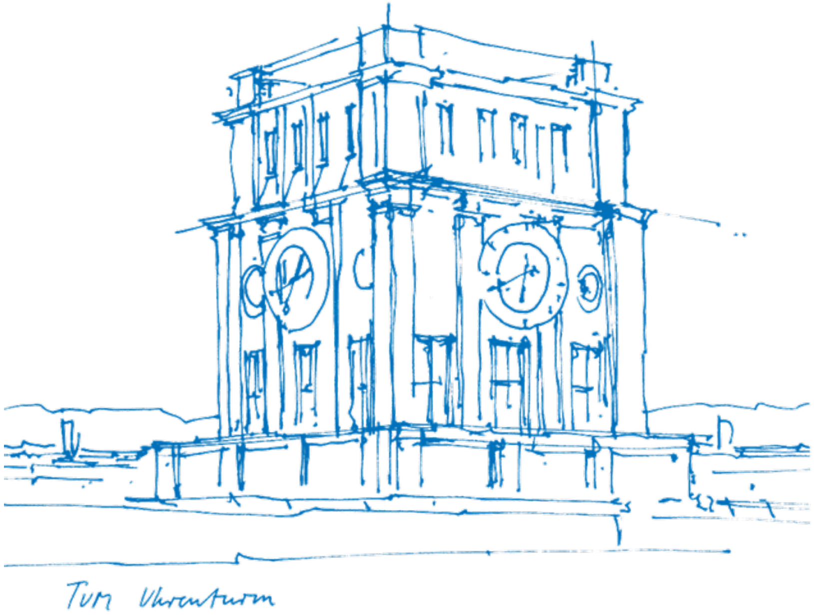
Half-Trek Criterion for Identifiability of Latent Variable Models

Mathias Drton

Department of Mathematics

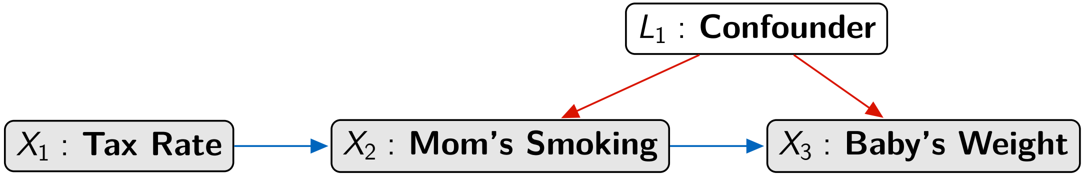
Technical University of Munich (TUM)

(joint work with Rina Foygel Barber, Nils Sturma, Luca Weihs)



Linear Structural Equation/Causal Models

Each model is induced by a directed graph:



Linear structural equations:

$$
\begin{aligned}
X_1 &= \lambda_{01} && && + \varepsilon_1, \\
X_2 &= \lambda_{02} + \lambda_{12}X_1 &&+ \gamma_2 L_1 &&+ \varepsilon_2, \\
X_3 &= \lambda_{03} + \lambda_{23}X_2 &&+ \gamma_3 L_1 &&+ \varepsilon_3, \\
L_1 &= \lambda_{0u} && && + \varepsilon_\ell.
\end{aligned}
$$

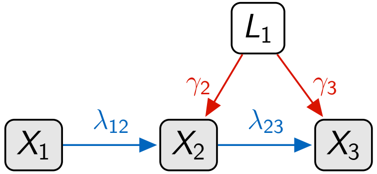
Independent errors:

$$\varepsilon_1 \perp\!\!\!\perp \varepsilon_2 \perp\!\!\!\perp \varepsilon_3 \perp\!\!\!\perp \varepsilon_\ell$$

$$\mathrm{Var}[\varepsilon_v] = \omega_v < \infty$$

Topic of the talk: If L_1 is latent, can we recover the direct effects $(\lambda_{12}, \lambda_{23})$ from $\Sigma = \mathrm{Var}[X]$?

Example: Instrumental Variable Model



$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 0 & \lambda_{12} & 0 \\ 0 & 0 & \lambda_{23} \\ 0 & 0 & 0 \end{pmatrix}^T \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} + \begin{pmatrix} 0 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} L_1 + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

Observed covariance matrix:

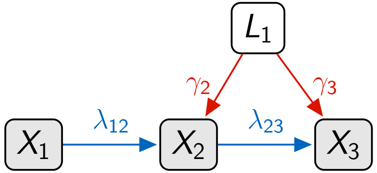
$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \cdot & \sigma_{22} & \sigma_{23} \\ \cdot & \cdot & \sigma_{33} \end{pmatrix} = \begin{pmatrix} \omega_1 & \boxed{\omega_1\lambda_{12}} & \boxed{\omega_1\lambda_{12}}\lambda_{23} \\ \cdot & \omega_2 + \gamma_2^2 + \omega_1\lambda_{12}^2 & \gamma_2\gamma_3 + \lambda_{23}\sigma_{22} \\ \cdot & \cdot & \omega_3 + \gamma_3^2 + 2\gamma_2\gamma_3\lambda_{23} + \lambda_{23}^2\sigma_{22} \end{pmatrix}$$

We see that

$$\lambda_{12} = \frac{\sigma_{12}}{\sigma_{11}} \qquad \text{with } \sigma_{11} > 0,$$

$$\lambda_{23} = \frac{\sigma_{13}}{\sigma_{12}} \qquad \text{with } \sigma_{12} = \omega_1\lambda_{12} \neq 0 \text{ `almost surely'}.$$

Example: Instrumental Variable Model



$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 0 & \lambda_{12} & 0 \\ 0 & 0 & \lambda_{23} \\ 0 & 0 & 0 \end{pmatrix}^T \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} + \begin{pmatrix} 0 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} L_1 + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

Observed covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \cdot & \sigma_{22} & \sigma_{23} \\ \cdot & \cdot & \sigma_{33} \end{pmatrix} = \begin{pmatrix} \omega_1 & \boxed{\omega_1\lambda_{12}} & \boxed{\omega_1\lambda_{12}}\lambda_{23} \\ \cdot & \omega_2 + \gamma_2^2 + \omega_1\lambda_{12}^2 & \gamma_2\gamma_3 + \lambda_{23}\sigma_{22} \\ \cdot & \cdot & \omega_3 + \gamma_3^2 + 2\gamma_2\gamma_3\lambda_{23} + \lambda_{23}^2\sigma_{22} \end{pmatrix}$$

We see that

$$\lambda_{12} = \frac{\sigma_{12}}{\sigma_{11}} \qquad \text{with } \sigma_{11} > 0,$$

$$\lambda_{23} = \frac{\sigma_{13}}{\sigma_{12}} \qquad \text{with } \sigma_{12} = \omega_1\lambda_{12} \neq 0 \ \text{'almost surely'}.$$

Questions?

Setup of the Paper

Variables:

Observed: $X = (X_v)_{v \in V}$ Latent: $L = (L_h)_{h \in \mathcal{L}}$

Graph:

Directed graph $G = (V \dot\cup \mathcal{L}, D)$ with directed cycles allowed

Latent-factor assumption:

All latent variables are latent factors $\equiv$ all nodes in $\mathcal{L}$ are source nodes of G.

Structural equation model:

$$X = \Lambda^\top X + \Gamma^\top L + \varepsilon$$

– all latent factors and error terms in (L, ε) are mutually independent, so $\Omega_{\text{diag}} = \text{Var}[\varepsilon] = \text{diag}(\omega_v : v \in V)$ diagonal, and $\text{Var}[L] = I$ without loss of generality.

– parameter matrices Λ and Γ are sparse and supported over edge set D.

Content of the Paper

Definition

Every latent-factor graph G yields a parametrization of the observed covariance matrix:

$$\phi_G : (\Lambda, \Gamma, \Omega_{\text{diag}}) \longmapsto \Sigma \equiv \text{Var}[X].$$

The model given by G is rationally identifiable if

$$\exists \text{ rational map } \psi_G \ : \quad \psi_G \circ \phi_G(\Lambda, \Gamma, \Omega_{\text{diag}}) = \Lambda \quad \text{for 'almost all'} \ (\Lambda, \Gamma, \Omega_{\text{diag}}).$$

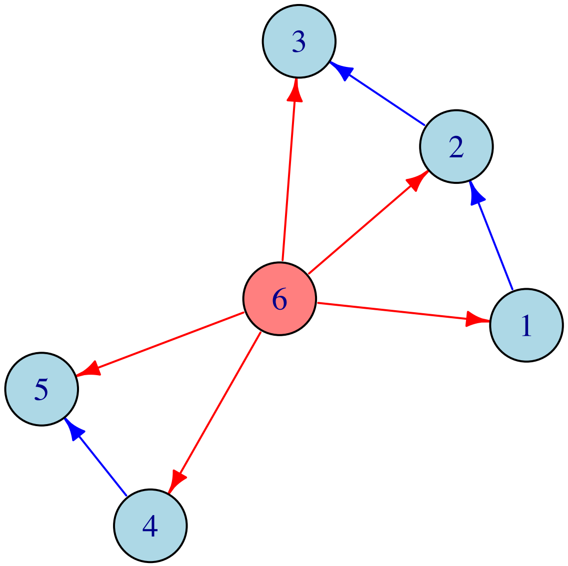
Main Contribution:

- Sufficient condition for rational identifiability.
- Recursive polynomial time algorithm.

  (caveat: polynomial time when bounding a matrix rank in a search step)
- Condition is not necessary but 'effective'; see simulations in paper.

Our Software: SEMID (R Package)

```
# Define graph
> Lambda = matrix(c(0, 1, 0, 0, 0, 0,
+                   0, 0, 1, 0, 0, 0,
+                   0, 0, 0, 0, 0, 0,
+                   0, 0, 0, 0, 1, 0,
+                   0, 0, 0, 0, 0, 0,
+                   1, 1, 1, 1, 1, 0),
+                 6, 6, byrow=TRUE)
> observedNodes = seq(1,5)
> latentNodes = c(6)
> g = LatentDigraph(Lambda, observedNodes, latentNodes)
> plot(g)
```

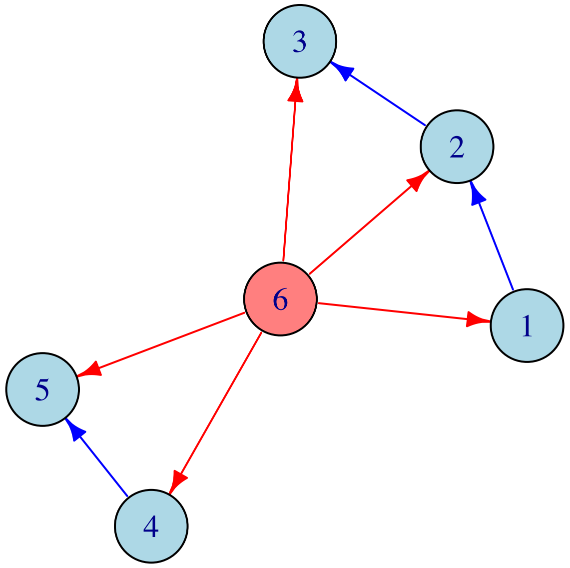


Our Software: SEMID (R Package)

```
> # Check identifiability
> res = lfhtcID(g)
> res
Call: lfhtcID(graph = g)

Latent Digraph Info
# observed nodes: 5
# latent nodes: 1
# total nr. of edges between observed nodes: 3

Generic Identifiability Summary
# nr. of edges between observed nodes shown gen. identifiable: 3
# gen. identifiable edges: 1->2, 2->3, 4->5
```



Latent and Observed Covariance Matrix

- Solving the model equation

$$X \;=\; \Lambda^\top X \;+\; \underbrace{\Gamma^\top L + \varepsilon}_{\text{unobserved}}$$

gives

$$X \;=\; (I-\Lambda)^{-\top}(\Gamma^\top L + \varepsilon).$$

- Latent covariance matrix

$$\Omega \;\equiv\; \mathrm{Var}[\Gamma^\top L + \varepsilon] \;=\; \mathrm{Var}[\varepsilon] + \Gamma^\top \mathrm{Var}[L]\Gamma \;=\; \Omega_{\text{diag}} + \Gamma^\top\Gamma.$$

Note that the matrix may be sparse and feature low-rank structure:

$$\Omega \;=\; \Omega_{\text{diag}} + \sum_{h\in\mathcal{L}} \gamma_h\gamma_h^\top \;=\; \text{diag} + \text{sum of sparse rank 1 matrices.}$$

- Observed covariance matrix

$$\Sigma = \mathrm{Var}[X] = (I-\Lambda)^{-\top}\Omega(I-\Lambda)^{-1}.$$

Using Algebraic Relations in Latent Covariance Matrix

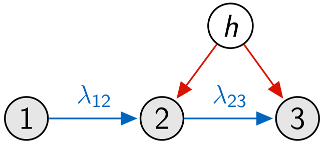
- Observe that

$$\Sigma = (I - \Lambda)^{-\top}\Omega(I - \Lambda)^{-1} \iff \boxed{\Omega \;=\; (I - \Lambda)^T\Sigma(I - \Lambda)}$$

- Algebraic relations between entries of $\Omega = \Omega_{\text{diag}} + \Gamma^\top\Gamma$ yield relations between entries of Λ and Σ:

$$f(\Omega) \;=\; 0 \iff f((I - \Lambda)^T\Sigma(I - \Lambda)) = 0.$$

- <u>IV Example:</u>



$$\Omega = \begin{pmatrix} \omega_1 & 0 & \mathbf{0} \\ 0 & \omega_2 + \gamma_{h2}^2 & \gamma_{h2}\gamma_{h3} \\ \mathbf{0} & \gamma_{h2}\gamma_{h3} & \omega_3 + \gamma_{h3}^2 \end{pmatrix}$$

$$\left[(I - \Lambda)^T\Sigma(I - \Lambda)\right]_{13} = \sigma_{13} - \lambda_{23}\,\sigma_{12} = 0$$

- The problem may be solved via a Gröbner basis computation. . . on small scale.

Using Algebraic Relations in Latent Covariance Matrix

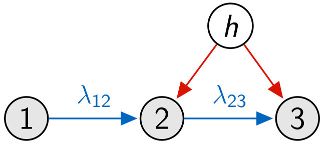
- Observe that

$$\Sigma = (I-\Lambda)^{-\top}\Omega(I-\Lambda)^{-1} \iff \boxed{\Omega \;=\; (I-\Lambda)^T\Sigma(I-\Lambda)}$$

- Algebraic relations between entries of $\Omega = \Omega_{\text{diag}} + \Gamma^\top\Gamma$ yield relations between entries of Λ and Σ:

$$f(\Omega) \;=\; 0 \iff f((I-\Lambda)^T\Sigma(I-\Lambda)) = 0.$$

- IV Example:



$$\Omega = \begin{pmatrix} \omega_1 & 0 & \mathbf{0} \\ 0 & \omega_2 + \gamma_{h2}^2 & \gamma_{h2}\gamma_{h3} \\ \mathbf{0} & \gamma_{h2}\gamma_{h3} & \omega_3 + \gamma_{h3}^2 \end{pmatrix}$$

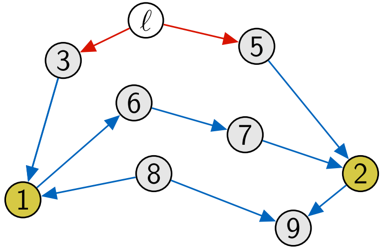
$$\left[(I-\Lambda)^T\Sigma(I-\Lambda)\right]_{13} = \sigma_{13} - \lambda_{23}\,\sigma_{12} = 0$$

- The problem may be solved via a Gröbner basis computation. . . on small scale.

Questions?

Combinatorics of Covariances

Example:

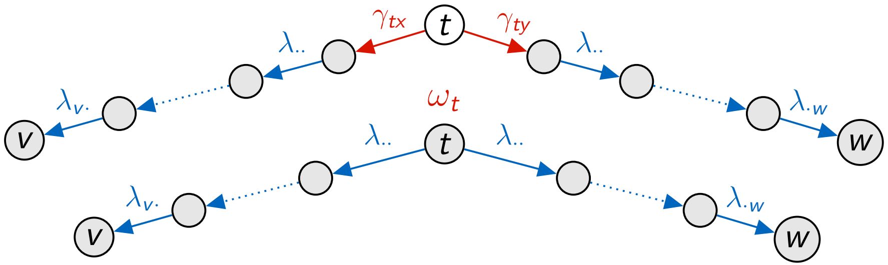


$$\Sigma_{12} = \lambda_{31}\gamma_{\ell 3}\gamma_{\ell 5}\lambda_{52} + \omega_1\lambda_{16}\lambda_{67}\lambda_{72}$$

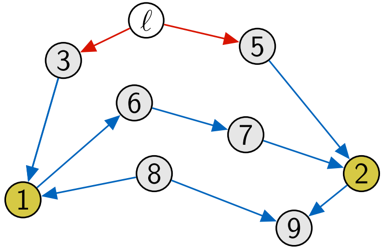
Combinatorics of Covariances

Trek rule (Wright, 1921, 1934)

$\Sigma_{vw} = \left[(I-\Lambda)^{-T}\Omega(I-\Lambda)^{-1}\right]_{vw}$ is sum of products along treks:



Example:



$$\Sigma_{12} = \lambda_{31}\gamma_{\ell 3}\gamma_{\ell 5}\lambda_{52} + \omega_1\lambda_{16}\lambda_{67}\lambda_{72}$$

Indeed, $(I-\Lambda)^{-1} = I + \Lambda + \Lambda^2 + \cdots + \Lambda^{m-1} + \cdots$ is a path matrix:

$$(I-\Lambda)^{-1}_{tw} = \sum_{P\in\mathcal{P}(t,w)} \prod_{x\to y\in P} \lambda_{xy}, \qquad \mathcal{P}(t,w) = \{\text{directed paths } t\to\cdots\to w\}.$$

Half-Trek Criterion (Foygel, Draisma, D., 2012)

- Starting point are the equations:

$$(I - \Lambda)^T \Sigma (I - \Lambda)_{vw} \;=\; 0, \quad v \leftarrow \circ \rightarrow w \notin G.$$

- For a sufficient condition, recursively visit nodes $v \in V$ and find linear equation systems:

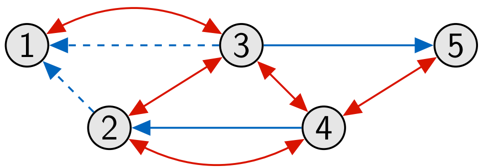
$$A(\Sigma) \cdot \Lambda_{\mathrm{pa}(v),v} = b(\Sigma).$$

Here, $\mathrm{pa}(v) \equiv \mathrm{pa}_V(v) = \{w \in V : w \rightarrow v \in G\}$.

- Check invertibility of $A(\Sigma)$ with the help of a 'Gessel-Viennot Lemma'.
- Computation: Find equation system in polynomial time via network-flow problems.

Subtlety: HTC from 2012 is based on latent projection to a mixed graph.

Half-Trek Criterion: Example



Solve at node #1: $(\lambda_{21}, \lambda_{31})$

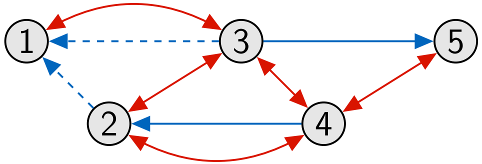
$$0 = \Omega_{51} = \left[(I-\Lambda)^T \Sigma (I-\Lambda)\right]_{51}$$

$$= \underbrace{\left(5 \overset{\sigma_{51}}{\cdots} 1\right)}_{\sigma_{51}} - \underbrace{\left(5 \overset{\sigma_{5w}}{\cdots} w \xrightarrow{\lambda_{w1}} 1\right)}_{\substack{\sigma_{52}\cdot\lambda_{21}\\ +\,\sigma_{53}\cdot\lambda_{31}}} - \underbrace{\left(5 \xleftarrow{\lambda_{u5}} u \overset{\sigma_{u1}}{\cdots} 1\right)}_{\lambda_{35}\cdot\sigma_{31}} + \underbrace{\left(5 \xleftarrow{\lambda_{u5}} u \overset{\sigma_{uw}}{\cdots} w \xrightarrow{\lambda_{w1}} 1\right)}_{\substack{\lambda_{35}\cdot\sigma_{32}\cdot\lambda_{21}\\ +\lambda_{35}\cdot\sigma_{33}\cdot\lambda_{31}}}$$

$$0 = \Omega_{21} = \left[(I-\Lambda)^T \Sigma (I-\Lambda)\right]_{21}$$

$$= \underbrace{\left(2 \overset{\sigma_{21}}{\cdots} 1\right)}_{\sigma_{21}} - \underbrace{\left(2 \overset{\sigma_{2w}}{\cdots} w \xrightarrow{\lambda_{w1}} 1\right)}_{\substack{\sigma_{22}\cdot\lambda_{21}\\ +\,\sigma_{23}\cdot\lambda_{31}}} - \underbrace{\left(2 \xleftarrow{\lambda_{u2}} u \overset{\sigma_{u1}}{\cdots} 1\right)}_{\cancel{\lambda_{42}\cdot\sigma_{41}}} + \underbrace{\left(2 \xleftarrow{\lambda_{u2}} u \overset{\sigma_{uw}}{\cdots} w \xrightarrow{\lambda_{w1}} 1\right)}_{\cancel{\substack{\lambda_{42}\cdot\sigma_{42}\cdot\lambda_{21}\\ \lambda_{42}\cdot\sigma_{43}\cdot\lambda_{31}}}}$$

Half-Trek Criterion: Example



Solve at node #1: $(\lambda_{21}, \lambda_{31})$

Coefficient matrix is generically invertible:

$$\begin{array}{c} \\ 5 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} \quad 2 \quad & \quad\quad 3 \end{array} \\ \begin{pmatrix} \sigma_{52} + \lambda_{35}\sigma_{32} & \sigma_{53} + \lambda_{35}\sigma_{33} \\ \sigma_{22} & \sigma_{23} \end{pmatrix} \end{array}$$

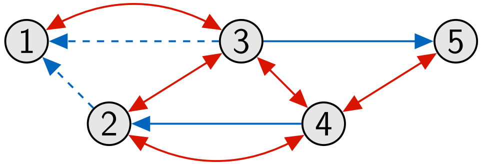
'Witness': A system of 'halftreks' without sided intersection

$$5 \leftarrow h \rightarrow 4 \rightarrow 2$$

$$2 \leftarrow h' \rightarrow 3$$

that connects $\{2, 5\}$ to $\text{pa}(1) = \{2, 3\}$.

Half-Trek Criterion: Example



Solve at node #1: $(\lambda_{21}, \lambda_{31})$

Coefficient matrix is generically invertible:

$$\begin{array}{c} \\ 5 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} \quad 2 \quad & \quad 3 \quad \end{array} \\ \begin{pmatrix} \sigma_{52} + \lambda_{35}\sigma_{32} & \sigma_{53} + \lambda_{35}\sigma_{33} \\ \sigma_{22} & \sigma_{23} \end{pmatrix} \end{array}$$

'Witness': A system of 'halftreks' without sided intersection

$$5 \leftarrow h \rightarrow 4 \rightarrow 2$$
$$2 \leftarrow h' \rightarrow 3$$

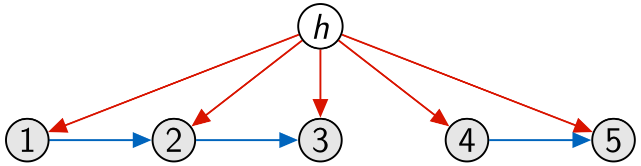
that connects $\{2, 5\}$ to $\text{pa}(1) = \{2, 3\}$.

Questions?

Latent Low Rank

- Lots of existing work is based on using zero entries in latent covariance matrix.
- However, the resulting methods cannot cover situations such as

  Example

  where the latent covariance matrix is dense:

$$\Omega \;=\; \Omega_{\text{diag}} + \gamma_h \gamma_h^\top = \text{diagonal} + \text{dense rank 1.}$$

- New paper: Generalize beyond zeros by exploiting

  latent low rank structure.

New Latent-Factor Half-Trek Criterion: Main Idea

- Digraph $(V + \mathcal{L}, D)$ with observed variables in V and latent variables in $\mathcal{L}$.
- Recursive search for linear equation systems that determine columns $\Lambda_{\text{pa}(v),v}$, $v \in V$.

New Latent-Factor Half-Trek Criterion: Main Idea

- Digraph $(V + \mathcal{L}, D)$ with observed variables in V and latent variables in $\mathcal{L}$.
- Recursive search for linear equation systems that determine columns $\Lambda_{\text{pa}(v),v}$, $v \in V$.
- To this end, we find a rank-deficient off-diagonal submatrix

$$\Omega_{Y,Z\cup\{v\}} = \left[(I-\Lambda)^T\Sigma(I-\Lambda)\right]_{Y,Z\cup\{v\}}.$$

  More precisely, the matrix is of rank $|Z|$ and such that

$$\left[(I-\Lambda)^T\Sigma(I-\Lambda)\right]_{Y,\{v\}} = \left[(I-\Lambda)^T\Sigma(I-\Lambda)\right]_{Y,Z} \cdot \psi \quad \text{for some } \psi \in \mathbb{R}^{|Z|}.$$

- If needed elements of Λ have already been solved for, we obtain a linear equation system

$$A(\Sigma) \cdot \Lambda_{\text{pa}(v),v} + B(\Sigma) \cdot \psi = c(\Sigma), \quad v \in V.$$

- Our combinatorial conditions ensure a generically unique solution.

Half-Treks

Definition

- A half-trek from node v to node w is a path of the form:

$$v \rightarrow x_1 \rightarrow \ldots \rightarrow x_\ell \rightarrow w \quad \text{or} \quad v \leftarrow \ell \rightarrow x_1 \rightarrow \ldots \rightarrow x_\ell \rightarrow w.$$

Relevance: Entries of $(I - \Lambda)^T \Sigma$ are sums over half-treks.

Half-Treks

Definition

- A half-trek from node v to node w is a path of the form:

$$v \to x_1 \to \dots \to x_\ell \to w \quad \text{or} \quad v \leftarrow \ell \to x_1 \to \dots \to x_\ell \to w.$$

Relevance: Entries of $(I - \Lambda)^T \Sigma$ are sums over half-treks.

- Each half-trek has two sides (away from source):

$$\text{Left}: \boxed{v} \to x_1 \to \dots \to x_\ell \to w, \quad \boxed{v \leftarrow \ell} \to x_1 \to \dots \to x_\ell \to w,$$
$$\text{Right}: \boxed{v \to x_1 \to \dots \to x_\ell \to w}, \quad v \leftarrow \boxed{\ell \to x_1 \to \dots \to x_\ell \to w}.$$

Half-Treks

Definition

- A half-trek from node v to node w is a path of the form:

$$v \to x_1 \to \dots \to x_\ell \to w \quad \text{or} \quad v \leftarrow \ell \to x_1 \to \dots \to x_\ell \to w.$$

Relevance: Entries of $(I - \Lambda)^T \Sigma$ are sums over half-treks.

- Each half-trek has two sides (away from source):

$$\begin{aligned} \text{Left}: &\ \boxed{v} \to x_1 \to \dots \to x_\ell \to w, \quad \boxed{v \leftarrow \ell} \to x_1 \to \dots \to x_\ell \to w, \\ \text{Right}: &\ \boxed{v \to x_1 \to \dots \to x_\ell \to w}, \quad v \leftarrow \boxed{\ell \to x_1 \to \dots \to x_\ell \to w}. \end{aligned}$$

- A system of half-treks has no sided intersection if neither the left sides nor the right sides intersect.

Latent-Factor Half-Trek Criterion (LF-HTC)

Definition

Let $v \in V$ and $Y, Z \subseteq V \setminus \{v\}$ and $H \subseteq \mathcal{L}$. Triple (Y, Z, H) satisfies latent-factor half-trek criterion for v if

1. $|Y| = |\mathrm{pa}(v)| + |H|$ and $|Z| = |H|$;
2. $Y \cap (Z \cup \{v\}) = \emptyset$;
3. $[\mathrm{pa}_{\mathcal{L}}(Y) \cap \mathrm{pa}_{\mathcal{L}}(Z \cup \{v\})] \subseteq H$;

By 1.-3.,

$$\Omega_{Y,Z\cup\{v\}} = \sum_{h\in H} (\Gamma_h^\top \Gamma_h)_{Y,Z\cup\{v\}} \quad \text{has rank } |H| = |Z|.$$

Latent-Factor Half-Trek Criterion (LF-HTC)

Definition

Let $v \in V$ and $Y, Z \subseteq V \setminus \{v\}$ and $H \subseteq \mathcal{L}$. Triple (Y, Z, H) satisfies latent-factor half-trek criterion for v if

1. $|Y| = |\text{pa}(v)| + |H|$ and $|Z| = |H|$;
2. $Y \cap (Z \cup \{v\}) = \emptyset$;
3. $[\text{pa}_{\mathcal{L}}(Y) \cap \text{pa}_{\mathcal{L}}(Z \cup \{v\})] \subseteq H$;
4. There is a system of half-treks from Y to $\text{pa}(v) \cup Z$ without sided intersection and all half-treks ending in Z have form $y \leftarrow \ell \rightarrow z$ for $\ell \in H$.

By 1.-3.,

$$\Omega_{Y,Z\cup\{v\}} = \sum_{h \in H} (\Gamma_h^\top \Gamma_h)_{Y,Z\cup\{v\}} \quad \text{has rank } |H| = |Z|.$$

Algorithm: Recursive Solving

Theorem (Latent-factor HTC-identifiability)

If the triple (Y, Z, H) satisfies the LF-HTC for $v \in V$, then column $\Lambda_{*,v}$ is a rational function of

- *the observed covariance matrix Σ,*
- *the columns $\Lambda_{*,z}$ for $z \in Z$, and*
- *the columns $\Lambda_{*,y}$ for those $y \in Y$ that can be reached from $Z \cup \{v\}$ using a half-trek that avoids H.*

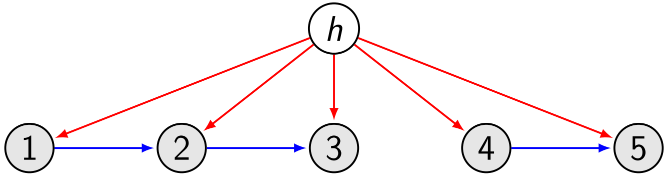
Algorithm

- Cycle through nodes v and search for LF-HTC triples that allow solving for $\Lambda_{*,v}$.
- Network-flow setup finds LF-HTC triples in polynomial time under a bound on $|Z| = |H|$.

Theorem

If we do not bound the rank $|Z| = |H|$, then LF−HTC is NP-complete.

Example: One Latent Factor

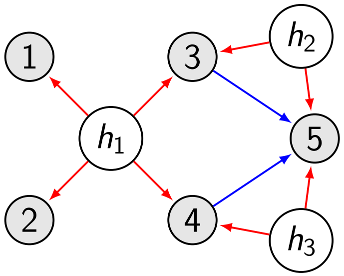


1. $\underline{v \in \{1,4\}}$: Trivially, $\Lambda_{*,1} = \Lambda_{*,4} = 0$.
2. $\underline{v = 3}$: Take $Y = \{1,2\}$, $Z = \{4\}$, $H = \{h\}$.

   (ii) $Y \cap (Z \cup \{3\}) = \{1,2\} \cap \{3,4\} = \emptyset$, (iii) void, (iv) $1 \leftarrow h \rightarrow 4$, $2 \equiv 2$
3. $\underline{v = 2}$: Take $Y = \{1,3\}$, $Z = \{4\}$, $H = \{h\}$.
4. $\underline{v = 5}$: Take $Y = \{3,4\}$, $Z = \{1\}$, $H = \{h\}$.

Subtleties of Latent Projections (Mixed Graphs)

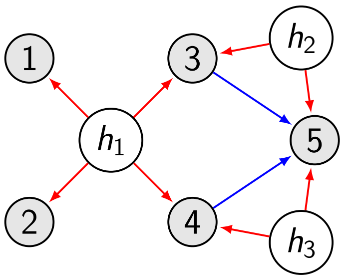
Original Graph



- LF-HTC not applicable and generically infinite-to-one
- Model-dimension: 11

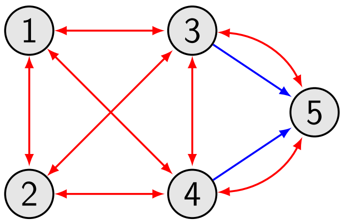
Subtleties of Latent Projections (Mixed Graphs)

Original Graph



- LF-HTC not applicable and generically infinite-to-one
- Model-dimension: 11

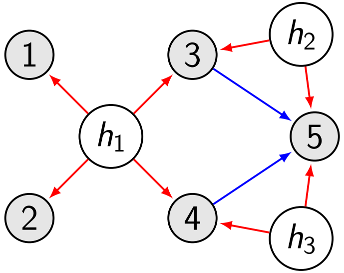
Projection (Mixed Graph)



- HTC-identifiable
- Model-dimension: 13

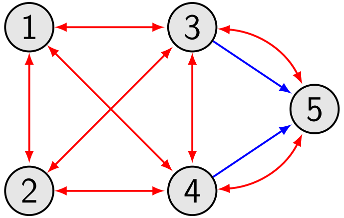
Subtleties of Latent Projections (Mixed Graphs)

Original Graph



- LF-HTC not applicable and generically infinite-to-one
- Model-dimension: 11

Projection (Mixed Graph)



- HTC-identifiable
- Model-dimension: 13

Subtlety: Identifiability of mixed graphs may be due to the assumption that confounding is caused by multiple factors!

Conclusion

- Many applications required modeling effects of latent variables.
- As projections, latent variable models may feature complicated parametrizations and geometry.
- Lots to explore still, in identification and for other problems...
- Some background reading:

Barber, Drton, Sturma, Weihs (2022).
Half-Trek Criterion for Identifiability of Latent Variable Models. arXiv:2201.04457. (Ann. Statist., forthcoming)

Foygel, Draisma, Drton (2012).
Half-Trek Criterion for Generic Identifiability of Linear Structural Equation Models. Ann. Statist. 40, no. 3, 1682–1713.

Drton (2018).
Algebraic Problems in Structural Equation Modeling.
The 50th Anniversary of Gröbner bases, Adv. Stud. Pure Math., 77, Math. Soc. Japan, Tokyo, pages 35–86.

European Research Council
Established by the European Commission