



UNIVERSITÄT ZU LÜBECK

Algorithmen für kausale Inferenz

Beitrag zum Darmstädter Ontologenkreis

Marcel Wienöbst

13. April 2022

IM FOCUS DAS LEBEN



Warum kausale Inferenz?

The New York Times

Another Benefit to Going to Museums? You May Live Longer

Researchers in Britain found that people who go to museums, the theater and the opera were less likely to die in the study period than those who didn't.



The screenshot shows the top portion of a Washington Post article. At the top, there is a navigation bar with 'Sections' on the left, the Washington Post logo and tagline 'Democracy Dies in Darkness' in the center, and a 'Try four weeks free' button on the right. Below the navigation bar is a yellow banner with a clock icon and the text 'This article was published more than 6 years ago'. The article is categorized under 'Climate and Environment'. The main headline is 'Why living around nature could make you live longer'. Below the headline, it says 'By Chelsea Harvey' and 'April 19, 2016'. There is a photograph of a sunlit forest. In the bottom right corner of the article preview, there is a small box that says 'MOST READ CLIMATE & ENVIRO'.

Warum kausale Inferenz?

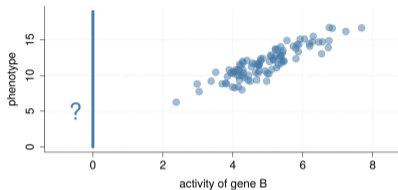
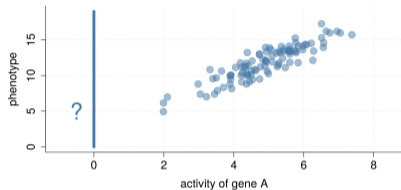


Warum kausale Inferenz?



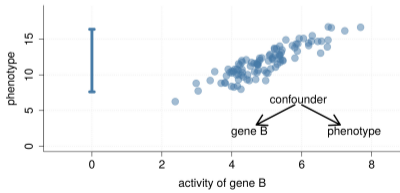
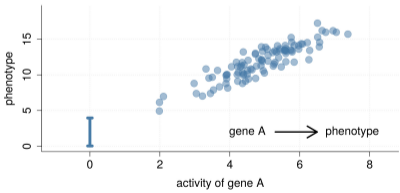
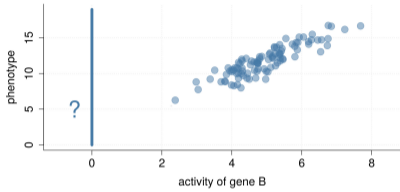
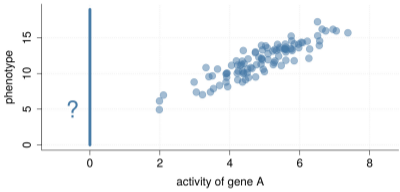
Ziel: Erkennen und quantifizieren von kausalen Beziehungen

Beispiel - Genaktivität ¹



¹Beispiel von Jonas Peters (<http://web.math.ku.dk/~peters/>)

Beispiel - Genaktivität ¹



¹Beispiel von Jonas Peters (<http://web.math.ku.dk/~peters/>)

Strukturelle kausale Modelle

Definition (Strukturelles kausales Modell (SKM); Pearl (2000))

Ein strukturelles kausales Modell besteht aus d Zuweisungen

$$X_j \leftarrow f_j(\mathbf{PA}_j, N_j) \quad \text{für } j = 1, \dots, d.$$

Strukturelle kausale Modelle

Definition (Strukturelles kausales Modell (SKM); Pearl (2000))

Ein strukturelles kausales Modell besteht aus d Zuweisungen

$$X_j \leftarrow f_j(\mathbf{PA}_j, N_j) \quad \text{für } j = 1, \dots, d.$$

Variablen:

- Sozialer Status (S)
- Gesundheit (G)
- Museumsbesuche (M)
- Lebenserwartung (L)

Strukturelles kausales Modell:

$$S \leftarrow f_S(N_S)$$

$$G \leftarrow f_G(S, N_G)$$

$$M \leftarrow f_M(S, G, N_M)$$

$$L \leftarrow f_L(G, N_L)$$

Beispiel — Museumsbesuche

Variablen:

- Sozialer Status (S)
- Gesundheit (G)
- Museumsbesuche (M)
- Lebenserwartung (L)

Strukturelles kausales Modell:

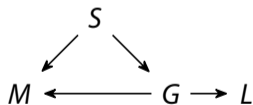
$$S \leftarrow f_S(N_S)$$

$$G \leftarrow f_G(S, N_G)$$

$$M \leftarrow f_M(S, G, N_M)$$

$$L \leftarrow f_L(G, N_L)$$

Gerichteter azyklischer Graph:

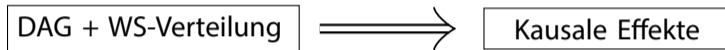


Wahrscheinlichkeitsverteilung:

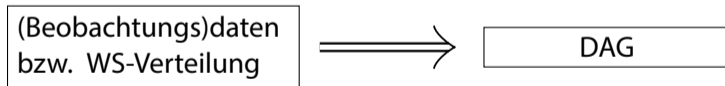
$$P(S, G, M, L) = P(S)P(G | S)P(M | S, G)P(L | G)$$

Probleme im Kontext von SCMs

Identifizierung:



(Causal) Discovery:



Causal Discovery

DAG D:

$$X_1 \rightarrow X_2 \rightarrow X_3$$

Verteilung:

$$P(X_1, X_2, X_3) = P(X_1)P(X_2 | X_1)P(X_3 | X_2)$$

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

DAG D:

$$X_1 \rightarrow X_2 \rightarrow X_3$$

Verteilung:

$$P(X_1, X_2, X_3) = P(X_1)P(X_2 | X_1)P(X_3 | X_2)$$

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

Zwei Annahmen für das Lernen von DAGs:

- $A \perp\!\!\!\perp_D B | \mathbf{C} \implies A \perp\!\!\!\perp_P B | \mathbf{C}$ *(global) Markov property*
- $A \perp\!\!\!\perp_P B | \mathbf{C} \implies A \perp\!\!\!\perp_D B | \mathbf{C}$ *Faithfulness*

Markoväquivalenzklassen

DAG D:

$$X_1 \rightarrow X_2 \rightarrow X_3$$

Verteilung bzgl. D:

$$P(X_1, X_2, X_3) = P(X_1)P(X_2 | X_1)P(X_3 | X_2)$$
$$\implies X_1 \perp\!\!\!\perp X_3 | X_2$$

Markoväquivalenzklassen

DAG D :

$$X_1 \rightarrow X_2 \rightarrow X_3$$

Verteilung bzgl. D :

$$\begin{aligned} P(X_1, X_2, X_3) &= P(X_1)P(X_2 | X_1)P(X_3 | X_2) \\ \implies X_1 &\perp\!\!\!\perp X_3 | X_2 \end{aligned}$$

DAG D' :

$$X_1 \rightarrow X_2 \leftarrow X_3$$

Verteilung bzgl. D' :

$$\begin{aligned} P(X_1, X_2, X_3) &= P(X_1)P(X_3)P(X_2 | X_1, X_3) \\ \implies X_1 &\not\perp\!\!\!\perp X_3 | X_2 \end{aligned}$$

Markoväquivalenzklassen

DAG D:

$$X_1 \rightarrow X_2 \rightarrow X_3$$

Verteilung bzgl. D:

$$\begin{aligned} P(X_1, X_2, X_3) &= P(X_1)P(X_2 | X_1)P(X_3 | X_2) \\ \implies X_1 &\perp\!\!\!\perp X_3 | X_2 \end{aligned}$$

DAG D':

$$X_1 \rightarrow X_2 \leftarrow X_3$$

Verteilung bzgl. D':

$$\begin{aligned} P(X_1, X_2, X_3) &= P(X_1)P(X_3)P(X_2 | X_1, X_3) \\ \implies X_1 &\not\perp\!\!\!\perp X_3 | X_2 \end{aligned}$$

DAG D'':

$$X_1 \leftarrow X_2 \leftarrow X_3$$

Verteilung bzgl. D'':

$$\begin{aligned} P(X_1, X_2, X_3) &= P(X_3)P(X_2 | X_3)P(X_1 | X_2) \\ \implies X_1 &\perp\!\!\!\perp X_3 | X_2 \end{aligned}$$

Algorithmische Probleme auf Markoväquivalenzklassen:

- Repräsentationen von Markoväquivalenzklassen ²
- Berechnung der Größe von Markoväquivalenzklassen ³
- Samplen von DAGs aus einer Markoväquivalenzklasse ³
- Effizientes Lernen von Markoväquivalenzklassen ⁴

²Extendability of causal graphical models: Algorithms and computational complexity; M.W., Max Bannach, Maciej Liśkiewicz; UAI'21

³Polynomial-time algorithms for counting and sampling Markov equivalent DAGs; M.W., Max Bannach, Maciej Liśkiewicz; AAI'21

⁴Recovering causal structures from low-order conditional independencies; M.W., Maciej Liśkiewicz; AAI'20

Berechnung der Größe einer Markoväquivalenzklasse

	Strategie	Komplexität
Meek '95	Brute-Force	$\mathcal{O}(n!)$
Gillespie, Perlman '01	Brute-Force	$\mathcal{O}(n!)$
He et al. '15	Root-Picking (RP)	$\mathcal{O}(n!)$
Talvitie, Koivisto '19	RP + Memoization (MM)	$\mathcal{O}(2^n \cdot n^4)$
Ghassami et al. '19	RP + MM + Clique-Tree	$\mathcal{O}(2^n \cdot n^4)$
Ganian et al. '20	RP + MM + Dom. Vertex	$\mathcal{O}(2^n \cdot n^4)$
Talvitie, Koivisto '19	DP auf dem Clique-Tree	$\mathcal{O}(k!2^k k^2 n)$
Teshnizi et al. '20	Int. Design + MM	$\mathcal{O}(2^{n+d}(nd + d^3))$
<i>W., Bannach, Liśkiewicz '21</i>	<i>Clique-Picking + MM</i>	$\mathcal{O}(n^4)$

Größe einer Markoväquivalenzklasse (MÄK)

Eingabe: CPDAG G (kompakte Repräsentation einer MÄK)

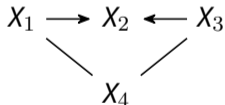
Ausgabe: Anzahl der DAGs in der MÄK repräsentiert durch G

Größe einer Markoväquivalenzklasse (MÄK)

Eingabe: CPDAG G (kompakte Repräsentation einer MÄK)

Ausgabe: Anzahl der DAGs in der MÄK repräsentiert durch G

Eingabe:



Ausgabe:

3

Anwendung: Active Learning

Eingabe: CPDAG G

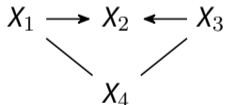
Ziel: Ermittle den 'true' DAG durch gezielte Experimente

Anwendung: Active Learning

Eingabe: CPDAG G

Ziel: Ermittle den 'true' DAG durch gezielte Experimente

Eingabe:



Experimente:

Intervenierte auf X_4

Wie können kausale Beziehungen in KI-Modellen repräsentiert werden?

- Welche Annahmen sind nötig/ausreichend?
- Nutzung von Daten aus verschiedenen Umgebungen
- Wiederverwendbare, robuste kausale Komponenten
- Agenten führen selbst Experimente durch

Vielen Dank für die Aufmerksamkeit!