Partial inversion and partial closure of paths on graphs: two matrix operators to study properties of large systems generated over graphs.

Nanny Wermuth
Chalmers/Gothenburg University

based on joint work with D.R. Cox and Michael Wiedenbeck

- Trying to understand short- and long-term effects of interventions is motivating empirical research in many fields of science

- Graphical Markov models are a flexible tool for the planning, analysis and interpretation of potential data generating processes

An example: a retrospective study with 283 females (Hardt, 2005)

Childhood adversities: first ordering

X, social retreat Y, activity retreat Z, vegetative symptoms	V, childhood happiness	S, love T, constraints U, role reversal W, social support	R, family distress A, sexual abuse	P, age B, schooling Q, family status

Graphical models combine hree simple but powerful concepts

- Wright (1923): used **directed graphs** to study linear processes by which his data might have been generated; **nodes represent variables, arrows dependencies**
- Markov (1908): used the notion of **conditional independence** to simplify and explain seemingly complex joint distributions
- Gibbs (1902): characterized the complexity of systems in **undirected graphs** with nodes differing in their nearest neighbors; **lines represent associations**

Stepwise generating processes

For an ordered set $\boldsymbol{V} = (\mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{d})$ and a vector variable $\boldsymbol{Y_V}$

each $\boldsymbol{Y_i}$ has potentially explanatory variables $\boldsymbol{Y_{r(i)}}$
where $\boldsymbol{r(i) = (i + 1, \ldots, d)}$ and

directly explanatory variables for $\boldsymbol{Y_i}$ are components of $\boldsymbol{Y_{r(i)}}$
needed in a given process to generate the dependencies of $\boldsymbol{Y_i}$

Parent graphs

are graphs in node set V with an arrow starting at **parent** node $j > i$

and pointing to **offspring** node i if and only if

variable Y_j is directly explanatory for Y_i

the graph is denoted by G^V_{par}

Joint densities generated over parent graphs

For the ij-arrow missing in G^V_{par}

a joint density f_V of the form

$$f_V = \prod_{i=1}^{d} f_{i|\text{par}(i)}$$

has Y_i independent of Y_j given $Y_{\text{par}(i)}$, written $i \perp\!\!\!\perp j | \text{par}(i)$

Joint densities generated over parent graphs

For the ij-arrow missing in G^V_{par}

a joint density f_V of the form

$$f_V = \prod_{i=1}^{d} f_{i|\text{par}(i)}$$

has Y_i independent of Y_j given $Y_{\text{par}(i)}$, written $i \perp\!\!\!\perp j | \text{par}(i)$

Note: any distributional form is possible for the densities

The edge matrix $\mathcal{A}$ of a parent graph, G^V_{par}

the edge matrix of G^V_{par} is the $d \times d$ **upper triangular binary matrix** with elements $\mathcal{A}_{ij}$ defined by

$$\mathcal{A}_{ij} = \begin{cases} 1 & \text{if } i \prec\!\!-\, j \text{ in } G^V_{\text{par}} \text{ or } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

The edge matrix $\mathcal{A}$ of a parent graph, G^V_{par}

the edge matrix of G^V_{par} is the $d \times d$ **upper triangular binary matrix** with elements $\mathcal{A}_{ij}$ defined by

$$\mathcal{A}_{ij} = \begin{cases} 1 & \text{if } i \prec\!\!- j \text{ in } G^V_{\text{par}} \text{ or } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

A **corresponding linear system** in mean-centred Y is

$$AY = \varepsilon \quad \text{with} \quad \text{cov}(\varepsilon) = \Delta \quad \text{diagonal}$$

A unit upper-triangular matrix and $\mathcal{A} = \text{In}[A]$

Question: What does the generating process imply?

For instance, **for a joint response model** in two components, $V = (a, b)$, for

$$f_V(y_V) = f_{a|b}(y_a|y_b)\, f_b(y_b)$$

and for **three types of parameter matrices** of a linear chain with

$$\Pi_{a|b} \qquad \Sigma_{aa|b} \qquad \Sigma_{bb}^{-1}$$

$$\mathrm{E}(Y_a|Y_b = y_b) = \Pi_{a|b}\, y_b$$

$$\mathrm{cov}(Y_a|Y_b = y_b) = \Sigma_{aa|b}$$

$$\mathrm{con}(Y_b) = \{\mathrm{cov}(Y_b)\}^{-1} = \Sigma_{bb}^{-1}$$

Partial correlations and independencies in the Gaussian case

for the **regression coefficient matrix** $\Pi_{a|b}$ where $i \in a, \; j \in b$

$$\rho_{ij|b\setminus j} = 0 \iff i \perp\!\!\!\perp j | b \setminus j$$

for the **conditional covariance matrix** $\Sigma_{aa|b}$

$$\rho_{ij|b} = 0 \iff i \perp\!\!\!\perp j | b$$

for the **marginal concentration matrix** Σ_{bb}^{-1}

$$\rho_{ij|b\setminus\{ij\}} = 0 \iff i \perp\!\!\!\perp j | b \setminus \{ij\}$$

Partial correlations and independencies in the Gaussian case

for the **regression coefficient matrix** $\Pi_{a|b}$ where $i \in a,\ j \in b$

$$\rho_{ij|b\setminus j} = 0 \iff i \perp\!\!\!\perp j | b \setminus j$$

for the **conditional covariance matrix** $\Sigma_{aa|b}$

$$\rho_{ij|b} = 0 \iff i \perp\!\!\!\perp j | b$$

for the **marginal concentration matrix** Σ_{bb}^{-1}

$$\rho_{ij|b\setminus\{ij\}} = 0 \iff i \perp\!\!\!\perp j | b \setminus \{ij\}$$

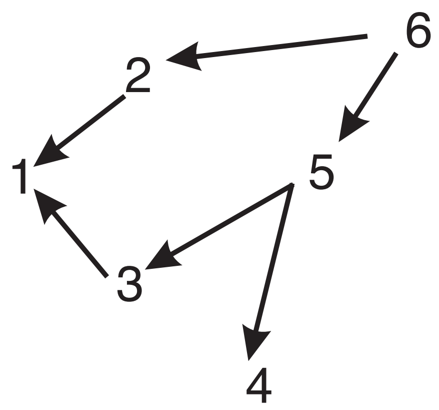
Note: **independencies and graphs generalize** to other distributions

Example: **Induced multivariate regression for Y_a given Y_b**

$$a = (1, 2, 3) \qquad b = (4, 5, 6)$$

parent graph

induced graph

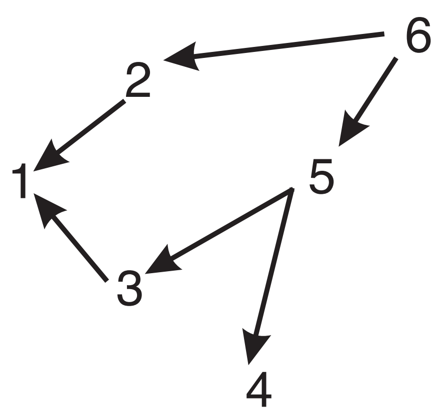


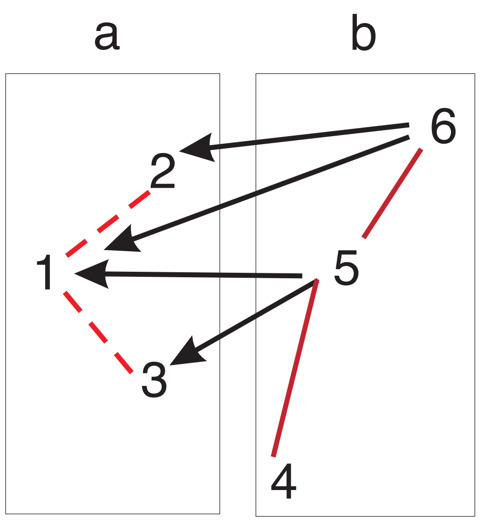
Example: **Induced multivariate regression for Y_a given Y_b**

$$a = (1, 2, 3) \qquad b = (4, 5, 6)$$

parent graph

induced graph

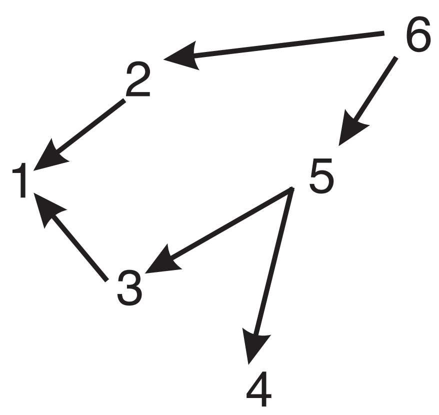




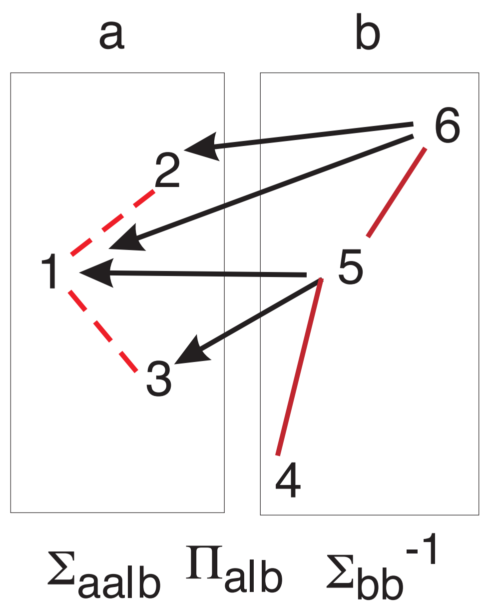
Example: **Induced multivariate regression for Y_a given Y_b**

$$a = (1, 2, 3) \qquad b = (4, 5, 6)$$

parent graph



induced graph



Partial inversion and partial closing of paths

Partial inversion, definition

For all $\boldsymbol{i, j \neq k}$, partial inversion of row and column $\boldsymbol{k}$ of a square matrix $\boldsymbol{M}$ gives $\boldsymbol{N}$ with

$$\begin{aligned}
N_{kk} &= 1/M_{kk} \\
N_{ik} &= M_{ik}/M_{kk} \\
N_{kj} &= -M_{kj}/M_{kj} \\
N_{ij} &= M_{ij} - M_{ik}M_{kj}/M_{kk}
\end{aligned}$$

Partial inversion, properties

Let arbitrary components a, b, c partition V and $G = \{a, b\}$,

then we have

commutativity

$$\mathrm{inv}_a\, \mathrm{inv}_b M = \mathrm{inv}_b\, \mathrm{inv}_a M = \mathrm{inv}_G M$$

symmetric difference

$$\mathrm{inv}_{ab}\, \mathrm{inv}_{bc} M = \mathrm{inv}_{ac} M$$

Partial inversion, properties

Let arbitrary components $\boldsymbol{a, b, c}$ partition $\boldsymbol{V}$ and $\boldsymbol{G = \{a, b\}}$,

then we have

commutativity

$$\mathbf{inv}_{\boldsymbol{a}}\, \mathbf{inv}_{\boldsymbol{b}} \boldsymbol{M} = \mathbf{inv}_{\boldsymbol{b}}\, \mathbf{inv}_{\boldsymbol{a}} \boldsymbol{M} = \mathbf{inv}_{\boldsymbol{G}} \boldsymbol{M}$$

symmetric difference

$$\mathbf{inv}_{\boldsymbol{ab}}\, \mathbf{inv}_{\boldsymbol{bc}} \boldsymbol{M} = \mathbf{inv}_{\boldsymbol{ac}} \boldsymbol{M}$$

Note: first property shared by the Beaton sweep-operator for symmetric matrices (Dempster, 1969) but not the second

Partial closure, definition

Let the unit binary and square matrix $\boldsymbol{\mathcal{M}}$ point with zeros to structural zeros in $\boldsymbol{M}$, then the structural zeros preserved after partial inversion of $\boldsymbol{M}$ on $\boldsymbol{k}$ are contained in $\boldsymbol{\mathcal{N}}$ where

$$
\begin{aligned}
\boldsymbol{\mathcal{N}}_{kk} &= \mathbf{1} \\
\boldsymbol{\mathcal{N}}_{ik} &= \boldsymbol{\mathcal{M}}_{ik} \\
\boldsymbol{\mathcal{N}}_{kj} &= \boldsymbol{\mathcal{M}}_{kj} \\
\boldsymbol{\mathcal{N}}_{ij} &= \begin{cases} \mathbf{1} & \textbf{if } \boldsymbol{\mathcal{M}}_{ij} = \mathbf{1} \textbf{ or } \boldsymbol{\mathcal{M}}_{ik}\boldsymbol{\mathcal{M}}_{kj} = \mathbf{1} \\ \mathbf{0} & \text{otherwise} \end{cases}
\end{aligned}
$$

Partial closure, properties

Let arbitrary components $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ partition $\boldsymbol{V}$ and $\boldsymbol{G} = \{\boldsymbol{a}, \boldsymbol{b}\}$,

then we have

– shared with partial inversion – **commutativity**

$$\mathrm{zer}_a\, \mathrm{zer}_b \mathcal{M} = \mathrm{zer}_b\, \mathrm{zer}_a \mathcal{M} = \mathrm{zer}_G \mathcal{M}$$

– different from partial inversion– **expansion**

$$\mathrm{zer}_{ab}\, \mathrm{zer}_{bc} \mathcal{M} = \mathrm{zer}_{abc} \mathcal{M}$$

Interpretation

Partial inversion applied to all indices of M gives the inverse of M

$$\mathbf{inv}_V \boldsymbol{M} = \boldsymbol{M}^{-1}$$

Partial closure applied to all indices of $\mathcal{M}$, where this is the edge matrix of a graph, gives the edge matrix of the transitive closure of this graph

$$\mathbf{zer}_V \boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{M}}^{-}$$

with

$$\boldsymbol{\mathcal{M}}^{-} = \mathrm{In}[(d\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{M}})^{-1}]$$

Edge matrix components for the induced graph of joint responses Y_a given Y_b

$$\mathrm{Ed}[(\mathcal{A}^{\mathrm{T}}\mathcal{A})] = \begin{pmatrix} \mathcal{S}_{aa|b} & \mathcal{P}_{a|b} \\ . & \mathcal{S}^{bb.a} \end{pmatrix}$$

expressible in terms of $\mathrm{zer}_a\mathcal{A}$

Edge matrix components for the induced graph of joint responses Y_a given Y_b

$$\mathrm{Ed}[(\boldsymbol{\mathcal{A}}^{\mathrm{T}}\boldsymbol{\mathcal{A}})] = \begin{pmatrix} \boldsymbol{\mathcal{S}}_{aa|b} & \boldsymbol{\mathcal{P}}_{a|b} \\ . & \boldsymbol{\mathcal{S}}^{bb.a} \end{pmatrix}$$

expressible in terms of $\mathrm{zer}_a\boldsymbol{\mathcal{A}}$

Parameter matrices in a linear regression chain induced by $\boldsymbol{A}, \boldsymbol{\Delta}$

$$\mathrm{inv}_a(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Delta}^{-1}\boldsymbol{A}) = \begin{pmatrix} \boldsymbol{\Sigma}_{aa|b} & \boldsymbol{\Pi}_{a|b} \\ \sim & \boldsymbol{\Sigma}_{bb}^{-1} \end{pmatrix}$$

expressible in terms of $\mathrm{inv}_a\boldsymbol{A}$

Example continued with $\boldsymbol{a} = \{1, 2, 3\}$, $\boldsymbol{b} = \{4, 5, 6\}$

for $\boldsymbol{A}$ and $\boldsymbol{B} = \mathbf{inv}_{\boldsymbol{a}} \boldsymbol{A}$

$$\begin{pmatrix} \boldsymbol{\Sigma}_{aa|b} & \boldsymbol{\Pi}_{a|b} \\ \sim & \boldsymbol{\Sigma}_{bb}^{-1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}_{aa} \boldsymbol{\Delta}_{aa} \boldsymbol{B}_{aa}^{T} & -\boldsymbol{A}_{aa}^{-1} \boldsymbol{A}_{ab} \\ \sim & \boldsymbol{A}_{bb}^{T} \boldsymbol{\Delta}_{bb}^{-1} \boldsymbol{A}_{bb} \end{pmatrix}$$

Example continued with $\boldsymbol{a} = \{1, 2, 3\}$, $\boldsymbol{b} = \{4, 5, 6\}$

for $\boldsymbol{A}$ and $\boldsymbol{B} = \mathrm{inv}_{\boldsymbol{a}} \boldsymbol{A}$

$$\begin{pmatrix} \Sigma_{aa|b} & \Pi_{a|b} \\ \sim & \Sigma_{bb}^{-1} \end{pmatrix} = \begin{pmatrix} B_{aa} \Delta_{aa} B_{aa}^{T} & -A_{aa}^{-1} A_{ab} \\ \sim & A_{bb}^{T} \Delta_{bb}^{-1} A_{bb} \end{pmatrix}$$

for $\mathcal{A}$ and $\mathcal{B} = \mathrm{zer}_{\boldsymbol{a}} \mathcal{A}$

$$\begin{pmatrix} \mathcal{S}_{aa|b} & \mathcal{P}_{a|b} \\ . & \mathcal{S}^{bb.a} \end{pmatrix} = \mathrm{In}[\begin{pmatrix} \mathcal{B}_{aa} \mathcal{B}_{aa}^{T} & \mathcal{A}_{aa}^{-} \mathcal{A}_{ab} \\ . & \mathcal{A}_{bb}^{T} \mathcal{A}_{bb} \end{pmatrix}]$$

The step from induced linear systems to induced edge matrices

Let defining equations for parameter matrices **without any matrix multiplied by its inverse** be induced by $MY = \eta$ with $\mathcal{M}$ given,

then the induced edge matrices are obtained by **replacing**

(i) every inverse matrix, say M_{aa}^{-1}, by the binary matrix of its structural zeros $\mathcal{M}_{aa}^{-}$,
(ii) every diagonal matrix by an identity matrix of the same dimension,
(iii) every other negative or positive submatrix, say $-M_{ab}$ or M_{ab}, by the corresponding submatrix of structural zeros, $\mathcal{M}_{ab}$, and **then applying the indicator function.**

Induced edge matrices may be used

- to distinguish zero structural and other missing dependencies
- to decide on independence statements implied by a given model
- to derive independence structures implied for other (joint response) models
- to study Markov equivalence
- to obtain graphical criteria for direct and indirect confounding due to ommitted variables

References

Marchetti (2006) J. Stat. Software

Wermuth and Cox (2004) JRSSB

Wermuth, Cox and Marchetti (2006) Bernoulli

Wermuth, Wiedenbeck and Cox (2006), BIT Numerical Mathematics