Causal discovery for linear cyclic models with latent variables

Antti Hyttinen[1], Frederick Eberhardt[2],
and Patrik O. Hoyer[1,3]

[1] HIIT / Dept. of Computer Science, University of Helsinki
[2] Dept. of Philosophy, Washington University in St. Louis
[3] CSAIL, Massachusetts Institute of Technology

PGM 2010
September 15th, 2010

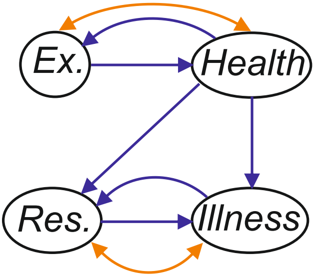
Outline

Example 1: Health, Exercise, Resistance, Illness

How to understand the interaction between Health, Exercise, Resistance, and Illness?

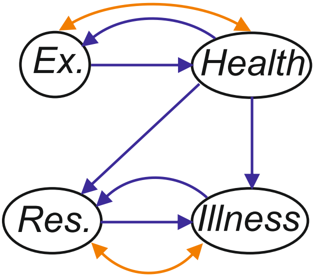
Example 1: Health, Exercise, Resistance, Illness

How to understand the interaction between Health, Exercise, Resistance, and Illness?



Example 1: Health, Exercise, Resistance, Illness

How to understand the interaction between Health, Exercise, Resistance, and Illness?



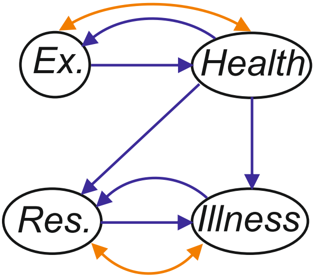
✓ Cycles

✓ Latent confounders

✓ Some experiments possible

Example 1: Health, Exercise, Resistance, Illness

How to understand the interaction between Health, Exercise, Resistance, and Illness?



✓ Cycles

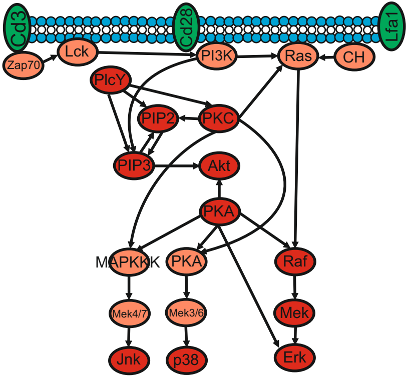
✓ Latent confounders

✓ Some experiments possible

Challenge: Select the experiments for learning the relationships as completely and accurately as possible!

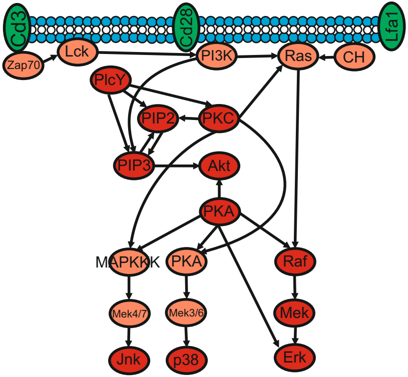
Example 2: Flow cytometry

How to understand cell signaling inside human T-cells?



Example 2: Flow cytometry

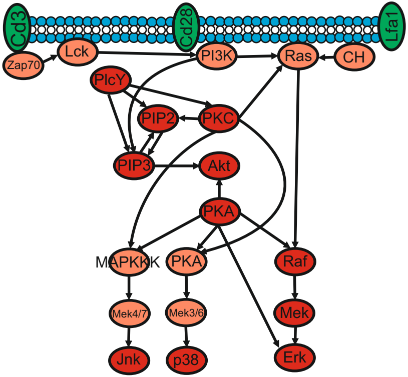
How to understand cell signaling inside human T-cells?



✓ Cycles
✓ Latent confounders
✓ Experiments available

Example 2: Flow cytometry

How to understand cell signaling inside human T-cells?



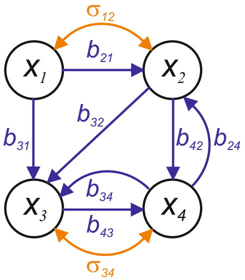
✓ Cycles
✓ Latent confounders
✓ Experiments available

Challenge: Given a set of experiments, learn as much of the structure as possible!

Earlier work

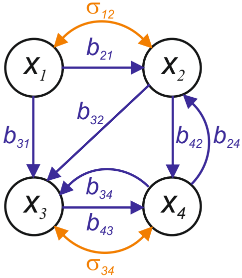
	exp.	latents	cycles	Cont. / Discr.
Spirtes et al. (FCI) (1993)		✓		C/D
Richardson (CCD) (1996)			✓	C/D
Schmidt and Murphy (2009)	✓		✓	D
Itani, Sachs et al. (2010)	✓		✓	D
Eberhardt, Hoyer, Scheines (2010)	✓	✓	✓	**C (lin.)**

An example of a linear cyclic model w. latents



$$
\begin{aligned}
x_1 &:= && && e_1\\
x_2 &:= b_{21}x_1 && && + b_{24}x_4 + e_2\\
x_3 &:= b_{31}x_1 + b_{32}x_2 && && + b_{34}x_4 + e_3\\
x_4 &:= && b_{42}x_2 + b_{43}x_3 && + e_4
\end{aligned}
$$

An example of a linear cyclic model w. latents

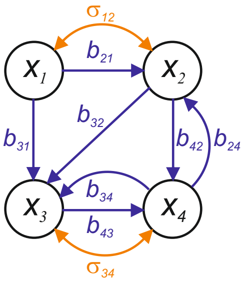


$$
\begin{aligned}
x_1 &:= && && && e_1 \\
x_2 &:= && b_{21}x_1 && && + b_{24}x_4 + e_2 \\
x_3 &:= && b_{31}x_1 + b_{32}x_2 && && + b_{34}x_4 + e_3 \\
x_4 &:= && && b_{42}x_2 + b_{43}x_3 && + e_4
\end{aligned}
$$

$$
\mathbf{x} := \mathbf{Bx} + \mathbf{e}
$$

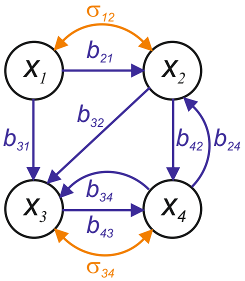
$$
\mathbf{B} = \begin{pmatrix} \mathbf{0} & 0 & 0 & 0 \\ b_{21} & \mathbf{0} & 0 & b_{24} \\ b_{31} & b_{32} & \mathbf{0} & b_{34} \\ 0 & b_{42} & b_{43} & \mathbf{0} \end{pmatrix}, \quad \mathbf{\Sigma}_e = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 \\ \sigma_{12} & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & \sigma_{34} \\ 0 & 0 & \sigma_{34} & \sigma_4^2 \end{pmatrix}
$$

Behaviour at equilibrium



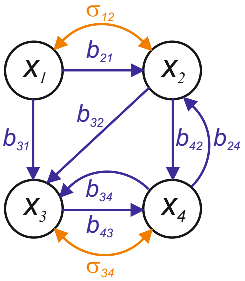
$$\mathbf{x} \;:=\; \mathbf{B}\mathbf{x} + \mathbf{e}$$

Behaviour at equilibrium



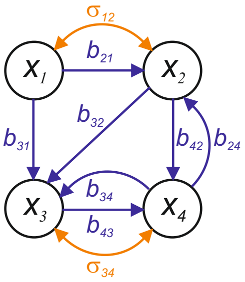
$$\mathbf{x}_t \quad = \quad \mathbf{B}\mathbf{x}_{t-1} + \mathbf{e} \leftarrow \text{background conditions invariant}$$

Behaviour at equilibrium



$$
\begin{aligned}
\mathbf{x}_t &= \mathbf{B}\mathbf{x}_{t-1} + \mathbf{e} \leftarrow \text{background conditions invariant} \\
\mathbf{x}_\infty &= \mathbf{B}^\infty \mathbf{x}_0 + (\mathbf{I} + \mathbf{B} + \mathbf{B}^2 + \cdots)\mathbf{e}
\end{aligned}
$$

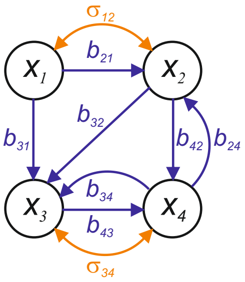
Behaviour at equilibrium



Stability Assumption: The model and all possible manipulated models are assumed stable: absolute values of the eigenvalues of **B** and manipulated **B**s must all be less than 1.

$$
\begin{aligned}
\mathbf{x}_t &= \mathbf{B}\mathbf{x}_{t-1} + \mathbf{e} \leftarrow \text{background conditions invariant} \\
\mathbf{x}_\infty &= \mathbf{B}^\infty \mathbf{x}_0 + (\mathbf{I} + \mathbf{B} + \mathbf{B}^2 + \cdots)\mathbf{e} \\
\mathbf{x}_\infty &= (\mathbf{I} - \mathbf{B})^{-1}\mathbf{e}
\end{aligned}
$$

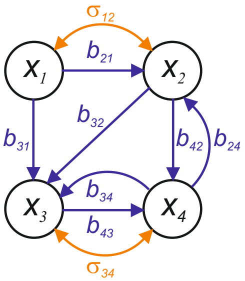
Behaviour at equilibrium



Stability Assumption: The model and all possible manipulated models are assumed stable: absolute values of the eigenvalues of $\mathbf{B}$ and manipulated $\mathbf{B}$s must all be less than 1.

$$\begin{aligned}
\mathbf{x}_t &= \mathbf{B}\mathbf{x}_{t-1} + \mathbf{e} \leftarrow \text{background conditions invariant} \\
\mathbf{x}_\infty &= \mathbf{B}^\infty \mathbf{x}_0 + (\mathbf{I} + \mathbf{B} + \mathbf{B}^2 + \cdots)\mathbf{e} \\
\mathbf{x}_\infty &= (\mathbf{I} - \mathbf{B})^{-1}\mathbf{e} \\
\mathrm{Cov}(\mathbf{x}_\infty) &= (\mathbf{I} - \mathbf{B})^{-1}\mathbf{\Sigma}_\mathbf{e}(\mathbf{I} - \mathbf{B})^{-T}
\end{aligned}$$

Behaviour at equilibrium



Stability Assumption: The model and all possible manipulated models are assumed stable: absolute values of the eigenvalues of **B** and manipulated **B**s must all be less than 1.

$$\mathbf{x} := \mathbf{B}\mathbf{x} + \mathbf{e}$$

$$\mathrm{Cov}(\mathbf{x}) = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Sigma}_{\mathbf{e}}(\mathbf{I} - \mathbf{B})^{-T}$$

Learning Method

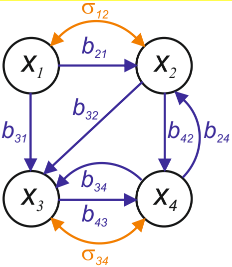
1. Input data from several randomized experiments, each intervening on possibly several different variables.

2. For each experiment:

   2a. Estimate the covariance matrix $\mathbf{C}_x^k$.

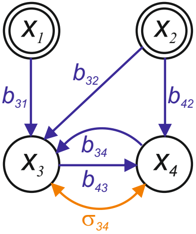
5. Output the estimated model $(\mathbf{B}, \mathbf{\Sigma_e})$.

Experimental effects



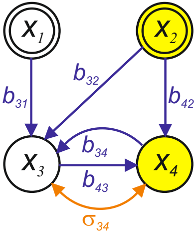
$$
\begin{aligned}
x_1 &:= && && && && e_1\\
x_2 &:= b_{21}x_1 && && && + b_{24}x_4 && + e_2\\
x_3 &:= b_{31}x_1 && + b_{32}x_2 && && + b_{34}x_4 && + e_3\\
x_4 &:= && b_{42}x_2 && + b_{43}x_3 && && + e_4
\end{aligned}
$$

Experimental effects



$$
\begin{aligned}
x_1 &:= \text{Randomized value}\\
x_2 &:= \text{Randomized value}\\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3\\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

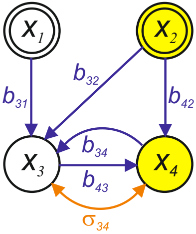
Experimental effects



$$
\begin{aligned}
x_1 &:= \text{Randomized value} \\
x_2 &:= \text{Randomized value} \\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3 \\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) = \text{Regression coefficient of } x_2 \text{ on } x_4$$

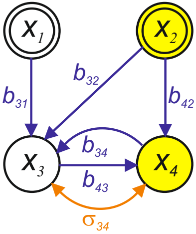
Experimental effects



$$
\begin{aligned}
x_1 &:= \text{Randomized value} \\
x_2 &:= \text{Randomized value} \\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3 \\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$
\begin{aligned}
t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) &= \text{Regression coefficient of } x_2 \text{ on } x_4 \\
&= \text{Sum-product of open paths from } x_2 \text{ to } x_4
\end{aligned}
$$

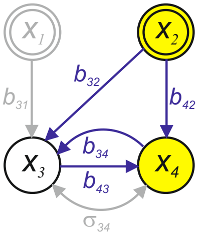
Experimental effects



$$x_1 := \text{Randomized value}$$
$$x_2 := \text{Randomized value}$$
$$x_3 := b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3$$
$$x_4 := b_{42}x_2 + b_{43}x_3 + e_4$$

$$t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) = \text{Regression coefficient of } x_2 \text{ on } x_4$$
$$= \text{Sum-product of open paths from } x_2 \text{ to } x_4$$
$$= b_{42} + b_{43}b_{32} + b_{43}b_{34}b_{42} + b_{43}b_{34}b_{43}b_{32} + \ldots$$

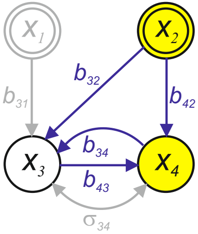
Experimental effects



$$
\begin{aligned}
x_1 &:= \text{Randomized value} \\
x_2 &:= \text{Randomized value} \\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3 \\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$
\begin{aligned}
t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) &= \text{Regression coefficient of } x_2 \text{ on } x_4 \\
&= \text{Sum-product of open paths from } x_2 \text{ to } x_4 \\
&= b_{42} + b_{43}b_{32} + b_{43}b_{34}b_{42} + b_{43}b_{34}b_{43}b_{32} +
\end{aligned}
$$

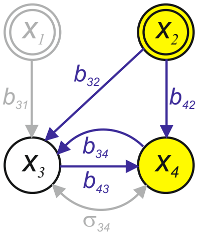
Experimental effects



$$
\begin{aligned}
x_1 &:= \text{Randomized value} \\
x_2 &:= \text{Randomized value} \\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3 \\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$
\begin{aligned}
t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) &= \text{Regression coefficient of } x_2 \text{ on } x_4 \\
&= \text{Sum-product of open paths from } x_2 \text{ to } x_4 \\
&= b_{42} + b_{43}b_{32} + b_{43}b_{34}b_{42} + b_{43}b_{34}b_{43}b_{32} + \\
&= (b_{42} + b_{43}b_{32})(1 + b_{43}b_{34} + b_{43}^2 b_{34}^2 + \cdots)
\end{aligned}
$$

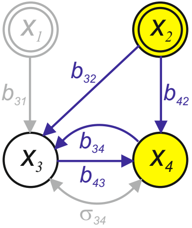
Experimental effects



Stability Assumption: The model and all possible manipulated models are assumed stable: absolute values of the eigenvalues of **B** and manipulated **B**s must all be less than 1.

$$
\begin{aligned}
t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) &= \text{Regression coefficient of } x_2 \text{ on } x_4 \\
&= \text{Sum-product of open paths from } x_2 \text{ to } x_4 \\
&= b_{42} + b_{43}b_{32} + b_{43}b_{34}b_{42} + b_{43}b_{34}b_{43}b_{32} + \\
&= (b_{42} + b_{43}b_{32})(1 + b_{43}b_{34} + b_{43}^2 b_{34}^2 + \cdots) \\
&= \frac{b_{42} + b_{43}b_{32}}{1 - b_{43}b_{34}}
\end{aligned}
$$

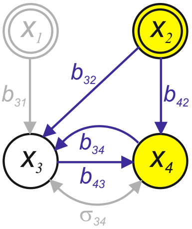
Linear equations



$$
\begin{aligned}
x_1 &:= \text{Randomized value} \\
x_2 &:= \text{Randomized value} \\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3 \\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) = \frac{b_{42} + b_{43}b_{32}}{1 - b_{43}b_{34}}$$

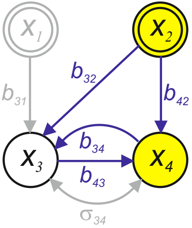
Linear equations



$$
\begin{aligned}
x_1 &:= \text{Randomized value} \\
x_2 &:= \text{Randomized value} \\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3 \\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$
\begin{aligned}
t(x_2 \rightsquigarrow x_4 || \{x_1, x_2\}) &= \frac{b_{42} + b_{43}b_{32}}{1 - b_{43}b_{34}} \\
&= b_{42} + \frac{b_{32} + b_{34}b_{42}}{1 - b_{43}b_{34}} b_{43}
\end{aligned}
$$

Linear equations



$$
\begin{aligned}
x_1 &:= \text{Randomized value}\\
x_2 &:= \text{Randomized value}\\
x_3 &:= b_{31}x_1 + b_{32}x_2 + b_{34}x_4 + e_3\\
x_4 &:= b_{42}x_2 + b_{43}x_3 + e_4
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{t}(\mathbf{x_2} \rightsquigarrow \mathbf{x_4} || \{\mathbf{x_1}, \mathbf{x_2}\}) &= \frac{b_{42} + b_{43}b_{32}}{1 - b_{43}b_{34}}\\
&= b_{42} + \frac{b_{32} + b_{34}b_{42}}{1 - b_{43}b_{34}} b_{43}\\
&\mathbf{=} \mathbf{b_{42}} + \mathbf{t}(\mathbf{x_2} \rightsquigarrow \mathbf{x_3} || \{\mathbf{x_1}, \mathbf{x_2}\})\mathbf{b_{43}}
\end{aligned}
$$

Learning Method

1. Input data from several randomized experiments, each intervening on possibly several different variables.

2. For each experiment:

    2a. Estimate the covariance matrix $\mathbf{C}_x^k$.

5. Output the estimated model $(\mathbf{B}, \mathbf{\Sigma_e})$.

Learning Method

1. Input data from several randomized experiments, each intervening on possibly several different variables.

2. For each experiment:

    2a. Estimate the covariance matrix $\mathbf{C}_x^k$.

    2b. Estimate the experimental effects $t(\bullet \rightsquigarrow \bullet || \mathbf{J}_k)$.

    2c. Form linear constraint equations on b_{ij}.

3. Solve the linear constraint equations for b_{ij} to get $\mathbf{B}$.

5. Output the estimated model $(\mathbf{B}, \mathbf{\Sigma_e})$.

Learning Method

1. Input data from several randomized experiments, each intervening on possibly several different variables.

2. For each experiment:

   2a. Estimate the covariance matrix $\mathbf{C}_x^k$.

   2b. Estimate the experimental effects $t(\bullet\rightsquigarrow\bullet||\mathbf{J}_k)$.

   2c. Form linear constraint equations on b_{ij}.

3. Solve the linear constraint equations for b_{ij} to get $\mathbf{B}$.

4. Given $\mathbf{B}$, calculate the covariances of the error terms σ_{ij} in $\mathbf{\Sigma_e}$ from an experiment where both x_i and x_j are observed.

5. Output the estimated model $(\mathbf{B}, \mathbf{\Sigma_e})$.

Theoretical Results

Theorem (Identifiability)

Given a sequence of experiments the model $(\mathbf{B}, \mathbf{\Sigma_e})$ *is fully identified by the method if and only if for each* **ordered** *pair of variables* (x_i, x_j)*:*

- *there is an experiment where* x_i *is* **intervened on** *and* x_j *is* **observed** *(Pair Condition), and*
- *another experiment where both* x_i *and* x_j *are* **observed** *(Covariance Condition).*

$$\mathrm{PC}: \begin{pmatrix} \cdot & \checkmark & \checkmark & \checkmark \\ \checkmark & \cdot & \checkmark & \checkmark \\ \checkmark & \checkmark & \cdot & \checkmark \\ \checkmark & \checkmark & \checkmark & \cdot \end{pmatrix}, \quad \mathrm{COV}: \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \checkmark & \cdot & \cdot & \cdot \\ \checkmark & \checkmark & \cdot & \cdot \\ \checkmark & \checkmark & \checkmark & \cdot \end{pmatrix}$$

Theoretical Results

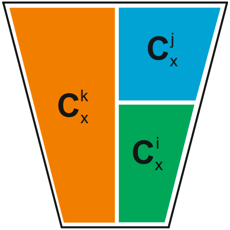
Theorem (Identifiability)

Given a sequence of experiments the model $(\mathbf{B}, \mathbf{\Sigma_e})$ *is fully identified by the method if and only if for each* **ordered** *pair of variables* (x_i, x_j)*:*

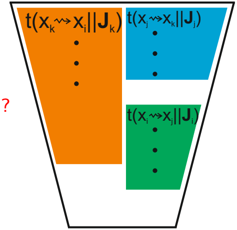
- *there is an experiment where* x_i *is* **intervened on** *and* x_j *is* **observed** *(Pair Condition), and*
- *another experiment where both* x_i *and* x_j *are* **observed** *(Covariance Condition).*

1. Singleton interventions $\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}$.
2. All but one interventions $\{x_1, x_2, x_3\}$, $\{x_1, x_2, x_4\}$, $\{x_1, x_3, x_4\}$, $\{x_2, x_3, x_4\}$ and passive observational dataset.
3. Something else like $\{x_1\}$, $\{x_2, x_3\}$, $\{x_3, x_4\}$ and $\{x_2, x_4\}$.

Completeness

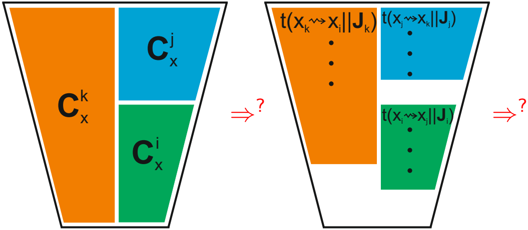


$\Rightarrow^?$



$\Rightarrow^?$ Linear equations on b_{ij}

Completeness



Linear equations on b_{ij}

Theorem (Completeness)

Given the data covariance matrices from a set of experiments, for determining the direct effects b_{ji}, the identifiability condition (Pair Condition) of the procedure is **necessary** *for any method.*

Assuming faithfulness

Theorem (Identifiability)

Given a sequence of experiments the model $(\mathbf{B}, \mathbf{\Sigma_e})$ *is fully identified by the method if and only if for each* **ordered** *pair of variables* (x_i, x_j)*:*

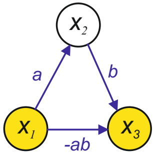
- *there is an experiment where* x_i *is* **intervened on** *and* x_j *is* **observed** *(Pair Condition), and*
- *another experiment where both* x_i *and* x_j *are* **observed** *(Covariance Condition).*

$$\mathrm{PC}: \begin{pmatrix} \cdot & \checkmark & \checkmark & \checkmark \\ \checkmark & \cdot & \checkmark & \checkmark \\ \checkmark & \checkmark & \cdot & \checkmark \\ \checkmark & \checkmark & \checkmark & \cdot \end{pmatrix}, \quad \mathrm{COV}: \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \checkmark & \cdot & \cdot & \cdot \\ \checkmark & \checkmark & \cdot & \cdot \\ \checkmark & \checkmark & \checkmark & \cdot \end{pmatrix}$$

Assuming faithfulness

Faithfulness in linear models

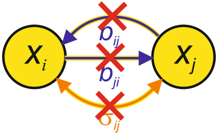
Any independence relation between variables is not the result of several exactly cancelling pathways.



Assuming faithfulness

Faithfulness in linear models

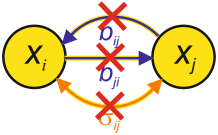
Any independence relation between variables is not the result of several exactly cancelling pathways.



Assuming faithfulness

Faithfulness in linear models

Any independence relation between variables is not the result of several exactly cancelling pathways.

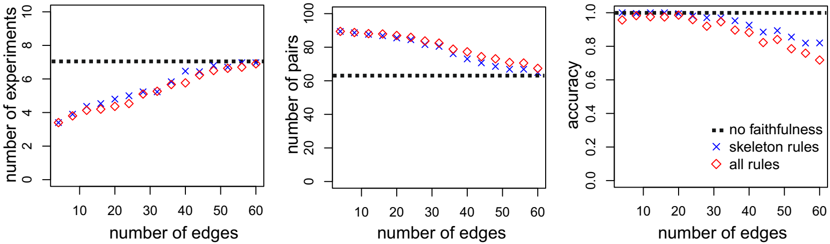


For every experimental dataset we can

1. Run a PC-type of search for finding independencies.
2. Add equations $b_{ji} = 0$ for any independencies found, not considering the edges that were broken by intervention.
3. Apply any valid orientation rules.
4. Take into account the additional structure found when selecting the next experiment (by maximizing the number of pairs for which the pair condition is satisfied).

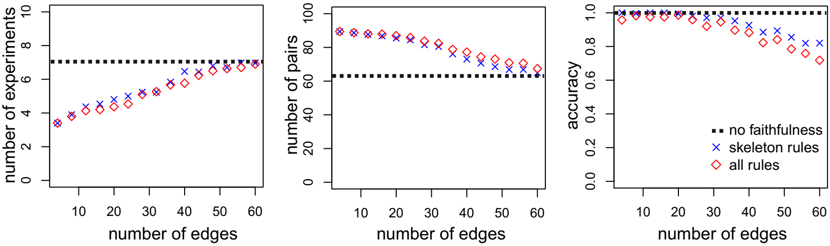
Test results

10 variables, max. 3 variables intervened



Test results

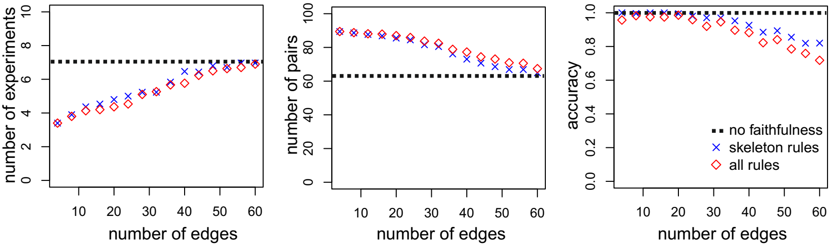
10 variables, max. 3 variables intervened



1. Fewer experiments are needed with sparse graphs.

Test results

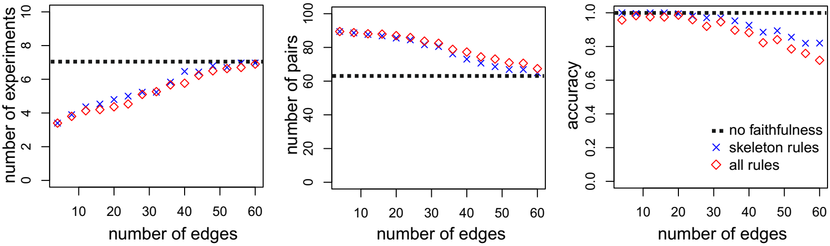
10 variables, max. 3 variables intervened



1. Fewer experiments are needed with sparse graphs.
2. More structure is discovered earlier on (e.g. after 3 experiments) with sparse graphs.

Test results

10 variables, max. 3 variables intervened



1. Fewer experiments are needed with sparse graphs.
2. More structure is discovered earlier on (e.g. after 3 experiments) with sparse graphs.
3. With denser graphs the accuracy gets worse.

Conclusion

Summary:

1. Revised the approach of Eberhardt, Hoyer, Scheines (2010) for learning linear cyclic models with latent variables.
2. Updated the identifiability condition.
3. Showed that the procedure is complete.
4. Incorporated the faithfulness assumption into the procedure.

Conclusion

Summary:

1. Revised the approach of Eberhardt, Hoyer, Scheines (2010) for learning linear cyclic models with latent variables.
2. Updated the identifiability condition.
3. Showed that the procedure is complete.
4. Incorporated the faithfulness assumption into the procedure.

Additionally in the poster:

1. A closer look at the linear equations.
2. Underdetermination and reasons behind completeness.
3. Faithfulness orientation rules and their justification.
4. Simulation & real world test results.

Conclusion

Summary:

1. Revised the approach of Eberhardt, Hoyer, Scheines (2010) for learning linear cyclic models with latent variables.
2. Updated the identifiability condition.
3. Showed that the procedure is complete.
4. Incorporated the faithfulness assumption into the procedure.

Additionally in the poster:

1. A closer look at the linear equations.
2. Underdetermination and reasons behind completeness.
3. Faithfulness orientation rules and their justification.
4. Simulation & real world test results.

Thank you!