Causal inference in Observational Time Series

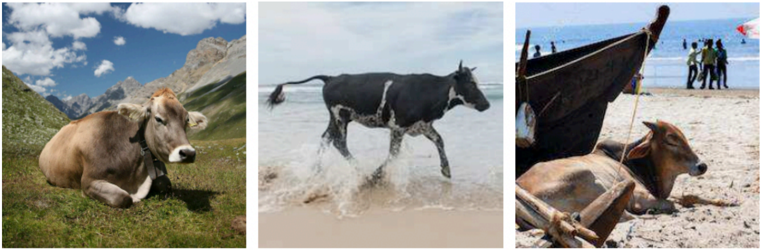
E. Devijver
C. K. Assaad, E. Gaussier, G. Gössler, A. Meynaoui, L. Zan

CNRS, Université Grenoble Alpes, France
Chaire Causalité, MIAI Institute

October 19, 2023

Why causality?

- Lack of generalization for ML models



(A) **Cow: 0.99**, Pasture: 0.99, Grass: 0.99, No Person: 0.98, Mammal: 0.98

(B) No Person: 0.99, Water: 0.98, Beach: 0.97, Outdoors: 0.97, Seashore: 0.97

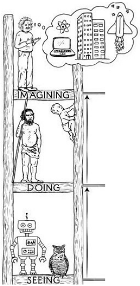
(C) No Person: 0.97, **Mammal: 0.96**, Water: 0.94, Beach: 0.94, Two: 0.94

Why causality?

- Lack of generalization for ML models
- Correlation is not causation

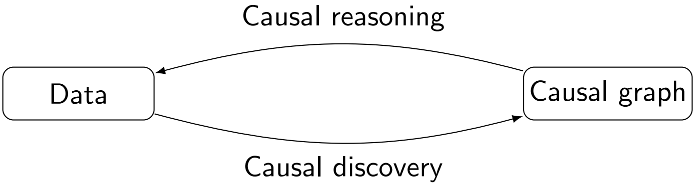
Why causality?

- Lack of generalization for ML models
- Correlation is not causation
- Want to answer causal questions
  - Counterfactual: Would my grandfather still be alive if he was vegetarian?
  - Intervention: How would my expected lifespan change if I become a vegetarian?
  - Association: What is the expected lifespan of somebody who is vegetarian?



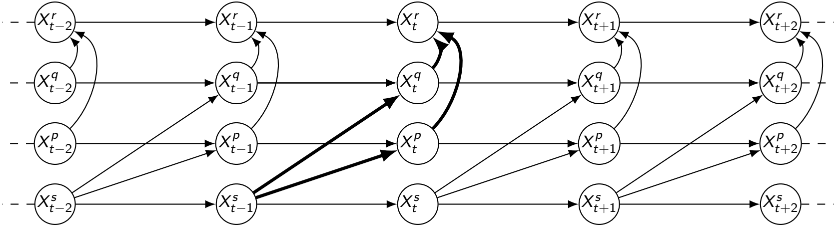
Why causality?

- Lack of generalization for ML models
- Correlation is not causation
- Want to answer causal questions



Outline

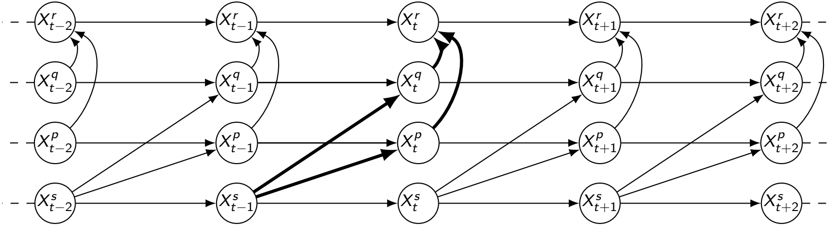
Causal graphs for time series



Full time causal graph.

A d-variate time series X
For a fixed t, each X_t is a vector $(X_t^1, \ldots, X_t^d)$,
in which X_t^p is the measurement of the pth time series at time t.

Causal graphs for time series

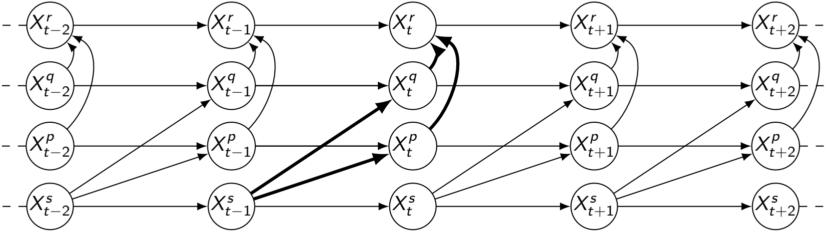


Full time causal graph.

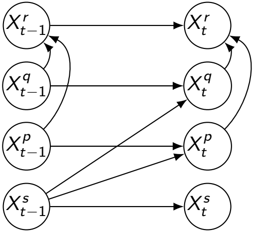
A d-variate time series X
For a fixed t, each X_t is a vector $(X_t^1, \ldots, X_t^d)$,
in which X_t^p is the measurement of the pth time series at time t.

A causal graph for a multivariate time series X is said to be *consistent throughout time* if all the causal relationships remain constant in direction throughout time.

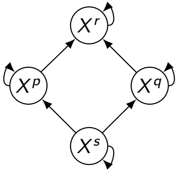
Causal graphs for time series



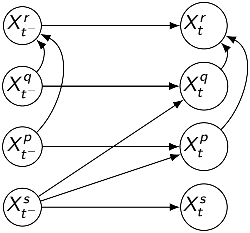
Full time causal graph.



Window Causal Graph



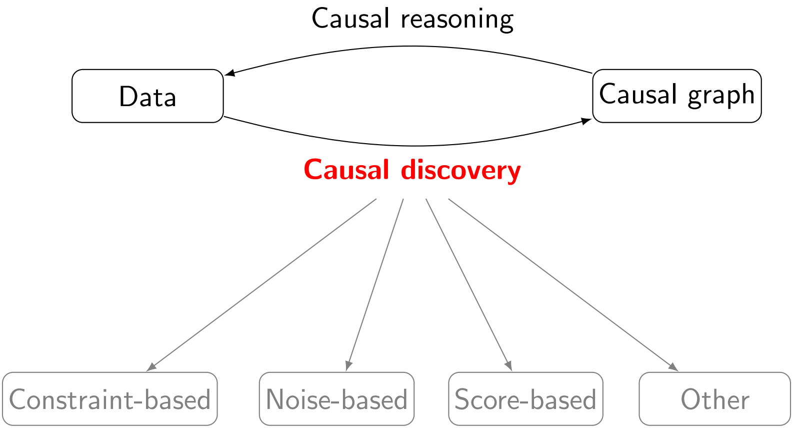
Summary Causal Graph



Extended Summary Causal Graph

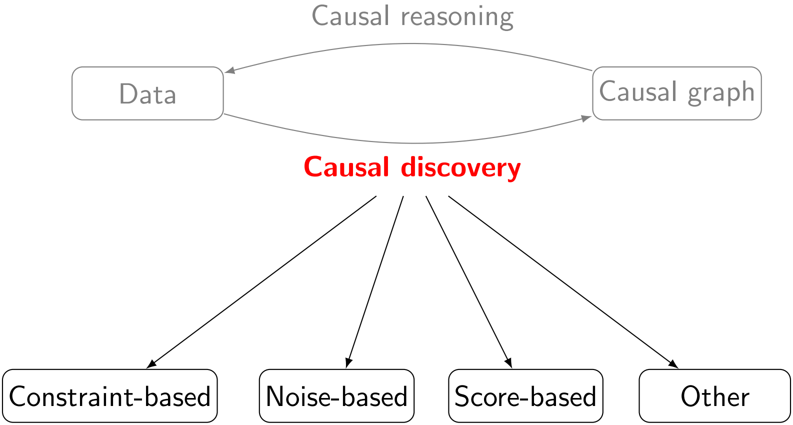
Outline

Causal discovery[1]



[1] C. K. Assaad, E. Devijver, and E. Gaussier. *Survey and evaluation of causal discovery methods for time series*. JAIR, 73, 2022.

Causal discovery[1]

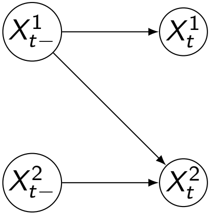


[1] C. K. Assaad, E. Devijver, and E. Gaussier. *Survey and evaluation of causal discovery methods for time series*. JAIR, 73, 2022.

Causal discovery

Classical Assumptions

- A set of variables is said to be *causally sufficient* if all common causes of all variables are observed.
- A causal relation between two variables is said to satisfy the *temporal priority* if it is oriented in such a way that the cause occurred before its effect.
- *Causal Markov Condition*: (conditional) independence in the graph leads to (conditional) independence in the data.



$$X^1_t \perp\!\!\!\perp X^2_t \mid X^1_{t-}$$

Causal discovery

Classical Assumptions

- A set of variables is said to be *causally sufficient* if all common causes of all variables are observed.
- A causal relation between two variables is said to satisfy the *temporal priority* if it is oriented in such a way that the cause occurred before its effect.
- *Causal Markov Condition*: (conditional) independence in the graph leads to (conditional) independence in the data.
- *Minimality condition*: the graph does not contain dependencies not present in the observational data.
- *Faithfulness*: only the conditional independence relations true in the data are entailed by the Causal Markov condition applied to the graph.

$$\begin{tikzcd}
 & X_t^1 \arrow[dl, "+"'] \arrow[dr, "-"] & \\
X_t^2 \arrow[dr, "+"'] & & X_t^3 \arrow[dl, "-"] \\
 & X_t^4 &
\end{tikzcd}$$

Causal discovery

Granger Causality[2]

[2] Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–38.

Causal discovery

Granger Causality[2]

X^p Granger-causes X^q if past values of X^p provide unique statistically significant information about future values of X^q.

Pairwise Granger causality

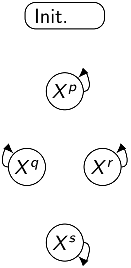


Figure: Running example: structure inferred by the pairwise Granger method (an arbitrary order has been chosen for the example).

[2]Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–38.

Causal discovery

Granger Causality[2]

X^p Granger-causes X^q if past values of X^p provide unique statistically significant information about future values of X^q.

Pairwise Granger causality

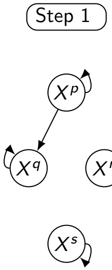


Figure: Running example: structure inferred by the pairwise Granger method (an arbitrary order has been chosen for the example).

[2] Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–38.

Causal discovery

Granger Causality[2]

X^p Granger-causes X^q if past values of X^p provide unique statistically significant information about future values of X^q.

Pairwise Granger causality

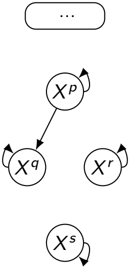


Figure: Running example: structure inferred by the pairwise Granger method (an arbitrary order has been chosen for the example).

[2] Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–38.

Causal discovery

Granger Causality[2]

X^p Granger-causes X^q if past values of X^p provide unique statistically significant information about future values of X^q.

Pairwise Granger causality

Step 6

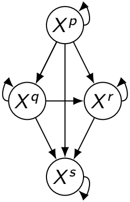


Figure: Running example: structure inferred by the pairwise Granger method (an arbitrary order has been chosen for the example).

[2] Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–38.

Multivariate Granger causality

$$X_t^q = a_{q,0} + \sum_{\substack{r=1 \\ r \neq p}}^{d} \sum_{i=1}^{\tau} a_{r,i} X_{t-i}^p + \xi_t^q, \qquad \text{(mvMres)}$$

$$X_t^q = a_{q,0} + \sum_{r=1}^{d} \sum_{i=1}^{\tau} a_{r,i} X_{t-i}^r + \xi_t^q, \qquad \text{(mvMfull)}$$

Extensions

- Non-linear associations
- Nonstationnarity

Causal discovery

Constraint-based approaches

- Exploiting conditional independencies to build a skeleton
- Skeleton is oriented according to a set of rules that define constraints on admissible orientations.

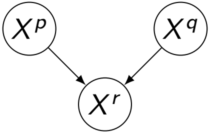
- Exploiting conditional independencies to build a skeleton
- Skeleton is oriented according to a set of rules that define constraints on admissible orientations.

Assumptions

- Causal Markov Condition
- Faithfulness

Structures with 3 nodes

$X^p \not\perp\!\!\!\perp X^q | X^r$

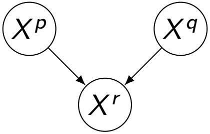


Causal discovery

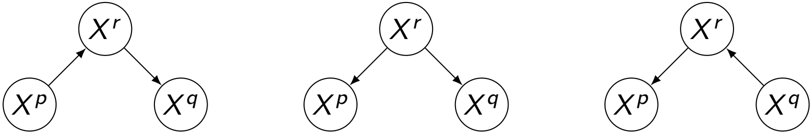
Constraint-based approaches

Structures with 3 nodes

$X^p \not\!\perp\!\!\!\perp X^q | X^r$



$X^p \perp\!\!\!\perp X^q | X^r$



Markov equivalence class: set of DAGs that encode the same set of conditional independencies.

Main difficulty when dealing with time series: determine a good measure of (conditional) dependencies

- Measure : (partial) correlation, entropy, mutual information ...
- Estimation
- Type of data (continuous, discrete, mixed)

Representation of the time series

Optimal lag γ_{pq} and $(\lambda_{pq}, \lambda_{qp})$ the *optimal* windows:

$$\gamma_{pq}, \lambda_{pq}, \lambda_{qp} = \underset{\gamma \geq 0, \lambda_1, \lambda_2}{\operatorname{argmax}} \; h(X^q_{t:t+\lambda_2} \mid X^q_{t-1}, X^p_{t-\gamma-1}) - h(X^q_{t:t+\lambda_2} \mid X^p_{t-\gamma-1:t-\gamma+\lambda_1}, X^q_{t-1}).$$

where h denotes the entropy.

Causal discovery

Constraint-based approaches: PCMCI[3] to discover a window causal graph

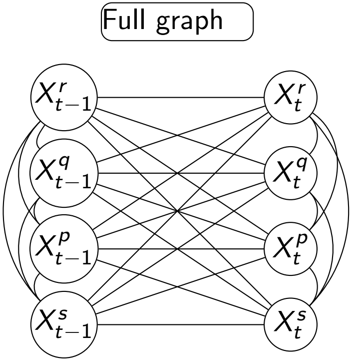


Figure: Running example: structure inferred by PCMCI with instantaneous relations.

[3] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11).

Causal discovery

Constraint-based approaches: PCMCI[3] to discover a window causal graph

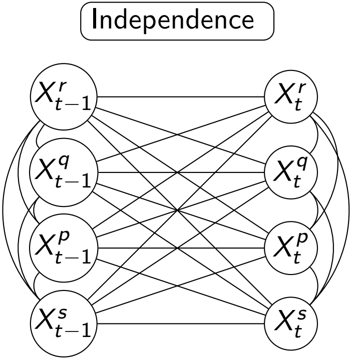


Figure: Running example: structure inferred by PCMCI with instantaneous relations.

[3] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11).

Causal discovery

Constraint-based approaches: PCMCI[3] to discover a window causal graph

Conditional independence

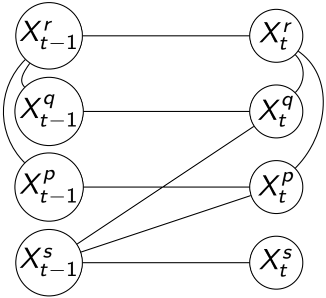


Figure: Running example: structure inferred by PCMCI with instantaneous relations.

[3] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11).

Causal discovery

Constraint-based approaches: PCMCI[3] to discover a window causal graph

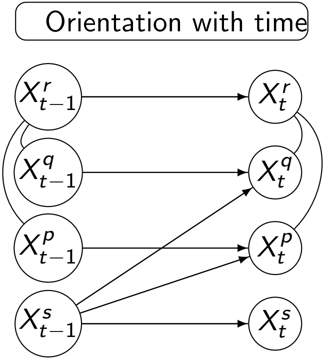


Figure: Running example: structure inferred by PCMCI with instantaneous relations.

[3] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11).

Causal discovery

Constraint-based approaches: PCMCI[3] to discover a window causal graph

Orientation with PC

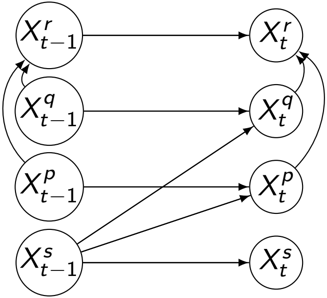


Figure: Running example: structure inferred by PCMCI with instantaneous relations.

[3] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11).

Causal discovery

Constraint-based approaches: PCGCE[4] to discover an extended summary causal graph

Assumptions

- Causal Markov condition
- Faithfulness
- Causal sufficiency for PCGCE (but extension to FCIGCE)

[4] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Constraint-based approaches: PCGCE[4] to discover an extended summary causal graph

Assumptions

- Causal Markov condition
- Faithfulness
- Causal sufficiency for PCGCE (but extension to FCIGCE)

Measure: need a specific one due to the graph structure
Proposed Greedy Causation Entropy (GCE)

$$\mathrm{GCE}(X^p \rightarrow X^q | X^{\mathbf{Pa}}, X^{\mathbf{Pr}})$$
$$:= I(X_t^q; X_{t-\gamma:t-1}^p | X_{t-}^{Pa_1}, \cdots, X_{t-}^{Pa_l}, X_t^{Pr_1}, \cdots, X_t^{Pr_m})$$

[4] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Constraint-based approaches: PCGCE[5] to discover an extended summary causal graph

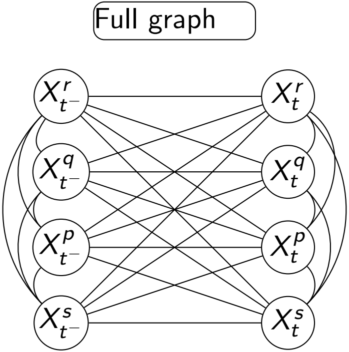


Figure: Running example: structure inferred by PCGCE with instantaneous relations.

[5] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Constraint-based approaches: PCGCE[5] to discover an extended summary causal graph

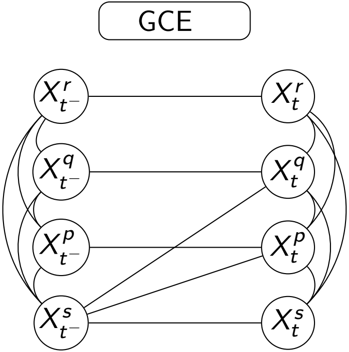


Figure: Running example: structure inferred by PCGCE with instantaneous relations.

[5] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Constraint-based approaches: PCGCE[5] to discover an extended summary causal graph

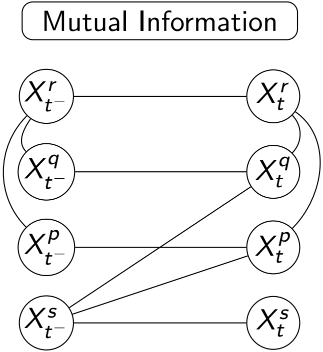


Figure: Running example: structure inferred by PCGCE with instantaneous relations.

[5] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Constraint-based approaches: PCGCE[5] to discover an extended summary causal graph

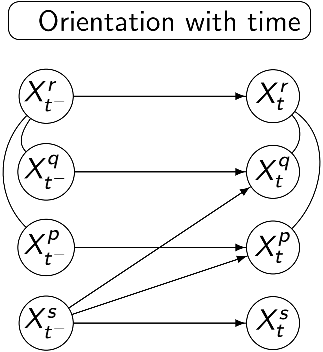


Figure: Running example: structure inferred by PCGCE with instantaneous relations.

[5] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Constraint-based approaches: PCGCE[5] to discover an extended summary causal graph

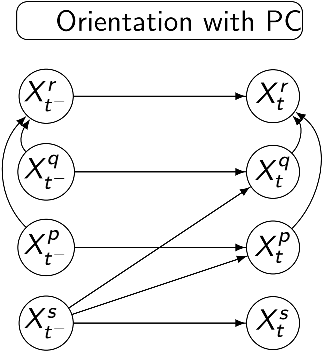


Figure: Running example: structure inferred by PCGCE with instantaneous relations.

[5] C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022

Causal discovery

Noise-based approaches

Causal discovery

Noise-based approaches

- Causal system described by a set of equations (SEM/SCM)
- each equation explains one variable of the system in terms of its direct causes and some additional noise.

Causal discovery

Noise-based approaches

- Causal system described by a set of equations (SEM/SCM)
- each equation explains one variable of the system in terms of its direct causes and some additional noise.

Assumptions

- Causal Markov Condition
- Minimality

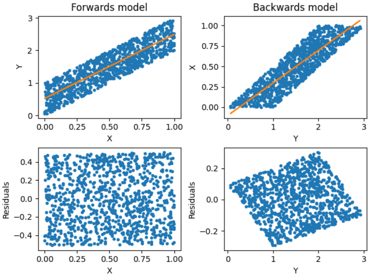
Can deal with 2 variables

Causal discovery

Noise-based approaches: Additive Noise Models

Additive noise model with nonlinear functions

$$
\begin{aligned}
X^p &= \xi^p, \\
X^q &= f_q(X^p) + \xi^q \qquad \text{with } X^p \perp\!\!\!\perp \xi_q.
\end{aligned}
$$

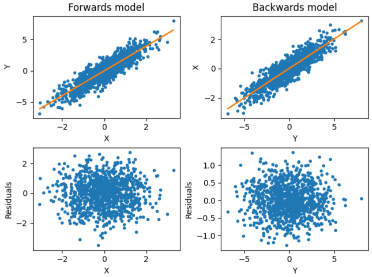


Causal discovery

Noise-based approaches: Additive Noise Models

Additive noise model with nonlinear functions

$$X^p = \xi^p,$$
$$X^q = f_q(X^p) + \xi^q \qquad \text{with } X^p \perp\!\!\!\perp \xi_q.$$



Causal discovery

Noise-based approaches: Additive Noise Models

Additive noise model with nonlinear functions

$$\begin{aligned} X^p &= \xi^p, \\ X^q &= f_q(X^p) + \xi^q \qquad \text{with } X^p \perp\!\!\!\perp \xi_q. \end{aligned}$$

Theorem (Identifiability of ANMs[6])

Assume that the conditional distribution of $X^q \mid X^p$ admits a smooth ANM, and that there exists $x_q \in \mathbb{R}$ such that, for almost all $x_p \in \mathbb{R}$,

$$(\log p_{\xi^q})''(x_q - f_q(x_p))f_q'(x_p) \neq 0.$$

Then, the set of log densities $\log p_X$ for which the obtained joint distribution P_{X^p,X^q} admits a smooth ANM from X^q to X^p is contained in a 3-dimensional affine space.

[6] Hoyer, P. O., Janzing, D., Mooij, J. M., Peters, J., Schölkopf, B. *Nonlinear causal discovery with additive noise models*. NeurIPS 2009

Causal discovery

Noise-based approaches: Additive Noise Models

Additive noise model with nonlinear functions

$$X^p = \xi^p,$$
$$X^q = f_q(X^p) + \xi^q \qquad \text{with } X^p \perp\!\!\!\perp \xi_q.$$

Principle (Multivariate additive noise principle)

Suppose we are given a joint distribution $P(X^1, \cdots, X^d)$. If it satisfies an identifiable Additive Noise Model such that $\{(X^p_{t-j})_{1\leq p\neq q\leq d, 0\leq j\leq\tau}, (X^q_{t-j})_{1\leq j\leq\tau}\} \rightarrow X^q$, then it is likely that $\{(X^p_{t-j})_{1\leq p\neq q\leq d, 0\leq j\leq\tau}, (X^q_{t-j})_{1\leq j\leq\tau}\}$ precedes X^q in the causal order.

Causal discovery

Noise-based approaches: VarLINGAM[6]

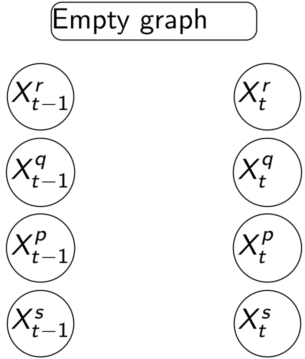


Figure: Running example: structured inferred by VarLiNGAM.

[6] Hyvarinen, A., Zhang, K., Shimizu, S., Hoyer, P. O. *Estimation of a structural vector autoregression model using non-gaussianity*. JMLR 2010

Causal discovery

Noise-based approaches: VarLINGAM[6]

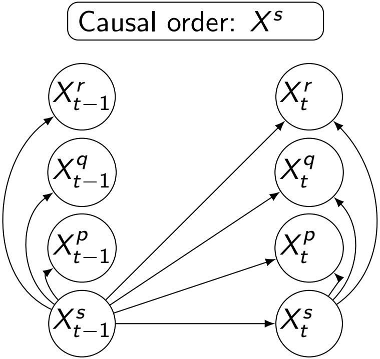


Figure: Running example: structured inferred by VarLiNGAM.

[6] Hyvarinen, A., Zhang, K., Shimizu, S., Hoyer, P. O. *Estimation of a structural vector autoregression model using non-gaussianity*. JMLR 2010

Causal discovery

Noise-based approaches: VarLINGAM[6]

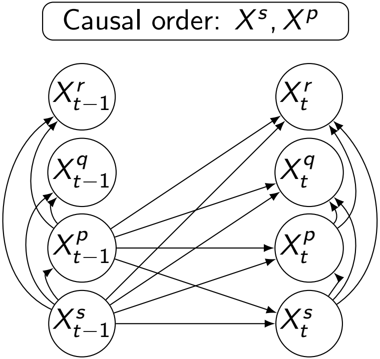


Figure: Running example: structured inferred by VarLiNGAM.

[6] Hyvarinen, A., Zhang, K., Shimizu, S., Hoyer, P. O. *Estimation of a structural vector autoregression model using non-gaussianity*. JMLR 2010

Causal discovery

Noise-based approaches: VarLINGAM[6]

Causal order: X^s, X^p, X^q

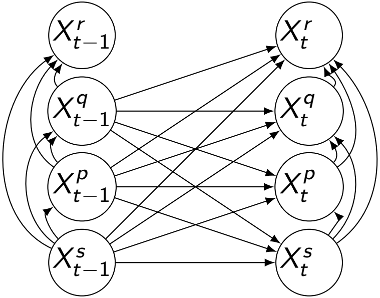


Figure: Running example: structured inferred by VarLiNGAM.

[6] Hyvarinen, A., Zhang, K., Shimizu, S., Hoyer, P. O. *Estimation of a structural vector autoregression model using non-gaussianity*. JMLR 2010

Causal discovery

Noise-based approaches: VarLINGAM[6]

Causal order: X^s, X^p, X^q, X^r

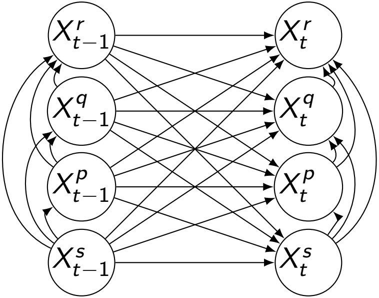


Figure: Running example: structured inferred by VarLiNGAM.

[6] Hyvarinen, A., Zhang, K., Shimizu, S., Hoyer, P. O. *Estimation of a structural vector autoregression model using non-gaussianity*. JMLR 2010

Causal discovery

Noise-based approaches: VarLINGAM[6]

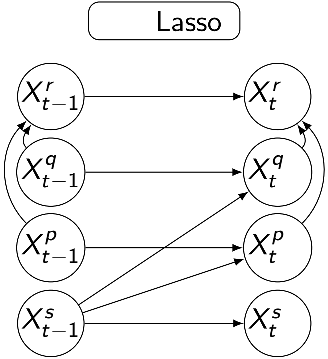


Figure: Running example: structured inferred by VarLiNGAM.

[6] Hyvarinen, A., Zhang, K., Shimizu, S., Hoyer, P. O. *Estimation of a structural vector autoregression model using non-gaussianity*. JMLR 2010

Causal discovery

NBCB[7]: a mix between noise-based and constraint-based approaches

[7] C. K. Assaad, E. Devijver, and E. Gaussier. *A Mixed Noise and Constraint Based Approach to Causal Inference in Time Series*, ECMLPKDD 2021

Causal discovery

NBCB[7]: a mix between noise-based and constraint-based approaches

Assumptions

- Causal Markov Condition
- Adjacency faithfulness: if X^p and X^q are adjacent, then they are not conditionally independent given any subset of vertices except Xp, Xq.
- Minimality

[7] C. K. Assaad, E. Devijver, and E. Gaussier. *A Mixed Noise and Constraint Based Approach to Causal Inference in Time Series*, ECMLPKDD 2021

Causal discovery

NBCB[7]: a mix between noise-based and constraint-based approaches

Assumptions

- Causal Markov Condition
- Adjacency faithfulness: if X^p and X^q are adjacent, then they are not conditionally independent given any subset of vertices except Xp, Xq.
- Minimality

Step 1: causal ordering (additive noise model)
Last place: time series which yields the residuals that are more independent to the other time series.
Step 2: pruning to remove spurious relations based on (conditional) independence measure.

[7] C. K. Assaad, E. Devijver, and E. Gaussier. *A Mixed Noise and Constraint Based Approach to Causal Inference in Time Series*, ECMLPKDD 2021

Causal discovery

NBCB: a mix between noise-based and constraint-based approaches

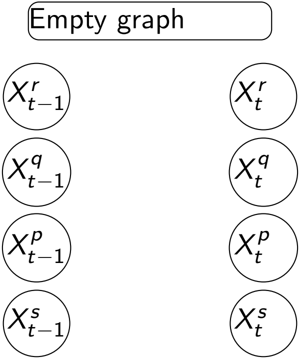


Figure: Running example: structured inferred by NBCB.

Causal discovery

NBCB: a mix between noise-based and constraint-based approaches

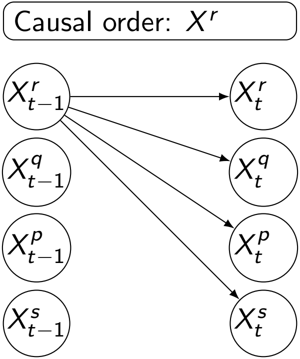


Figure: Running example: structured inferred by NBCB.

Causal discovery

NBCB: a mix between noise-based and constraint-based approaches

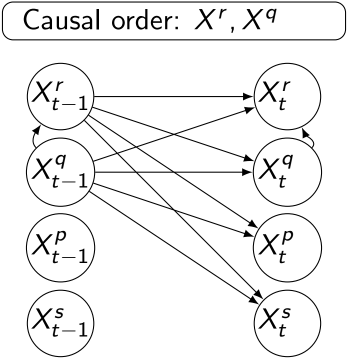


Figure: Running example: structured inferred by NBCB.

Causal discovery

NBCB: a mix between noise-based and constraint-based approaches

Causal order: X^r, X^q, X^p

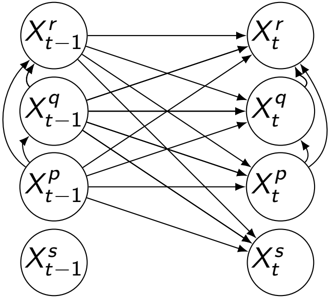


Figure: Running example: structured inferred by NBCB.

Causal discovery

NBCB: a mix between noise-based and constraint-based approaches

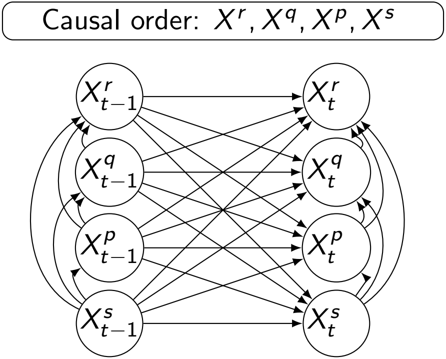


Figure: Running example: structured inferred by NBCB.

Causal discovery

NBCB: a mix between noise-based and constraint-based approaches

Conditional independence using TCE

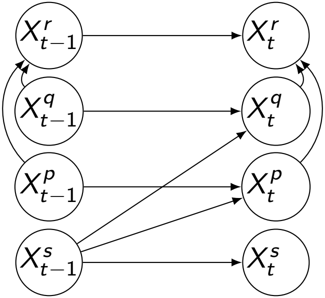
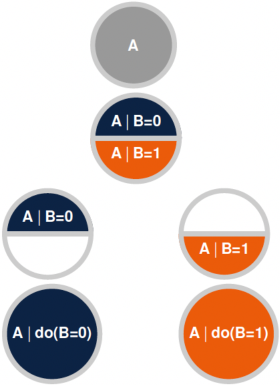


Figure: Running example: structured inferred by NBCB.

Outline

Causal reasoning
Intervention

- Population
- Sub-populations
- Conditioning
- Intervening



Causal reasoning

Problem Statement

Estimating without bias the total effect of an intervention

$$P(Y_t = y_t | do(X_{t-\gamma} = x_{t-\gamma})) = P(y_t | do(x_{t-\gamma})).$$

- Experimentation? Can be costly, unethical or even unfeasible.
- Directly from observational data? Identifiability

The total effect $P(y_t \mid do(x_{t-\gamma}))$ is said to be *identifiable* from a graph if it can be uniquely computed from the observed distribution, without any further assumption on the distribution.

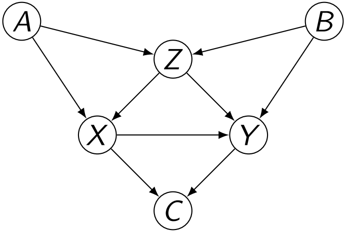
Causal reasoning

Problem Statement

Example: back-door criterion

Consider a causal graph and the total effect $P(y|do(x))$. Let $\mathcal{Z}$ a set of variables with no descendant of X and that blocks every path between X and Y that contains an arrow into X. Then, if $P(x,z) > 0$,

$$P(y|do(x)) = \sum_z P(y|x,z)P(z).$$



Assumptions: causal sufficiency, consistency throughout time.

Theorem 1[8]: Consider an FTCG $\mathcal{G}^f$ (or equivalently a WCG). The total effect $P(y_t|do(x_{t-\gamma}))$, with $\gamma \geq 0$ is identifiable in $\mathcal{G}^f$.

Theorem 2: Consider an ECG $\mathcal{G}^e$. The total effect $P(y_t|do(x_{t-\gamma}))$, with $\gamma \geq 0$ is identifiable in $\mathcal{G}^e$.

[8] Blondel et al. 2016,Shpitser et al. 2008

Causal reasoning

Identifiability in SCG[9]

Assumptions: causal sufficiency, consistency throughout time.

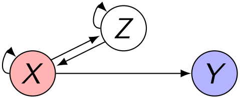
Theorem 3: Consider an SCG $\mathcal{G}^s = (\mathcal{V}^s, \mathcal{E}^s)$. The total effect $P(y_t | do(x_{t-\gamma}))$, with $\gamma \geq 0$, is not identifiable if and only if $X \in Anc(Y, \mathcal{G}^s)$ and one of the following holds:

- $\gamma \neq 0$ and $Cycles^{>}(X, \mathcal{G}^s \backslash \{Y\}) \neq \emptyset$, or
- there exists a σ-active back-door path $\pi^s = \langle V^1 = X, \cdots, V^n = Y \rangle$ such that $\langle V^2, \cdots, V^{n-1} \rangle \subseteq Desc(X, \mathcal{G}^s)$ and one of the following holds:
  - $n > 2$, ie $\langle V^2, \cdots, V^{n-1} \rangle \neq \emptyset$, or
  - $\gamma \neq 1$, or
  - $\gamma = 1$, $n = 2$ and $Cycles(Y, \mathcal{G}^s \backslash \{X\}) \neq \emptyset$.

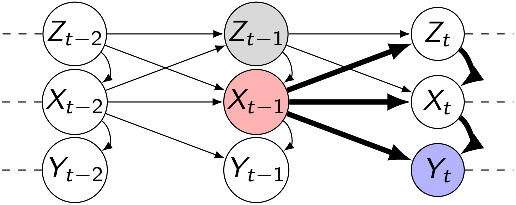
[9]A. Meynaoui et al., *Identifiability of total effects from abstractions of time series causal graphs*, submitted

Causal reasoning

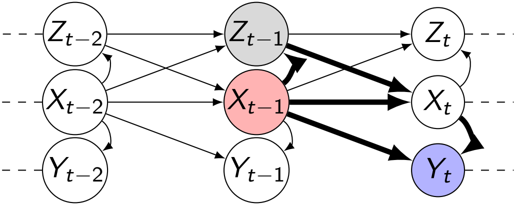
Identifiability in SCG: non identifiable example 1



An SCG $\mathcal{G}_1^s$ and the total effect $P(y_t|do(x_{t-1}))$.



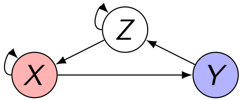
An FTCG compatible with the SCG $\mathcal{G}_1^s$.



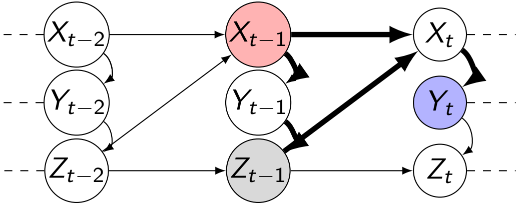
Another FTCG compatible with the SCG $\mathcal{G}_1^s$.

Causal reasoning

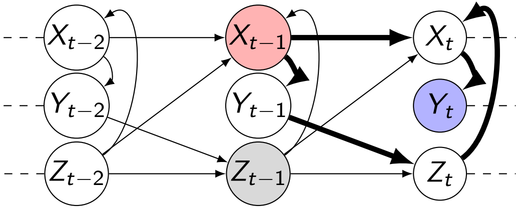
Identifiability in SCG: non identifiable example 2



An SCG $\mathcal{G}_2^s$ and the total effect $P(y_t|do(x_{t-1}))$.



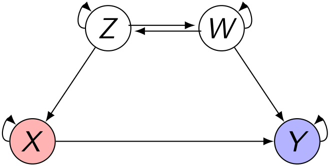
An FTCG compatible with the SCG $\mathcal{G}_2^s$.



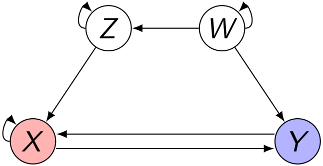
Another FTCG compatible with the SCG $\mathcal{G}_2^s$.

Causal reasoning

Identifiability in SCG: identifiable examples



$P(y_t | do(x_{t-\gamma}))$



$P(y_t | do(x_{t-1}))$

Outline

Conclusion and perspectives

- Which causal graph do we want to infer?
- The representation of time series is essential (windows - lags)
- Many families to discover causal graph for time series (also score-based, logic-based, topology-based, difference-based)
- Hybrid methods can take benefit of several worlds

Conclusion and perspectives

- Which causal graph do we want to infer?
- The representation of time series is essential (windows - lags)
- Many families to discover causal graph for time series (also score-based, logic-based, topology-based, difference-based)
- Hybrid methods can take benefit of several worlds

- Stationarity assumption for time series? Ongoing work with L. Zan
- Estimation of total effects?
- Counterfactual analysis?

References (Causal discovery)

- P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. MIT press, 2000.
- C. K. Assaad, E. Devijver, and E. Gaussier. *A Mixed Noise and Constraint Based Approach to Causal Inference in Time Series*, ECMLPKDD 2021
- C. K. Assaad, E. Devijver, and E. Gaussier. *Survey and evaluation of causal discovery methods for time series*. JAIR, 73, 2022.
- C. K. Assaad, E. Devijver, and E. Gaussier. *Causal Discovery of Extended Summary Graphs in Time Series*, UAI 2022
- L. Zan, A. Meynaoui, C.K. Assaad, E. Devijver, E. Gaussier, *A Conditional Mutual Information Estimator for Mixed Data and an Associated Conditional Independence Test*, Entropy 2022

References (Causal reasoning)

- Shpitser, I. and Pearl, J., *Complete identification methods for the causal hierarchy*, JMLR 2008
- G. Blondel, M. Arias, R. Gavalda, *Identifiability and transportability in dynamic causal networks*, International Journal of Data Science and Analytics, 2016
- Perkovic, E., *Identifying causal effects in maximally oriented partially directed acyclic graphs*, UAI 2020
- A. Meynaoui, C. K. Assaad, E. Devijver, E. Gaussier, G. Gössler, *Identifiability of total effects from abstractions of time series causal graphs*, submitted