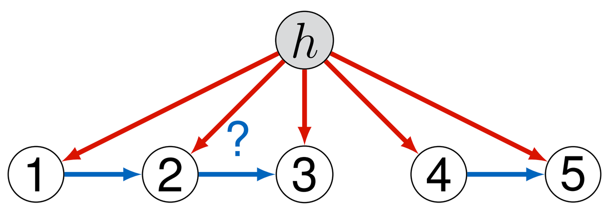
Research group Mathematical Statistics
TUM School of Computation, Information and Technology
Technical University of Munich

Half-Trek Criterion for Identifiability of Latent Variable Models

Rina Foygel Barber[1], Mathias Drton[2], Nils Sturma[2], Luca Weihs[3]

1. Motivation

Assumption: Known causal structure between observed and latent variables.



Latent-factor graph with 5 observed nodes and one latent node h.

Aim: Identify the direct causal effects between the observed variables based on the observed covariance matrix.
(identify = uniquely recover)

Main contributions:

- Sufficient condition for rational identifiability in a linear setting.
- Applicable in settings where latent variables may also have dense effects on many or even all of the observables.
- Recursive polynomial time algorithm.
  (when bounding a matrix rank in a search step)

2. Setup

Linear structural equation model with observed variables $X = (X_v)_{v \in V}$ and latent variables $L = (L_h)_{h \in \mathcal{L}}$:

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

- **Sparsity:** Parameter matrices Λ and Γ are supported over the edge set D of a directed graph $G = (V \dot\cup \mathcal{L}, D)$.
- **Latent-factor assumption**: All nodes in $\mathcal{L}$ are source nodes of G.
- **Independence** of the latent factors and the error terms: $\text{Var}[\varepsilon] =: \Omega_{\text{diag}}$ is diagonal and $\text{Var}[L] = I$.

Latent covariance matrix:

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

Observed covariance matrix:

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

3. Rational Identifiability

Given: Latent-factor graph $G = (V \dot\cup \mathcal{L}, D)$.

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

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

Definition. The model given by G is rationally identifiable if there is a rational map ψ_G such that

$$\psi_G \circ \varphi_G(\Lambda, \Gamma, \Omega_{\text{diag}}) = \Lambda$$

for 'almost all' $(\Lambda, \Gamma, \Omega_{\text{diag}})$.

Remark. Always solvable via Gröbner basis computations.
- Double-exponential complexity.
- Only feasible on small graphs.

Software

SEMID
An R-package for parameter identifiability in linear structural equation models.
⟶ available on CRAN and GitHub

4. Key Idea

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

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

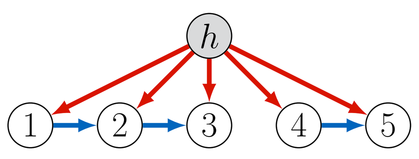
Observation:
The latent covariance matrix may be sparse and feature low-rank structure:

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

⟶ We exploit algebraic relations that are vanishing off-diagonal sub-determinants of Ω.

Example:

$\text{rank}(\Omega_{\{1,2\},\{3,4\}}) = 1$
$\implies \det(\Omega_{\{1,2\},\{3,4\}}) = 0.$



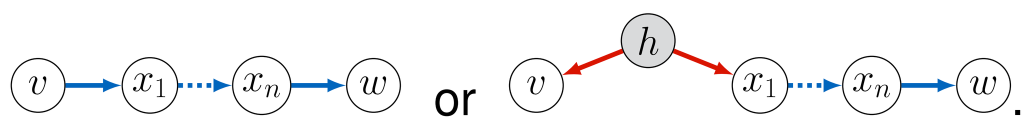
We have the following relations among Λ and Σ:

$$\det\left(\left[(I - \Lambda)^T \Sigma (I - \Lambda)\right]_{\{1,2\},\{3,4\}}\right)$$
$$= \lambda_{23}\sigma_{12}\sigma_{24} - \lambda_{23}\sigma_{14}\sigma_{22} - \sigma_{13}\sigma_{24} + \sigma_{14}\sigma_{23} = 0,$$

which we can then solve for λ_{23}.

5. LF Half-Trek Criterion

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



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

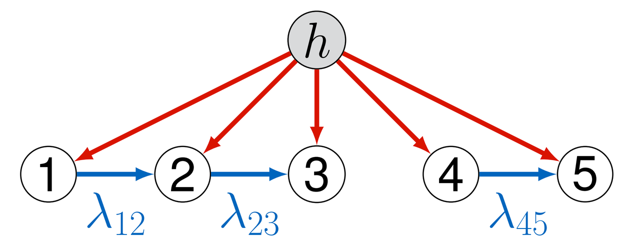
Definition. Let $v \in V$ and $Y, Z \subseteq V \setminus \{v\}$ and $H \subseteq \mathcal{L}$. The triple (Y, Z, H) satisfies the latent-factor half-trek criterion (LF-HTC) for v if

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

Theorem. *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 $(\Lambda_{*,z})_{z \in Z}$ 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. Recursively 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 $|Z| = |H|$.

Example I



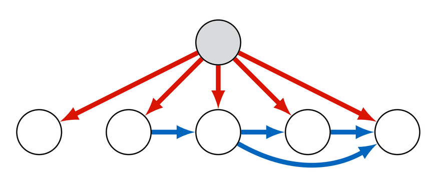
$v \in \{1, 4\}$: Trivially, $\Lambda_{*,1} = \Lambda_{*,4} = 0$.

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

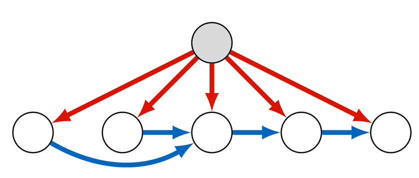
$v = 2$ and $v = 5$: Can find (Y, Z, H) similarly.

$\implies$ The model is LF-HTC-identifiable, that is, the parameters λ_{12}, λ_{23}, and λ_{45} are recovered by rational functions in the entries of Σ.

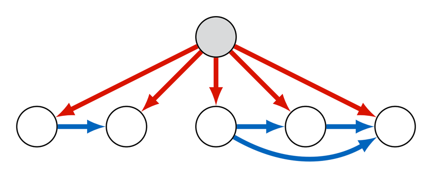
Example II



(a) Rationally identifiable.



(b) Generically finite-to-one but not rationally identifiable.



(c) Generically infinite-to-one.

[1]Department of Statistics, University of Chicago
[2]TUM School of Computation, Information and Technology, Technical University of Munich
[3]Allen Institute for AI

YES Causal Inference Workshop
March 13-15, 2023, Eindhoven