On perfectness in Gaussian graphical models

Arash A. Amini[†] Bryon Aragam[‡] Qing Zhou[†]

[†]University of California, Los Angeles [‡]University of Chicago

March 14, 2022

Gaussian graphical model

- Consider a regular (i.e., non-singular) Gaussian distribution :

$$X = (X_1, \ldots, X_d) \sim N(\mu, \Gamma^{-1}).$$

- Zero pattern of precision matrix Γ defines a graph $G = (V, E)$.
- $N(\mu, \Gamma^{-1})$ is G-Markov when

$$C \text{ separates } A \text{ and } B \text{ in } G \Longrightarrow X_A \perp\!\!\!\perp X_B \mid X_C$$

- $N(\mu, \Gamma^{-1})$ is G-perfect when

$$C \text{ separates } A \text{ and } B \text{ in } G \Longleftrightarrow X_A \perp\!\!\!\perp X_B \mid X_C$$

- Question: When is $N(\mu, \Gamma^{-1})$ G-perfect?
- Answer (this paper): Almost always.

- Follows from existing results on chain graphs (including DAGs and UGs).
- New direct proof via a transparent parametrization of Gaussian models

Main results

- Let G be any undirected graph on d nodes.
- PD $=$ positive definite

Theorem 1

The set of PD precision matrices Γ that are G-Markov but not G-perfect has Lebesgue measure zero.

Corollary 1

Let μ_G be **any** *continuous distribution on G-Markov precision matrices. Then,*

$$\Gamma \sim \mu_G \implies \mathbb{P}(\Gamma \text{ is } G\text{-perfect}) = 1.$$

Parametrization

- Consider the graph G on 3 nodes

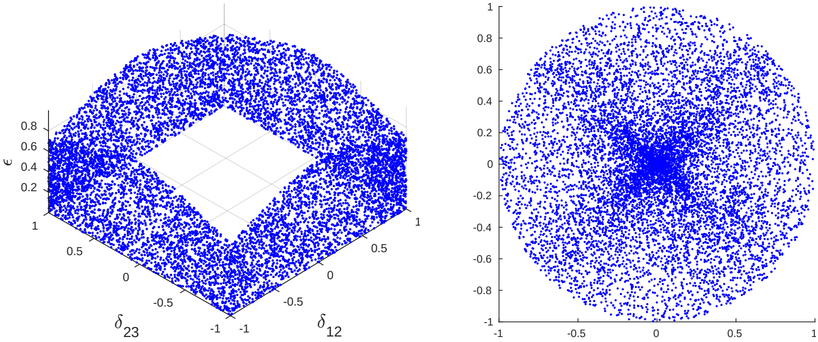
- Parametrize G-Markov (normalized) precision matrices as

$$\Gamma^{G,\delta,\varepsilon} = \begin{pmatrix} 1 & \varepsilon\delta_{12} & 0 \\ \varepsilon\delta_{12} & 1 & \varepsilon\delta_{23} \\ 0 & \varepsilon\delta_{23} & 1 \end{pmatrix},$$

where $\delta \in [-1,1]^G_*$ and ε is at most

$$\varepsilon_G(\delta) := \sup\{\varepsilon > 0 : \ \Gamma^{G,\delta,\varepsilon} \text{ is PD}\}.$$

- Decouples the PD constraints (on ε) and perfectness constraints (on δ).
- Let $\mathcal{M}^G$ be the set of valid (δ, ε).
- Normalizes $\mathcal{M}^G$ further by restricting to $\mathbb{S}^G_\infty := \{\delta \in \mathbb{R}^G : \ \|\delta\|_\infty = 1\}$.
- We construct a probability measure μ_G over $\mathcal{M}^G_\infty$.



- (Left) Samples from μ_G on $\mathcal{M}^G_\infty$ for the graph $G = 1 - 2 - 3$.
- We have $\varepsilon_G((\delta_{12}, \pm 1)) = 1/\sqrt{1 + \delta_{12}^2}$ and similarly for $\varepsilon_G((\pm 1, \delta_{23}))$.
- μ_G is supported on $\mathbb{S}^2_\infty \times [0, 1]$. Singular w.r.t. $\mathcal{L}^3$ but a.c. w.r.t. $\mathcal{H}^2$.
- (Right) Pushforward of μ_G by the map $(\varepsilon, \delta) \to \Gamma^{G,\delta,\varepsilon}$.

Fine structure of the non-perfect set

- The set of non-perfect parameters:

$$\mathcal{N}^G = \{(\delta, \varepsilon) : \ (\Gamma^{G,\delta,\varepsilon})^{-1} \text{ is not perfect}\}.$$

Theorem 2

Let G be a graph with $g := |G| \geq 2$. Let

$$B_\delta := \{\varepsilon : \ (\delta, \varepsilon) \in \mathcal{N}^G\}.$$

There is a "good" set $\mathcal{D}_G$ of δ explicitly given in (8) in the paper such that:

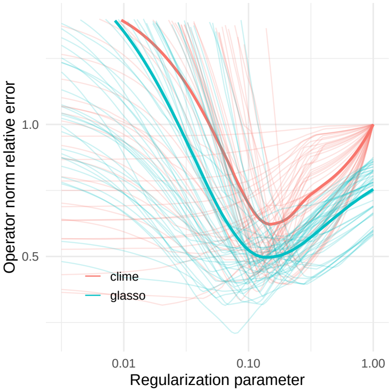
(a) *$\mathcal{D}_G^c$ is an $\mathcal{L}^g$-null set,* — *Set of "bad" δ is small*

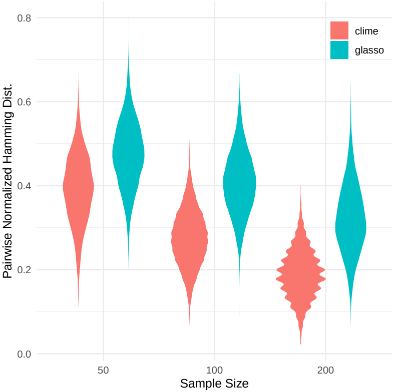
(b) *for every $\delta \in \mathcal{D}_G$, B_δ is finite.* — *For good δ, only finitely many bad ε*

In particular,

(c) *$\mathcal{N}^G$ is a $\mathcal{L}^{g+1}$-null set.* — *Set of "bad pairs" (δ, ε) is small*

Applications





- Using μ_G as a trial distribution in rejection sampling or MCMC, can sample from any continuous distribution over G-perfect models.
- Uncertainty quantification. Simulation-based uncertainty quantification given an estimated graph $\widehat{G}$ by re-sampling from $\mu_{\widehat{G}}$. (Figures above)
- Bayesian inference. Use μ_G as a trial in the Metropolis-Hastings algorithm for efficient posterior sampling.