Rigidity theory for Gaussian graphical models: the maximum likelihood threshold of a graph

Daniel Irving Bernstein (MIT and Fields), Sean Dewar (RICAM and Fields), Shlomo Gortler (Harvard), Tony Nixon (Lancaster University), Meera Sitharam (University of Florida), and Louis Theran (University of St. Andrews)

dibernst@mit.edu
https://dibernstein.github.io

Overview

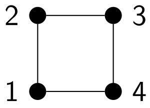
- Broad motivating problem: fit model to relatively small dataset (fewer observations than variables)
- Gaussian graphical models: family of multivariate normal distributions satisfying independence constraints given by a graph
- Goal: use combinatorics of the graph to determine how few observations are needed to be able to fit the graphical model
- Take-home message: rigidity theory offers many tools

Gaussian graphical models

Let $\mu \in \mathbb{R}^V$ and $\Sigma \in \mathbb{R}^{V \times V}$ be positive definite. The ***multivariate normal distribution*** $\mathcal{N}(\mu, \Sigma)$ ***with mean*** μ ***and covariance*** Σ has density

$$f_{\mu,\Sigma}(x) := \frac{\exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu))}{\sqrt{(2\pi)^V \det(\sigma)}}$$

Given a graph $G = (V, E)$, the ***Gaussian graphical model*** $\mathcal{M}_G$ consists of all multivariate normal distributions $\mathcal{N}(\mu, \Sigma)$, with random variables V, such that $(\Sigma)^{-1}_{uv} = 0$ whenever uv is ***not*** an edge of G.



$$\Sigma^{-1} = \begin{pmatrix} x_{11} & x_{12} & 0 & x_{14} \\ x_{12} & x_{22} & x_{23} & 0 \\ 0 & x_{23} & x_{33} & x_{34} \\ x_{14} & 0 & x_{34} & x_{44} \end{pmatrix}$$

Interpretation: $uv \notin E$ means $u \perp\!\!\!\perp v \mid V \setminus \{u, v\}$ for distributions in $\mathcal{M}_G$

Maximum likelihood estimation: definition

Suppose we are given:

- A graph $G = (V, E)$, and
- datapoints $x_1, \dots, x_n$, supposedly iid from some distribution in $\mathcal{M}_G$.

The ***maximum likelihood estimate (MLE)*** is the solution to the following optimization problem, if it exists:

$$\begin{aligned} \max_{\mu, \Sigma} \quad & \prod_{i=1}^{n} f_{\mu,\Sigma}(x_i) \\ \text{s.t.} \quad & (\Sigma^{-1})_{uv} = 0 \text{ for all } uv \notin E, \\ & \Sigma \succ 0 \end{aligned}$$

This can be found via convex optimization.

Maximum likelihood estimation: convex optimization

Let $\hat{\mu}$ and S be the sample mean and covariance (note: $\text{rank}(S) = n$ a.s.)

$$\hat{\mu} := \frac{1}{n}\sum_{i=1}^{n} x_i \qquad\qquad S := \sum_{i=1}^{n}(x_i - \hat{\mu})^T(x_i - \hat{\mu}).$$

The MLE in $\mathcal{M}_G$ exists if and only if the following can be solved:

$$\begin{aligned} \max_{K} \quad & \text{Tr}(SK) + \log\det K \\ \text{s.t.} \quad & K \succ 0 \quad \text{and} \quad K_{uv} = 0 \text{ for all non-edges } uv \text{ of } G. \end{aligned}$$

Theorem (Dempster 1972)

The MLE exists iff there exists $A \succ 0$ satisfying

$$A_{ij} = S_{ij} \qquad \text{if } i = j \text{ or } ij \text{ is an edge of } G.$$

Definition (Maximum likelihood threshold)

Given $G = (V, E)$, $\text{MLT}(G)$ is the minimum n such that the maximum likelihood estimate in $\mathcal{M}_G$ exists almost surely given n datapoints.

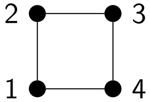
First bounds on MLT

Theorem (Dempster 1972)

MLT(G) *is the minimum* r *such that for almost every* $S \succeq 0$ *of rank* r*, there exists* $A \succ 0$ *such that*

$$A_{ij} = S_{ij} \qquad \text{if } i = j \text{ or } ij \text{ is an edge of } G.$$

- $\mathrm{MLT}(K_n) = n$.
- If G has a k-clique, then $\mathrm{MLT}(G) \geq k$.



$$\begin{pmatrix} x_{11} & x_{12} & \mathbf{x_{13}} & x_{14} \\ x_{12} & x_{22} & x_{23} & \mathbf{x_{24}} \\ \mathbf{x_{13}} & x_{23} & x_{33} & x_{34} \\ x_{14} & \mathbf{x_{24}} & x_{34} & x_{44} \end{pmatrix}$$

Theorem (Buhl 1993)

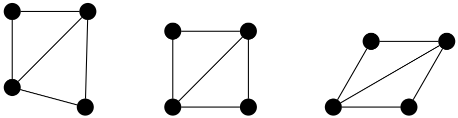
Let $\omega(G)$ *and* $\tau(G)$ *denote the clique number and treewidth of* G*. Then*

$$\omega(G) \leq \mathrm{MLT}(G) \leq \tau(G) + 1.$$

Rigidity theory basics

Definition

A ***bar and joint framework in*** d ***dimensions*** consists of a graph G, and a map $p : V(G) \to \mathbb{R}^d$. Such a framework is ***independent*** if the edge-lengths can be independently perturbed.



Theorem (Asimov and Roth 1978)

Given a graph G*, then if* $p : V(G) \to \mathbb{R}^d$ *is "generic," then whether the framework* (G, p) *is independent in* $\mathbb{R}^d$ *does not depend on* p*.*

One says that G is (generically) independent in $\mathbb{R}^d$ if (G, p) is independent for all generic $p : V(G) \to \mathbb{R}^d$.

Upper bounds via generic independence

Definition (Generic completion rank)

The ***generic completion rank of*** G, denoted $\mathrm{GCR}(G)$, is the minimum d such that G is generically independent in $\mathbb{R}^{d-1}$.

$\mathrm{GCR}(G)$ is also the minimum k such that every generic partial symmetric matrix whose missing entries correspond to the non-edges of G can be completed to rank k.

Theorem (Uhler 2012, Gross and Sullivant 2018)

$\mathrm{MLT}(G) \leq \mathrm{GCR}(G)$.

$\mathrm{GCR}(G)$ can be computed in RP time, so it would be great if the above inequality were sharp. However...

Theorem (Blekherman and Sinn 2019)

$\mathrm{MLT}(K_{5,5}) = 4$ *but* $\mathrm{GCR}(K_{5,5}) = 5$.
$\mathrm{MLT}(K_{n,n})$ *grows linearly with* n *whereas* $\mathrm{GCR}(K_{n,n})$ *grows quadratically.*

MLT in rigidity-theoretic terms

Definition

Let (G, p) and (G, q) be frameworks in $\mathbb{R}^d$ and $\mathbb{R}^e$. Consider the equality

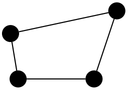
$$\|p(u) - p(v)\| = \|q(u) - q(v)\|.$$

If it holds for all edges uv of G, the frameworks are ***equivalent***. If it moreover holds for ***all pairs*** of vertices, the frameworks are ***congruent***.

A framework (G, p) in $\mathbb{R}^k$ has ***full affine span*** if $\{p(v) : v \in V(G)\}$ affinely spans $\mathbb{R}^k$.

Theorem (BDGNST 2021+)

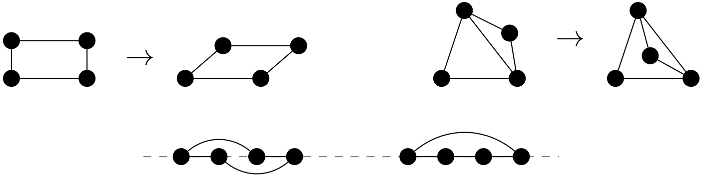
Let G be a graph with n vertices. MLT(G) *is the smallest d such that every generic framework in $\mathbb{R}^{d-1}$ is equivalent to a framework in $\mathbb{R}^{n-1}$ with full affine span.*



More rigidity

One says that (G, p) is:

- universally rigid if (G, p) and (G, q) are congruent when equivalent
- globally rigid if (G, p) and (G, q) are congruent when they are equivalent frameworks in the same dimension
- locally rigid if (G, p) and (G, q) are congruent when they are equivalent frameworks in the same dimension and sufficiently close



- Local and global rigidity are generic properties (Asimov and Roth 1978; Connelly 2005; Gortler, Healy, and Thurston 2010)
- If G has an open set of frameworks in $\mathbb{R}^{d-1}$ that are all universally rigid, then $\mathrm{MLT}(G) > d$

Importing results from rigidity

Theorem (Connelly, Gortler, and Theran 2020)

G is generically globally rigid in $\mathbb{R}^{d-1}$ if and only if there exists an open set of configurations on G in $\mathbb{R}^{d-1}$ that are all universally rigid.

Theorem (BDGNST 2021+)

If a subgraph of G on at least $d+1$ vertices is generically globally rigid in $\mathbb{R}^{d-1}$, then $\mathrm{MLT}(G) > d$.

Implications:

- Lower bounds on MLT generalizing Buhl's result $\omega(G) \leq \mathrm{MLT}(G)$
- If G has fewer than 9 vertices, then $\mathrm{MLT}(G) = \mathrm{GCR}(G)$
- If $\mathrm{GCR}(G) \leq 4$ or $\mathrm{MLT}(G) \leq 3$, then $\mathrm{MLT}(G) = \mathrm{GCR}(G)$

Importing results from low-dimensional rigidity

Proposition (Folklore)

Let G be a graph with n vertices. Then

- *G is independent in $\mathbb{R}^1$ iff G has no cycles*
- *G is globally rigid in $\mathbb{R}^1$ iff G is 2-connected*

If $\mathrm{GCR}(G) = 3$, then $\mathrm{MLT}(G) = 3$:

- $\mathrm{GCR}(G) = 3$ implies G has a cycle
- cycles are globally rigid in $\mathbb{R}^1$, so $\mathrm{MLT}(G) > 2$
- $\mathrm{MLT}(G) \leq \mathrm{GCR}(G)$, so $\mathrm{MLT}(G) = 3$.

Theorem (Berg and Jordán 2003)

If G is 3-connected and minimally dependent in $\mathbb{R}^2$, then G is globally rigid in $\mathbb{R}^2$.

Theorem (BDGNST 2021+)

If $\mathrm{GCR}(G) = 4$ then $\mathrm{MLT}(G) = 4$.

Weak maximum likelihood thresholds

Definition (Weak maximum likelihood threshold)

Given a graph G, $\mathrm{WMLT}(G)$ denotes the minimum n such that the maximum likelihood estimate in $\mathcal{M}_G$ given n datapoints exists with positive probability.

Proposition (Folklore)

$\mathrm{WMLT}(G) = 1$ *iff* G *has no edges.*

Proposition (BDGNST)

If $\mathrm{WMLT}(G) = 2$, *then there exists an orientation of the edges of* G *yielding the order diagram of a partially ordered set.*

- Conjecture: the converse is true too
- If the above conjecture holds, then computing $\mathrm{WMLT}(G)$ is NP-hard

Future work

- Find an algorithm for computing MLT(G) and WMLT(G)
- Find new examples of graphs where MLT(G) $<$ GCR(G)
- Determine if there exists an efficient algorithm for finding the largest d such that G contains a subgraph that is globally rigid in $\mathbb{R}^d$
- Bound MLT(G) in terms of the genus of G (Dewar 2021$+$)
- Can ***coordinated*** rigidity (Schulze, Serocold, and Theran 2018) be used to understand MLTs of ***colored*** Gaussian graphical models?
- Are there any subfields of rigidity theory that could be used to understand MLTs of ***directed*** Gaussian graphical models?

The end

Thank you for your attention!

(Preprint coming soon)