Symmetries in directed Gaussian graphical models

Visu Makam
visu@umich.edu
Radix Trading Europe B.V.

Philipp Reichenbach
reichenbach@tu-berlin.de
TU Berlin

Anna Seigal
aseigal@math.harvard.edu
Harvard University

European Research Council
Established by the European Commission

Directed acyclic graph (DAG): $\mathcal{G}(I, E)$

- **set of vertices:** $I = [m] := \{1, 2, \dots, m\}$
- **set of directed edges:** $E \subseteq \{i \leftarrow j \mid i, j \in I\}$
- **acyclic:** $\mathcal{G}$ doesn't contain any cycles $i_1 \to i_2 \to \cdots \to i_d \to i_1$

Gaussian model given by a DAG

- **Linear structural equation:** $y = \Lambda y + \varepsilon$, i.e. $y_i = \sum_{(j \to i) \in E} \lambda_{ij} y_j + \varepsilon_i$.
- $y \in \mathbb{R}^m$ is ***data*** on the set of vertices $I = [m]$
- $\lambda_{ij} \in \mathbb{R}$ is ***"effect"*** of vertex j on vertex i ($\lambda_{ij} = 0$ for $j \not\to i$ in $\mathcal{G}$)
- ***independent noise*** on vertices: $\varepsilon \sim \mathcal{N}(0, \Omega)$, diagonal $\Omega \in \mathrm{PD}_m$
- $\mathcal{M} = \{\Psi = (\mathrm{id} - \Lambda)^{\mathsf{T}} \Omega^{-1} (\mathrm{id} - \Lambda)\}$

Introducing restricted DAG (RDAG) models

Idea: Introduce symmetries on the parameters via a graph colouring c.

Motivation:

- vertex and edge symmetries appear in various applications
- natural analog of undirected coloured graphical models [1]
- decrease maximum likelihood thresholds

Maximum Likelihood (ML) estimation on a Gaussian model

- $\mathcal{M} \subseteq \mathrm{PD}_m(\mathbb{R})$ parameter set of *precision matrices*, i.e., $P_\Psi = N(0, \Psi^{-1})$
- $Y = (Y_1, Y_2, \dots, Y_n) \in \mathbb{R}^{m \times n}$, matrix of n **i.i.d. samples**
- **log-likelihood fct.** $\ell_Y : \mathcal{M} \to \mathbb{R}$ measures likeliness of Y
- **MLE given Y:** $\hat{\Psi} \in \mathcal{M}$ such that $\ell_Y(\hat{\Psi}) = \sup_{\Psi \in \mathcal{M}} \ell_Y(\Psi)$

Example 1: DAG model

Consider the DAG $1 \leftarrow 3 \to 2$. The (unrestricted) DAG model is

$$y_1 = \lambda_{13} y_3 + \varepsilon_1, \qquad y_2 = \lambda_{23} y_3 + \varepsilon_2, \qquad y_3 = \varepsilon_3,$$

where $\varepsilon_i \sim N(0, \omega_{ii})$ are independent.

Example 2: Simple RDAG model

Consider the coloured DAG $\textcircled{1} \leftarrow \boxed{3} \to \textcircled{2}$. The RDAG model is

$$y_1 = \lambda y_3 + \varepsilon_1, \qquad y_2 = \lambda y_3 + \varepsilon_2, \qquad y_3 = \varepsilon_3,$$

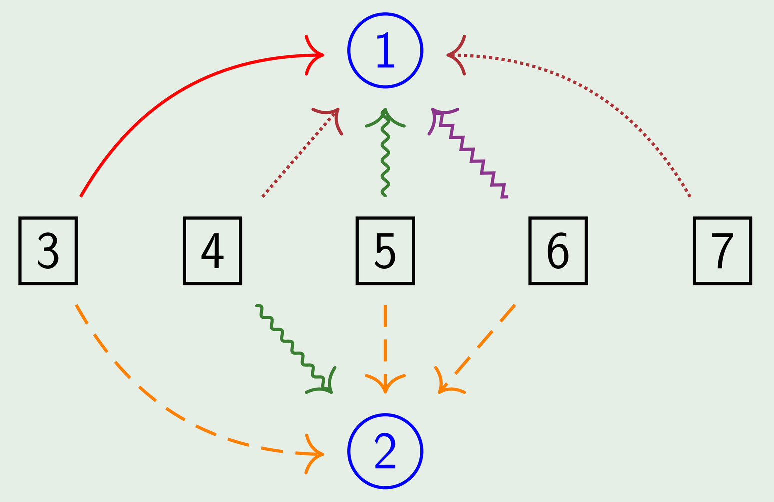
where $\varepsilon_1, \varepsilon_2 \sim N(0, \omega)$ and $\varepsilon_3 \sim N(0, \omega')$ are independent.

Remark: RDAG models include (unrestricted) DAG models

Usual DAG models arise when all vertices and all edges have *distinct* colours.

Example 3: Augmented sample matrix $M_{Y,s}$ for vertex colour s

A sample matrix $Y \in \mathbb{R}^{7 \times n}$, with rows $Y^{(i)}$, for RDAG model on the coloured DAG



has

$$M_{Y,\circ} = \begin{pmatrix} Y^{(1)} & Y^{(2)} \\ Y^{(3)} & 0 \\ 0 & Y^{(3)} + Y^{(5)} + Y^{(6)} \\ Y^{(5)} & Y^{(4)} \\ Y^{(6)} & 0 \\ Y^{(4)} + Y^{(7)} & 0 \end{pmatrix}$$

Definition: Compatible Colouring

A colouring $c : I \cup E \to \{\text{colours}\}$ is **compatible**, if

(i) vertex colours and edge colours are disjoint

(ii) whenever $c(i \leftarrow j) = c(k \leftarrow l)$, then $c(i) = c(k)$.

Consequence: Obtain a partition of the edge colours

$$c(E) = \bigsqcup_{s \in c(I)} \mathrm{prc}(s),$$

where $\mathrm{prc}(s) = \{c(i \leftarrow j) \mid (i \leftarrow j) \in E,\ c(i) = s\}$ are the **parent relationship colours**.

Notation: for vertex colour s

- α_s is the number of vertices of colour s
- $\beta_s := |\mathrm{prc}(s)|$ is the number of parent relationship colours for s
- $M_{Y,s}^{(0)}, M_{Y,s}^{(1)}, \dots, M_{Y,s}^{(\beta_s)}$ are the rows of $M_{Y,s} \in \mathbb{R}^{(\beta_s+1) \times (\alpha_s n)}$
- r_s is the dim'n of $\mathrm{span}\{M_{Y,s}^{(1)}, \dots, M_{Y,s}^{(\beta_s)}\}$ for $n = 1$ and generic $Y \mathbb{R}^{m \times n}$
- **DAG case:** $\alpha_s = 1$; β_s is number of parents; $r_s \in \{0, 1\}$; $M_{Y,s}^{(0)} = Y^{(s)}$; $M_{Y,s}^{(t)}$, $t \in [\beta_s]$ are the parent rows

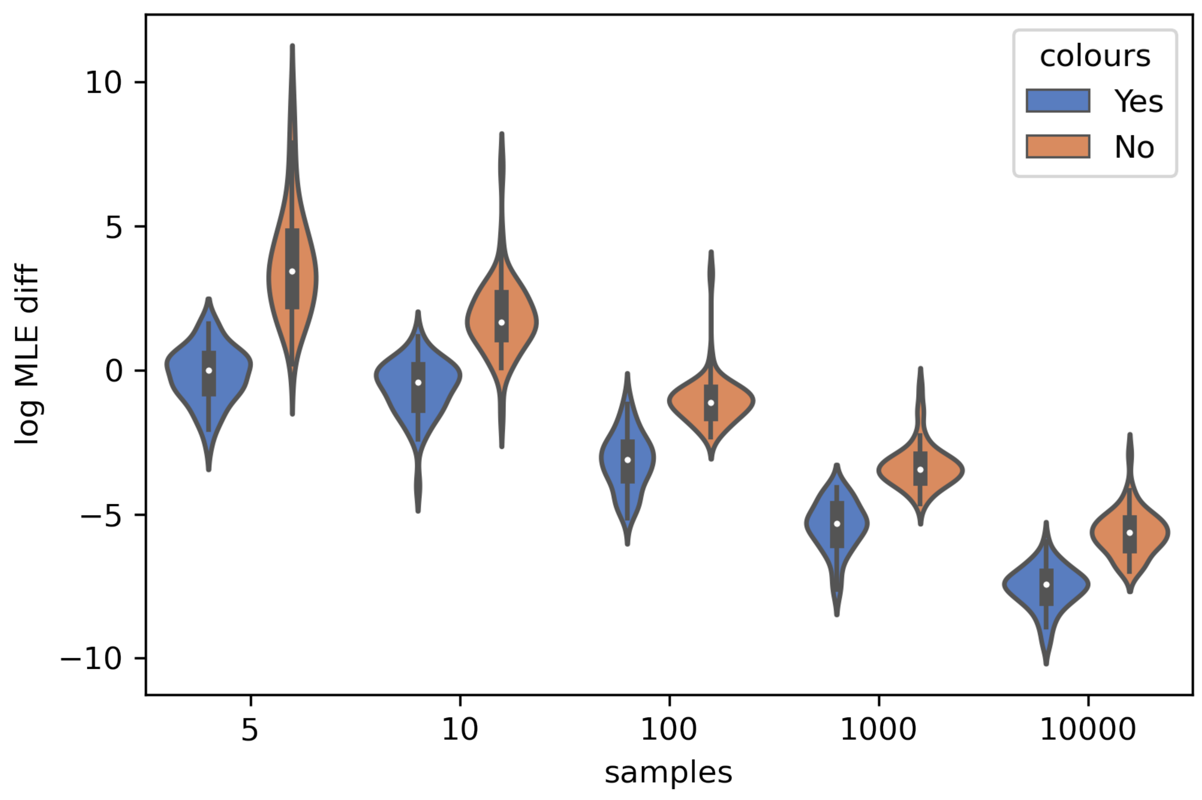


Figure: We generated RDAGs on 10 vertices, with each edge present with probability 0.5 and 5 edge colours. We sampled from the distribution $n \in \{5, 10, 100, 1000, 10000\}$ times. For each n we generated 50 random graphs and computed the RDAG MLE and the DAG MLE.

Theorem 1: (Linear independence conditions for ML estimation)

Consider the RDAG model on $(\mathcal{G}, c)$ where colouring c is compatible, and fix sample matrix $Y \in \mathbb{R}^{m \times n}$. For ML estimation given Y we have:

(a) ℓ_Y unbounded from above $\Leftrightarrow \exists\, s \in c(I)$: $M_{Y,s}^{(0)} \in \mathrm{span}\{M_{Y,s}^{(i)} : i \in [\beta_s]\}$

(b) MLE exists $\Leftrightarrow \forall\, s \in c(I)$: $M_{Y,s}^{(0)} \notin \mathrm{span}\{M_{Y,s}^{(i)} : i \in [\beta_s]\}$

(c) MLE exists uniquely $\Leftrightarrow \forall\, s \in c(I)$: $M_{Y,s}$ has full row rank.

Algorithm to compute MLEs

The proof of Theorem 1 leads to a closed-form formula for MLEs in an RDAG model, as a collection of least squares estimators.

Theorem 2: (Bounds on ML thresholds)

Consider the RDAG model on $(\mathcal{G}, c)$ where colouring c is compatible, and $(\mathcal{G}, c)$ has no edges between vertices of the same colour. Then

$$\max_s \left\lceil \frac{r_s - 1}{\alpha_s - 1} \right\rceil + 1 \le \mathrm{mlt}_e \le \max_s \left\lceil \frac{\beta_s}{\alpha_s} \right\rceil + 1, \qquad (1)$$

$$\max_s \left\lfloor \frac{\beta_s}{\alpha_s} \right\rfloor + 1 \le \mathrm{mlt}_u \le \max_s (\beta_s + 2 - r_s). \qquad (2)$$

What's more?

- Relations to undirected coloured graphical models [2, Section 3]
- further Simulations [2, Section 6]
- Connections to Invariant Theory & Gaussian group models [2, Appendices]

Open Problem

Provide *exact* formulae for ML thresholds [2, Problem 5.4].

References

[1] S. Højsgaard and S. L. Lauritzen. "Graphical Gaussian models with edge and vertex symmetries". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 70.5 (2008), pp. 1005–1027.

[2] V. Makam, P. Reichenbach, and A. Seigal. "Symmetries in Directed Gaussian Graphical Models". 2021. URL: https://arxiv.org/abs/2108.10058.