

Monnin, Kretschmer, Polichtchouk, Int. J. Clim (2021)
SUDDEN STRATOSPHERIC WARMINGS (SSWs)

SSWs := Days when the winds turn negative

EARLY WINTER SSWs SHOW STRONGER SURFACE IMPACTS
EARLY WINTER SSWs HAVE STRONGER WIND ANOMALIES...

...due to the SSW definition

This explains the stronger impacts

**SSW IS A COMMON EFFECT**

By conditioning on the common effect “SSW”, we introduce a non-causal association between “SPV anomaly” and “month”
CONDITIONING ON A COMMON EFFECT

All winter days

Only SSWs

SPV anomaly (m/s)

SPV anomaly (m/s)

NAO anomaly (hPa)

2) Controlling for confounders (in a Nutshell)
CAUSAL NETWORKS

- A causal network consists of **nodes** (representing variables, e.g. ENSO) and **links** (indicating the direction of causality)
- causal network = directed acyclic graph (DAG)
- a sequence of links “connecting” two nodes in the network is called a **path** (regardless of the direction of arrows!)

Paths from X to Y:

X --> U --> Y
X <-- Z --> Y
**Basic Causal Structures and Their Implications**

- **Z is a common driver of X and Y**
  - The path from X to Y is open
  - X and Y are dependent

- **Z is a mediator of X and Y**
  - The path from X to Y is blocked by conditioning on Z
  - X and Y are independent conditional on Z

- **Z is a common effect of X and Y**
  - The path from X to Y is blocked by Z
  - X and Y are independent
  - The path from X to Y is opened by conditioning on Z
  - X and Y are dependent conditional on Z
- There is an open path DK <--NAO -->MED
- conditioning on NAO blocks this path

- There is an open path ENSO --> Jet --> CA
- conditioning on Jet blocks this path

- The path SPV --> SSW <-- month is blocked
- conditioning on SSW opens this path
### Rules of Do-Calculus

Aim: Express $P(Y|\text{do}(X))$ such that it does not contain any “do” expressions.

1. Insertion/deletion of observations

   $$P(Y|\text{do}(X),Z,W) = P(Y|\text{do}(X),Z)$$

   *If $W$ is irrelevant to $Y$*

2. Action/observation exchange

   $$P(Y|\text{do}(X),Z) = P(Y|X,Z)$$

   *If $Z$ blocks all back-door paths from $X$ to $Y$*

3. Insertion/deletion of actions

   $$P(Y|\text{do}(X)) = P(Y)$$

   *If there is no causal path from $X$ to $Y$*
The Backdoor Criterion

To quantify the causal effect of $X$ on $Y$, one needs to control for all confounding factors. Confounding is anything that leads to $P(Y|X)$ being different than $P(Y|\text{do}(X))$.

To quantify the causal effect of $X$ on $Y$, one needs to block all open paths between them (other than the one of interest).
EXAMPLES (OF GOOD CONTROLS)

To block the path $X \leftarrow Z \rightarrow Y$

To block the path $X \leftarrow Z \leftarrow U \rightarrow Y$

Source: http://causality.cs.ucla.edu/blog/index.php/category/back-door-criterion/
EXAMPLES (OF GOOD CONTROLS)

To block the path $X \leftarrow Z \rightarrow M \rightarrow Y$

To block the path $X \leftarrow U \rightarrow Z \rightarrow M \rightarrow Y$

Source: http://causality.cs.ucla.edu/blog/index.php/category/back-door-criterion/
EXAMPLES (OF BAD CONTROLS)

Because this opens the path
X <--U1 --> Z <--U2 --> Y

Source: http://causality.cs.ucla.edu/blog/index.php/category/back-door-criterion/
EXAMPLES (OF BAD CONTROLS)

Because this blocks the path $X \rightarrow Z \rightarrow Y$

Because this (partially) blocks the path $X \rightarrow M \rightarrow Y$ (as $Z$ is evidence for $M$)

Source: http://causality.cs.ucla.edu/blog/index.php/category/back-door-criterion/
RECOMMENDATIONS

+ many tutorials in the internet!
3) An example from climate science
INFLUENCE OF SEA ICE ON THE POLAR VORTEX

How strong is the causal effect of Barents Kara sea ice (BK) in autumn on the winter stratospheric polar vortex (SPV)?
How strong is the causal effect of Barents Kara sea ice (BK) in autumn on the winter stratospheric polar vortex (SPV)?
CONTROLLING FOR THE BACK DOOR PATHS

Open Paths from BK to SPV:

\[ \text{BK} \rightarrow \text{Ural} \rightarrow \text{SPV} \]
\[ \text{BK} \leftarrow \text{AL} \rightarrow \text{SPV} \]
\[ \text{BK} \leftarrow \text{Ural} \rightarrow \text{SPV} \]

\[
\text{SPV}_{\text{JFM}} = a \text{BK}_{\text{OND}} + \text{confounders}
\]
\[
\text{SPV}_{\text{JFM}} = a \text{BK}_{\text{OND}} + b \text{AL}_{\text{OND}}
\]
\[
\text{SPV}_{\text{JFM}} = a \text{BK}_{\text{OND}} + b \text{AL}_{\text{OND}} + c \text{URAL}_{\text{OND}} + \epsilon
\]
CONTROLLING FOR THE BACK DOOR PATHS

Causal networks make scientific assumptions transparent and help to identify where information is propagating.

Kretschmer et al. WCD (2020), Kretschmer et al. BAMS (2021)
4) Non-linear dependencies
Precipitation in Australia (AU) is affected by ENSO and by the Indian Ocean Dipole (IOD)

The relationships likely involve non-linearities
**NON-LINEAR DEPENDENCIES**

We stratify the data into different categories:

- **AU**: below/above average
- **IOD**: negative/neutral/positive phase
- **ENSO**: La Niña/neutral/El Niño

Conditional probabilities for above average AU (default = ½)

<table>
<thead>
<tr>
<th></th>
<th>La Niña</th>
<th>Neutral</th>
<th>El Niño</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOD -</td>
<td>0.83</td>
<td>0.50</td>
<td>-</td>
<td>0.67</td>
</tr>
<tr>
<td>Neutral</td>
<td>0.80</td>
<td>0.43</td>
<td>0.17</td>
<td>0.52</td>
</tr>
<tr>
<td>IOD +</td>
<td>1.0</td>
<td>0.25</td>
<td>0.24</td>
<td>0.30</td>
</tr>
<tr>
<td>Marginal</td>
<td>0.83</td>
<td>0.43</td>
<td>0.22</td>
<td>0.50</td>
</tr>
</tbody>
</table>
NON-LINEAR DEPENDENCIES

Conditional probabilities for above average AU (default = \( \frac{1}{2} \))

<table>
<thead>
<tr>
<th></th>
<th>La Niña</th>
<th>Neutral</th>
<th>El Niño</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOD -</td>
<td>0.83</td>
<td>0.50</td>
<td>-</td>
<td>0.67</td>
</tr>
<tr>
<td>Neutral</td>
<td>0.80</td>
<td>0.43</td>
<td>0.17</td>
<td>0.52</td>
</tr>
<tr>
<td>IOD +</td>
<td>1.0</td>
<td>0.25</td>
<td>0.24</td>
<td>0.30</td>
</tr>
<tr>
<td>Marginal</td>
<td>0.83</td>
<td>0.43</td>
<td>0.22</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Above average precipitation is unlikely during El Niño

\[
P(AU^+ \mid \text{El Niño}) = 0.22
\]

Above average precipitation is likely during IOD-

\[
P(AU^+ \mid \text{IOD-}) = 0.67
\]

Kretschmer et al., BAMS (2021)
NON-LINEAR DEPENDENCIES

What is the added information provided by IOD, given ENSO?

\[ P(\text{AU}+ | \text{IOD}+, \text{El Niño}) = 0.24 \]
\[ P(\text{AU}+ | \text{El Niño}) = 0.22 \]

\[ \frac{0.24}{0.22} = 1.09 \]

Interpretation of data depends on causal assumptions!
5) Conclusions
Question: What is the (average) causal effect of X on Y?

1. Use expert knowledge to set a (plausible) causal model

2. Collect data

3. Control for confounders to isolate the causal effect

P(Y \mid \text{do}(X)) = P(Y \mid X, Z)

Confounding is anything that leads to P(Y \mid X) being different than P(Y \mid \text{do}(X))

linear case:
Y = aX + bZ
SUMMARY

- Scientific data science is never fully “objective”
- We should be transparent about our assumptions (by using causal networks)

- Causal networks help to identify where information is propagating and to extract the causal effect of interest
- Conditioning on confounders = blocking the “open” paths in the network
- Its fully non-parametric
- Implementing a causal framework only involves small changes in scientific practice but allows to draw stronger, causal statements
OUTLOOK: LEARNING CAUSAL STRUCTURES FROM DATA

Input: Time-series

Causal Discovery

PCMCI Algorithm

\[ \text{corr}(X_{t-\tau}, Y_t) \quad \text{Iterate through combinations of conditions} \]

Identifies spurious correlations due to

- common drivers
- mediators
- auto-correlation effects

Output: causal model/network

Kretschmer et al. (2016, 2018, 2019), Runge et al. (2019a, 2019b), Saggio et al. (2020)
APPLICATION REQUIRES PROCESS UNDERSTANDING

Indian Summer Monsoon
Di Capua et al. (2019), ESD

Hurricane Activity
Pfleiderer et al. (2020), WCD

Morocco Crop yield
Lehmann et al. (2020), GRL

Marine cold air outbreaks
Polkova et al. (2021), QJRMS
Call for Papers!

Special Issue: Novel data science approaches to evaluate weather and climate extremes

Submit now

Guest Editors
Marlene Kretschmer,
Aglaé Jezequel
Zach Labe
Danielle Touma

m.j.a.kretschmer@reading.ac.uk
@Marlene_Climate
THANK YOU FOR YOUR ATTENTION!