Identifiability of Gaussian DAG models with one latent source

Hisayuki Hara

Niigata University
http://www.econ.niigata-u.ac.jp/~hara/
hara@econ.niigata-u.ac.jp

Joint work with Dennis Leung and Mathias Drton

Contents

Contents

DAG model with one latent variable

$$X = \Lambda^T X + \delta L + \epsilon.$$

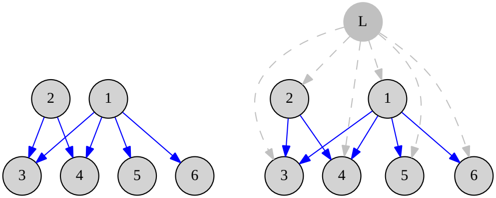
- $X = (X_1, \dots, X_m)^T$: observable variables
- L : a latent variable
  - $L \sim N(0, 1)$
- $\epsilon = (\epsilon_1, \dots, \epsilon_m)^T$
  - $\epsilon \sim N(0, \Omega)$, $\Omega = \mathrm{diag}(\omega_1, \dots, \omega_m)$
- $\Lambda = \{\lambda_{vw}\}$: strictly upper triangular
- $\delta = (\delta_1, \delta_2, \dots, \delta_m)$: factor loads
- $X \sim N(0, \Sigma)$,

$$\Sigma = (I_m - \Lambda^T)^{-1}(\Omega + \delta\delta^T)(I_m - \Lambda)^{-1}.$$

The DAG model with a latent variable

$$X = \Lambda^T X + \delta L + \epsilon.$$

- A factor analysis model s.t.
  - one latent variable
  - DAG structure among observable variables



G $\quad$ G^*

- $G = (V, E)$: DAG for observable variables
- G^* : DAG for the model with a latent variable

parametrization map

- $\theta := (\Lambda, \Omega, \delta) \in \Theta := \mathbb{R}^{|E|} \times \mathbb{R}^m_{>0} \times \mathbb{R}^m$.
- $\dim\Theta = |E| + 2m$.
- parametrization map :

$$\phi_G : \theta \mapsto (I_m - \Lambda^T)^{-1}(\Omega + \delta\delta^T)(I_m - \Lambda)^{-1}.$$

- Λ : strictly upper triangular

$$(I_m - \Lambda)^{-1} = I_m + \Lambda + \Lambda^2 + \cdots + \Lambda^{m-1}.$$

- ϕ_G is a polynomial map on θ.
- The model is called globally identifiable when ϕ_G is one-to-one.

Identifiability of models

- $\phi_G(\Lambda, \Omega, \delta) = \phi_G(\Lambda, \Omega, -\delta)$
  - not globally identifiable
- When $\Lambda = 0$, ϕ_G is 2-to-1
  - Anderson and Rubin(1956)
- When $\Lambda \neq 0$, ϕ_G could be
  - ∞-to-1
  - generically k-to-1 with $2 < k < \infty$
  - not necessarily 2-to-1

Generially finite identifiability

When ϕ_G is generically finite-to-one, the model is called generically finite identifiable(GFI).

Computational algebraic considerations

- $F(\theta, \Sigma) := (I_m - \Lambda^T)^{-1}(\delta\delta^T + \Omega)(I_m - \Lambda)^{-1} - \Sigma$
- $f_{ij}(\theta)$: (i, j) element of $F(\theta, \Sigma)$
- I_G : an ideal generated by $\{f_{ij}(\theta) : i > j\}$

$$I_G = \langle f_{11}, f_{12}, \ldots, f_{mm} \rangle$$

Proposition(e.g. Cox et al.)

When I_G is zero-dimensional, $F(\theta, \Sigma) = 0$ has at most finitely many solutions.

Question

Under what conditions on G (or G^*) ϕ_G is finite-to-one?

Contents

G_{con} and $G_{|L,cov}$

- $G_{con} = (V, E_{con})$: conditional independence graph of G
- $G_{|L,cov} = (V, E_{|L,cov})$ represents marginal dependency of variable pairs after conditioning on L.
  - G_{con} and $G_{|L,cov}$ are easily obtained from G
- $G_{con}^c = (V, E_{con}^c)$: complementary graph of $G_{con} = (V, E_{con})$
- $G_{|L,cov}^c = (V, E_{|L,cov}^c)$: complementary graph of $G_{|L,cov}$

SW condition

Theorem(Stanghellini and Wermuth)

Suppose that G satisfies either of the following conditions,

1. every connected components of G^c_{con} has an odd cycle,
2. every connected components of $G^c_{|L,cov}$ has an odd cycle.

Then the model defined by G^* is generically finite identifiable.

- SW condition is applicable to any DAG model with one latent variable.

SW condition

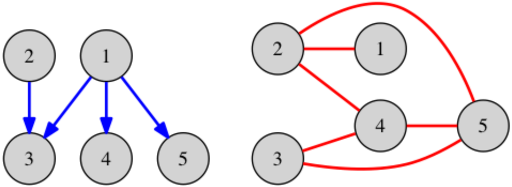
m	4	5	6
GFI models	5	95	3344
SW condition	5	49	985

- The number of GFI models with $m = 4, 5, 6$ computed by Singular.
- SW condition does not look so good.
- Here we provide better conditions.

Contents

Complementary graph of DAG G

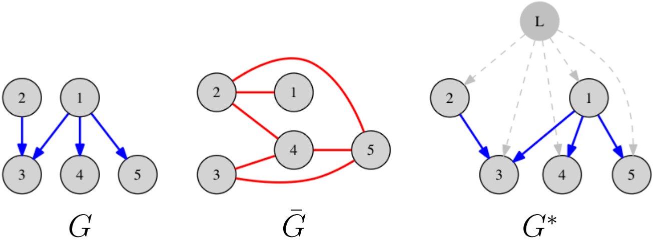
- G : DAG
- $\bar{G}$: complementary graph of an undirected graph obtained by replacing all directed edges of G with undirected edges.



Theorem 1

If every connected component of $\bar{G}$ has an odd cycle, the model defined by G^* is generically finite identifiable.

- $\bar{G}$ has one connected component.
- $2-4-5-2$, $3-4-5-3$ are odd cycles.
- the model associated with G is GFI



Parametrization map for Σ^{-1}

- Σ is
$$\Sigma = (I_m - \Lambda^T)^{-1}(\Omega + \delta\delta^T)(I_m - \Lambda)^{-1}.$$
- Σ^{-1} is
$$\Sigma^{-1} = (I_m - \Lambda)\Omega^{-1}(I_m - \Lambda^T) - \gamma\gamma^T.$$
- $\tilde{\theta} := (\Lambda, \Omega, \gamma)$
$$\tilde{\phi}_G(\tilde{\theta}) : \tilde{\theta} \mapsto (I_m - \Lambda)\Omega^{-1}(I_m - \Lambda^T) - \gamma\gamma^T.$$
- ϕ_G is finite-to-one if and only if $\tilde{\phi}_G$ is finite-to-one.

Jacobian of $\tilde{\phi}_G$ and GFI

Proposition

$\tilde{\phi}_G$ is generically finite-to-one if and only if its Jacobian matrix

$$J(\tilde{\phi}_G) = \frac{\partial \tilde{\phi}_G}{\partial \theta}$$

is generically column full-rank.

- The condition of Theorem 1 is a sufficient condition on $J(\tilde{\phi}_G)$ to be column full rank.

LDH condition 1

m	4	5	6
GFI models	5	95	3344
SW condition	5	49	985
LDH condition 1	5	88	2957
SW and LDH	5	88	2957

- We can see that our condition is better than SW condition.
- There still exist some GFI models that do not satisfy our condition.
- When m is bigger, the ratio of GFI models that do not satisfy our condition increases.

Contents

Theorems

Theorem 2

Suppose that G has a sink node v satisfying

- $\mathrm{pa}(v) \neq V \setminus \{v\}$,
- the model defined by $G^*(V \setminus \{v\})$ is GFI.

Then the model defined by G^* is GFI.

Theorem 3

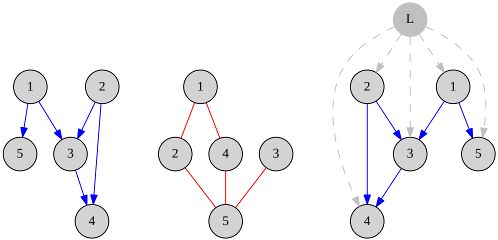
Suppose that G has a source node v satisfying

- $\mathrm{ch}(v) \neq V \setminus \{v\}$.
- the model defined by $G^*(V \setminus \{v\})$ is GFI.

Then the model defined by G^* is GFI.

Sufficient condition 2

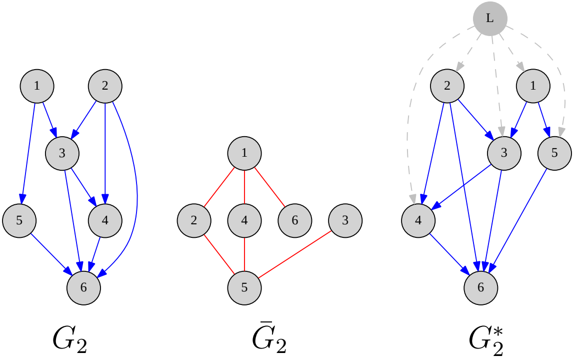
- $\bar{G}_1$ does not have an odd cycle.
- Symbolic computation shows that the model associated with G_1^* is GFI.



G_1 $\quad\bar{G}_1\quad$ G_1^*

LDH condition 2

- Add a sink node $\{6\}$ to G_1.
- $\mathrm{pa}(6) = \{2, 3, 4, 5\}$.
- $\bar{G}_2$ also have no odd cydle.
- But the model defined by G_2^* is also GFI.



tetrad

Lemma

- s_{ij} : (i,j) element of $(I_m - \Lambda^T)\Sigma(I_m - \Lambda)$
- $$\begin{aligned}(I_m - \Lambda^T)\Sigma(I_m - \Lambda) &= \Omega + \delta\delta^T \\ &= \text{diagonal} + \mathrm{rank1}\end{aligned}$$

  if and only if

  $$\tau_{(ik),(jl)}(\Lambda) = s_{ij}s_{kl} - s_{il}s_{jk} = 0,$$

  $$i < j < k < l \text{ or } i < k < j < l.$$

- 2×2 off-diagonal minors are called tetrads.
- Tetrads are all zeros.

tetrad

Proposition

- $\tau_{(ik),(jl)}(\Lambda) = 0$ is a quartic equation on Λ and the model is GFI if and only if

$$\tau_{(ik),(jl)}(\Lambda) = 0,$$

$$i < j < k < l \text{ or } i < k < j < l$$

has finitely many solution.

- For a given Λ,

$$(I_m - \Lambda^T)\Sigma(I_m - \Lambda) = \Omega + \delta\delta^T$$

is 2-to-1 on (Ω, δ).
- By using this fact, we can obtain the conditions.

LDH condition 2

m	4	5	6
GFI models	5	95	3344
SW condition	5	49	985
LDH condition 1	5	88	2957
SW and LDH	5	88	2957

- For $m = 6$, $387 = 3344 - 2957$ models are GFI but do not satisfy Theorem 1.
- It turns out that 194 of them are shown to be GFI in this way.

References

D. Leung, M. Drton, and H. Hara.(2015).
Identifiability of directed Gaussian graphical models with one latent source.
arXiv 1505.01583, submitted.