Causal Discovery for Linear Cyclic Models with Latent Variables

Antti Hyttinen ★ Frederick Eberhardt ★ Patrik O. Hoyer

HIIT ★ University of Helsinki ★ Washington University in St. Louis ★ MIT

Linear Cyclic Model with Latent Variables

\[
B = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

\[
\Sigma = \begin{pmatrix}
\sigma_{11}^2 & 0 & 0 & 0 \\
0 & \sigma_{22}^2 & 0 & 0 \\
0 & 0 & \sigma_{33}^2 & 0 \\
0 & 0 & 0 & \sigma_{44}^2 \\
\end{pmatrix}
\]

\[
x \rightarrow Bx + e \leftrightarrow
\]

\* The model is parametrized by \(B\) and \(\text{cov}(e) = \Sigma\).

\* Behaviour at equilibrium:

\[
x_0 \Rightarrow Bx_0 + e \rightarrow \text{background conditions invariant}
\]

\[
x_{\infty} \Rightarrow B^{-1}x_{\infty} + (I + B^{-1})^\infty e
\]

\[
e \sim N(0, \Sigma_e) \Rightarrow x_{\infty} \sim N(0, (I - B^{-1})^\infty \Sigma_e(I - B^{-1})^\infty)
\]

Self cycles are not identifiable from equilibrium data, so assuming \(v_i = b_i = 0\).

The model and all possible manipulated models are assumed stable: absolute of the eigenvalues of \(B\) and \(B^{-1}\) must all be less than 1.

Interventions & Experimental Effects

\[
B^\ast = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

\[
\Sigma^\ast = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

\[
C_i^\ast = (I - B)^{-1}\Sigma^\ast (I - B^{-1})^{-1} - \text{independent randomizations}
\]

Given a sequence of experiments the model \((B, \Sigma_e)\) 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 \(B\), and \(x_j\) is observed \(C_i^\ast\) (Pair Condition), and

\* Another experiment where both \(x_i\) and \(x_j\) are observed \(C_i^\ast\) (Covariance Condition).

Identifiability theorem

\[
\begin{pmatrix}
t(x_{\rightarrow}x_{\leftarrow}\{x_1, x_2\}) \\
t(x_{\rightarrow}x_{\leftarrow}\{x_1, x_3\}) \\
t(x_{\rightarrow}x_{\leftarrow}\{x_2, x_3\}) \\
\end{pmatrix}
\]

Experimental effects between observed

\[
t(x_{\rightarrow}x_{\leftarrow}\{x_1, x_2\})
\]

Regression coefficient of \(x_1\) on \(x_2\)

2. Which parameters are identified in the general case?

\[
B = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

\[
\Sigma_e = \begin{pmatrix}
\sigma_{11}^2 & 0 & 0 & 0 \\
0 & \sigma_{22}^2 & 0 & 0 \\
0 & 0 & \sigma_{33}^2 & 0 \\
0 & 0 & 0 & \sigma_{44}^2 \\
\end{pmatrix}
\]

\* Generally, for identifying \(b_{ij}\), pair condition must be satisfied for all pairs \((i, j)\).

\* If pair condition for \((i, j)\) is not satisfied, then \(b_{ij}\) is never identified.

Identifiability of \((\Sigma, \Sigma_e)\)

Selecting the experiment in such a way that the model is learned accurately with the fewest number of experiments.

\* Fewer experiments are needed with sparse graphs.

\* More structure is discovered earlier on with sparse graphs.

\* With denser graphs the accuracy gets worse.

Method

1. Input covariance matrices \(C_1^\ast, \ldots, C_n^\ast\).

2. Form linear equations

\[
\begin{pmatrix}
k_1 & k_2 \\
\vdots & \vdots \\
k_{n-1} & k_n \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
t(x_{\rightarrow}x_{\leftarrow}\{x_1, x_2\}) \\
t(x_{\rightarrow}x_{\leftarrow}\{x_1, x_3\}) \\
t(x_{\rightarrow}x_{\leftarrow}\{x_2, x_3\}) \\
\end{pmatrix}
\]

3. Solve for direct effects \(b_{ij}\).

4. Get the covariances of the error terms \(\sigma_{ij}\) from an experiment where both \(x_i\) and \(x_j\) are observed with the formula

\[
\sigma_{ij} = \text{cov}(x_i, x_j) + \text{cov}(x_i, \epsilon) + \text{cov}(x_j, \epsilon)
\]

Assuming Faithfulness

Faithfulness in linear models

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

For every experimental dataset

1. Run a search for finding independencies. Add constraint equations from skeleton rule.

2. Add more constraint equations from orientation rules.

3. Take into account the additional structure found when selecting the next experiment.

Experiment Selection

1. Select the experiment that satisfies the pair condition for most new pairs.

2. If any parameters are identified, consider the pair condition for the corresponding pairs as satisfied.

Test Results

Sachs et al. Flow Cytometry Data

Learning as much of the structure as possible given only 5 experiments, intervening on \{X, Y, Z\} and \{X, Y, Z\}.