HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Causal Inference
Theory and Applications in Enterprise Computing

Christopher Hagedorn, Johannes Huegle, Dr. Michael Perscheid

May 06, 2020

Agenda
May 05, 2020

- **Introduction to Group Topics**
  1. Causal Structure Learning I
  2. Causal Structure Learning II
  3. Application Scenario
  4. Causal Inference

- **Recap: Causal Graphical Models**

- **Jupyter Lab**
  1. Causal Inference in a Nutshell - Cooling House Scenario
  2. Causal Graphical Models - Exercises

Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Introduction to Group Topics

Constraint-Based Causal Structure Learning Within Non-Linear Settings

In real world applications, the data is often not following a mutlivariate normal distribution such as in our cooling house example, but is more complex. Recently developed conditional independence tests consider these settings. Yet, they are computational expensive, yielding long execution times.

Introduction to Group Topics
Causal Structure Learning I

Research Question/Task:

Determine bottlenecks in an existing test method, in the context of constraint-based causal structure learning, in particular the PC-Algorithm. Address the determined bottlenecks, by using efficient data structures or applying tailored parallel execution strategies with the goal to reduce computation time.

Reading Material:

- CI-test based on a nearest-neighbor estimator of conditional mutual information
  https://elib.dlr.de/126425/1/Runge18.pdf
- Framework for Parallel Constraint-Based Learning
  https://www.jstatsoft.org/article/view/v077i02

Heterogeneous Computing for Constraint-Based Causal Structure Learning

GPUs provide massive parallel compute power to conduct independence tests in parallel, which have been adopted in the PC-Algorithm. Yet, during learning CPU resources remain widely unused.

Introduction to Group Topics
Causal Structure Learning II

Research Question/Task:

How can CPU resources be efficiently utilized during the GPU-accelerated algorithm‘s execution to reduce computation time? Further, is there a sweet spot between GPU/CPU execution, in which switching between the mode‘s is preferable?

Reading Material:

- cuPC: CUDA-based Parallel PC Algorithm for Causal Structure Learning on GPU
  https://arxiv.org/abs/1812.08491
  Implementation: https://github.com/LIS-Laboratory/cupc
- PC-Algorithm R-Implementation
  https://cran.r-project.org/web/packages/pcalg/index.html
  Framework for Parallel Constraint-Based Learning
  https://www.jstatsoft.org/article/view/v077i02

From Data to Causal Inference:
A Printing Press Example

Learning causal models has a large potential in real world settings. Despite the limitations, given that data is usually not captured with the application of causal inference in mind, it is important to understand the value of learnt causal graphical models.

Introduction to Group Topics
Application Scenario

Research Question/Task:

Understand the provided real world use case and its data. What needs to be considered during data preprocessing and handling in order to apply causal structure learning and further apply causal inference on the learnt model?

Reading Material:

- Master Thesis by Daniel Thevessen
  excerpts provided via e-mail
- Slides from use case
  provided via e-mail

Causal Inference,
Do-Operator, etc.
in non-trivial Settings

In real world applications, the data is often not following a multivariate normal distribution such as in our cooling house example, but is rather more complex. Thus, causal inference methods need to address these settings to become relevant in practice.

Introduction to Group Topics
Causal Inference

Research Question/Task:

Understand and compare existing approaches for causal inference within different assumptions on underlying functional relationships,
e.g., in the context of heterogeneous data.
Determine limitations in different settings and if possible suggest improvements.

Reading Material:

- Probabilistic Active Learning of Functions in Structural Causal Models
  https://arxiv.org/pdf/1706.10234.pdf
- Quantifying causal influences
  https://arxiv.org/abs/1203.6502

3. Introduction to Research Topics
Topic Application

How to work on a topic?

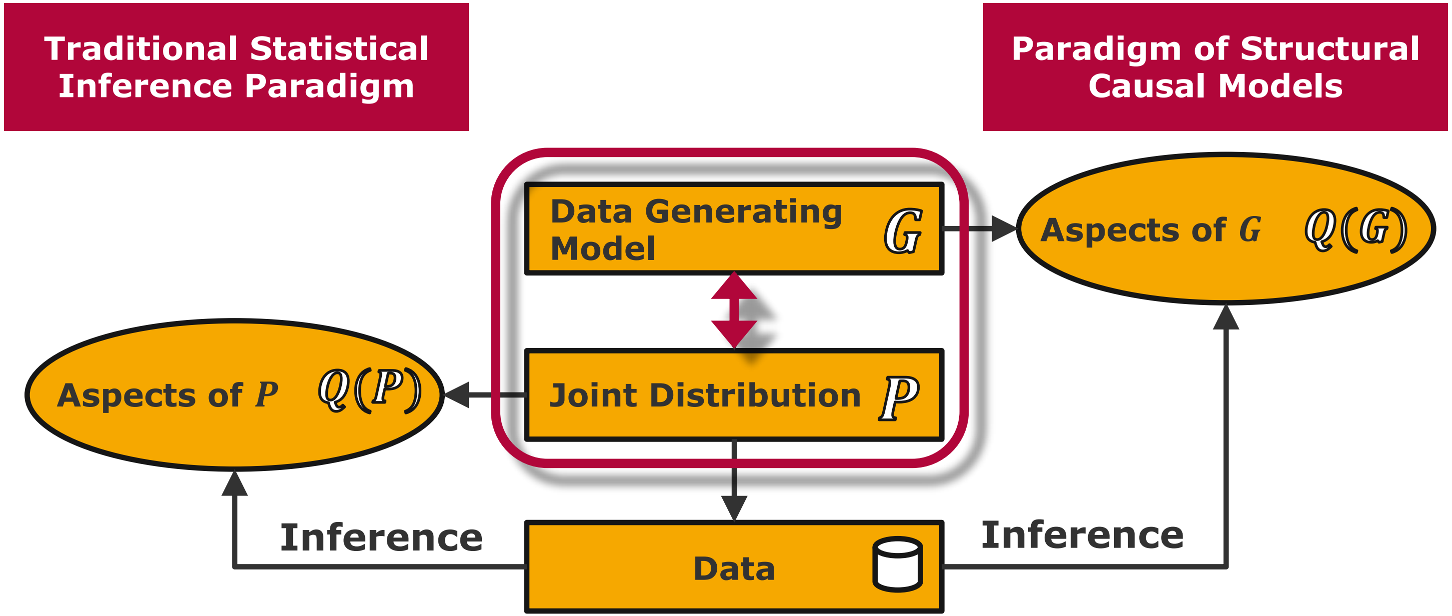
1. Understand theoretic basis and your selected topic
2. Work on implementation
3. Present results
4. Write scientific report in a review process

How to apply for a topic?

- Build groups of around two or three students
- Send prioritized list of 3 topics to Johannes Huegle until: *Sun May 10, 11.59 PM*
- Topic Assignments: *Wed May 13, 9:15 AM*

Recap: Causal Inference in a Nutshell

Recap: Causal Inference in a Nutshell
Concept



E.g., what is the sailors' probability of recovery when **we see** a treatment with lemons?

$$Q(P) = P(recovery|lemons)$$

E.g., what is the sailors' probability of recovery if **we do** treat them with lemons?

$$Q(G) = P(recovery|do(lemons))$$

Recap: Causal Inference in a Nutshell
Causal Graphical Models

Data Generating Model $\mathcal{G}$

↓

Joint Distribution $\mathcal{P}$

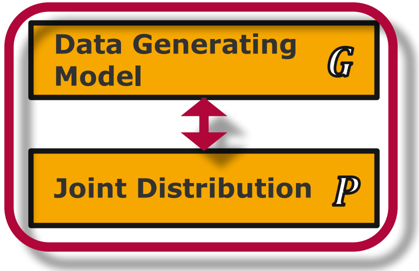
Causal Graphical Model

- *Directed Acyclic Graph (DAG)* $G = (V, E)$
  - *Vertices* $V_1, \dots, V_n$
  - *Directed edges* $E = (V_i, V_j)$, i.e., $V_i \to V_j$
  - *No cycles*
- *Directed Edges* encode direct causes via
  - $V_j = f_j(\mathrm{Pa}(\mathrm{V_j}), \mathrm{N_j})$ with independent noise $N_1, \dots, N_n$

Causal Sufficiency

- All relevant variables are included in the DAG G

Recap: Causal Inference in a Nutshell
Connecting G and P



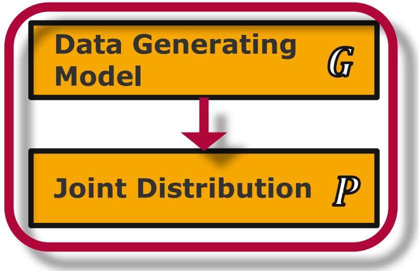
$$(X \perp\!\!\!\perp Y|Z)_G \Rightarrow (X \perp\!\!\!\perp Y|Z)_P$$

- Key Postulate: *(Local) Markov Condition*
- Essential mathematical concept: *d-Separation*
  - Idea: *Blocking* of paths
  - Implication: *Global Markov Condition*

$$(X \perp\!\!\!\perp Y|Z)_G \Leftarrow (X \perp\!\!\!\perp Y|Z)_P$$

- Key Postulate: *Causal Faithfulness*

Recap: Causal Inference in a Nutshell
Local Markov Condition and Blocking of Paths



(Local) Markov Condition

V_j independent of nondescendants $ND(V_j)$ given parents $Pa(V_j)$, i.e.,

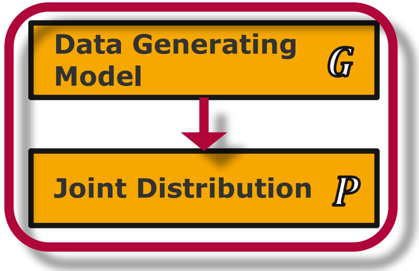
$$\boldsymbol{V_j \perp\!\!\!\perp V_{V/(Des(V_j) \cup Pa(V_j))} | Pa(V_j)}.$$

Blocking of Paths

A path q is said to be *blocked* by a set S if

- q contains a *chain* $V_i \rightarrow V_j \rightarrow V_k$ or a *fork* $V_i \leftarrow V_j \rightarrow V_k$ such that the middle node is in S, or
- q contains a *collider* $V_i \rightarrow V_j \leftarrow V_k$ such that the middle node is not in S and such that no descendant of V_j is in S.

Recap: Causal Inference in a Nutshell
D-Separation and Global Markov Condition



D-Separation

S is said to *d-separate* X and Y in the DAG G, i.e.,

$$(X \perp\!\!\!\perp Y \mid S)_G$$

if S blocks every path from a vertex in X to a vertex in Y.

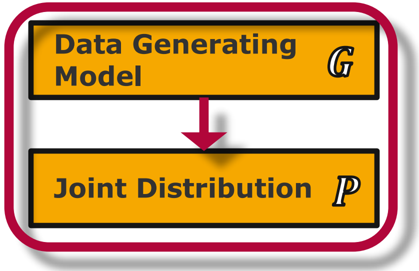
Global Markov Condition

For all disjoint subsets of vertices X, Y and S we have that

$$X, Y \text{ d-separated by } S \Rightarrow (X \perp\!\!\!\perp Y \mid S)_P.$$

Recap: Causal Inference in a Nutshell

$(X \perp\!\!\!\perp Y|S)_G \Rightarrow (X \perp\!\!\!\perp Y|S)_P$



Theorem

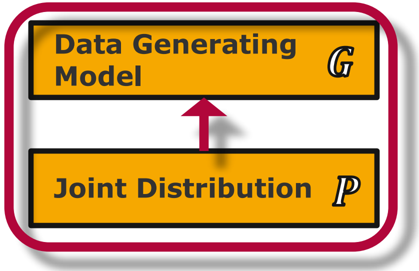
The following are equivalent:

- Existence of a *functional Causal Model* G;
- *(Local) Markov condition*: statistical independence of nondescendants given parents (i.e.: every information exchange with its nondescendants involves its parents)
- *Global Markov condition*: d-separation (characterizes the set of independences implied by local Markov condition)
- *Factorization:* $p(v_1, \ldots, v_n) = \prod_{i=1}^{n} p(v_i|Pa(v_i))$.

(subject to technical conditions)

Recap: Causal Inference in a Nutshell

$(X \perp\!\!\!\perp Y|S)_G \Leftarrow (X \perp\!\!\!\perp Y|S)_P$



Causal Faithfulness

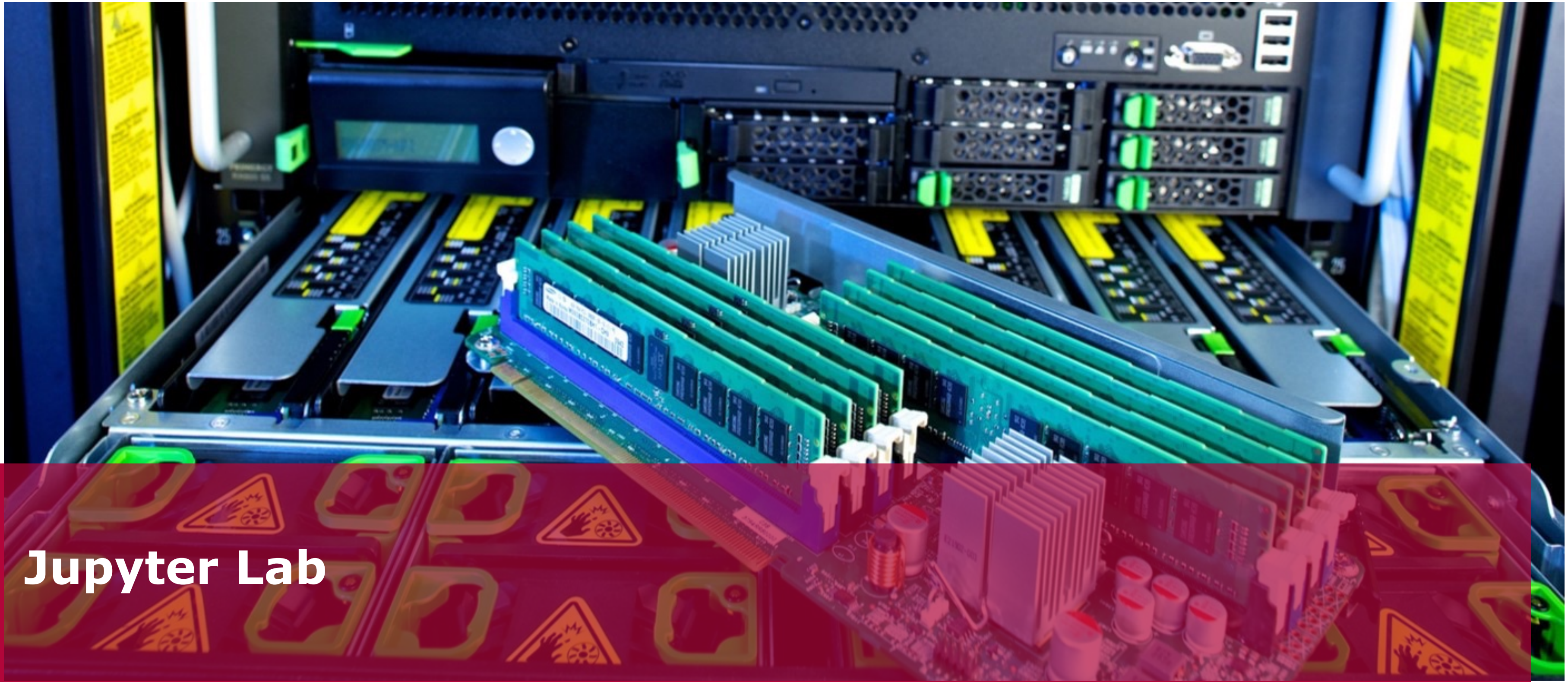
p is called faithful relative to G if only those independencies hold true that are implied by the Markov condition, i.e.,

$$(X \perp\!\!\!\perp Y|\, S)_G \Leftarrow (X \perp\!\!\!\perp Y|\, S)_P$$

Markov Equivalence Class

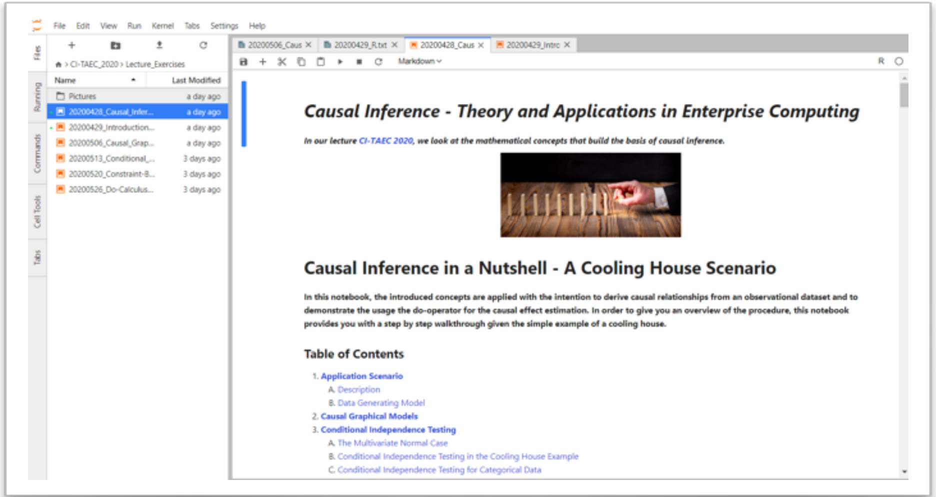
Two DAGs are Markov equivalent if and only if they have the same skeleton and the same v-structures

Hasso Plattner Institut
Digital Engineering · Universität Potsdam



Jupyter Lab

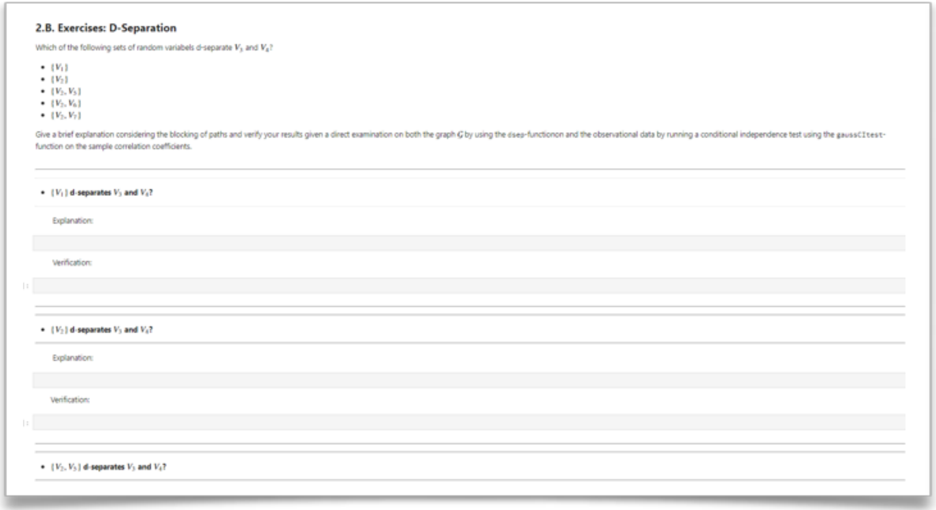
Jupyter Lab
Causal Inference in a Nutshell - Cooling House



Topics

- Data generating model
- D-separation in application
- Conditional independence in P

Jupyter Lab
Causal Graphical Models - Exercises


2.B. Exercises: D-Separation

Which of the following sets of random variabels d-separate V_3 and V_4?

- $\{V_1\}$
- $\{V_2\}$
- $\{V_2, V_5\}$
- $\{V_2, V_6\}$
- $\{V_2, V_7\}$

Give a brief explanation considering the blocking of paths and verify your results given a direct examination on both the graph G by using the `dsep`-functionon and the observational data by running a conditional independence test using the `gaussCItest`-function on the sample correlation coefficients.

- **$\{V_1\}$ d-separates V_3 and V_4?**

Explanation:

Verification:

- **$\{V_2\}$ d-separates V_3 and V_4?**

Explanation:

Verification:

- **$\{V_2, V_5\}$ d-separates V_3 and V_4?**

Exercises

- Causal Graphical Models
- D-Separation

Thank you
for your attention!