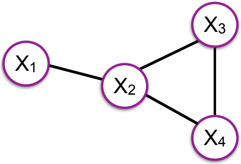
Maximum likelihood threshold of a graph

Elizabeth Gross
University of Hawaiʻi at Mānoa

Joint work with
Seth Sullivant, North Carolina State University

October 29, 2020

Gaussian graphical models



$$X = (X_1, X_2, X_3, X_4) \sim \mathcal{N}(0, \Sigma)$$

The non-edges of G record the conditional independence structure of X:

$$X_1 \perp\!\!\!\perp X_4 \mid (X_2, X_3)$$

$$X_1 \perp\!\!\!\perp X_3 \mid (X_2, X_4)$$

$$\Rightarrow (\Sigma^{-1})_{14} = 0, \ (\Sigma^{-1})_{13} = 0.$$

$\mathbb{S}^m = m \times m$ symmetric real matrices
$\mathbb{S}^m_{>0} =$ pos. def. matrices in $\mathbb{S}^m$
$\mathbb{S}^m_{\geq 0} =$ psd matrices in $\mathbb{S}^m$

Let $G = (V, E)$ with $|V| = m$.

$$\mathcal{M}_G = \{\Sigma \in \mathbb{S}^m_{>0} : (\Sigma^{-1})_{ij} = 0 \text{ for all } i, j \text{ s.t. } i \neq j, ij \notin E\}$$

Definition

The centered **Gaussian graphical model** associated to the graph G is the set of all multivariate normal distributions $\mathcal{N}(0, \Sigma)$ such that $\Sigma \in \mathcal{M}_{\mathcal{G}}$.

Maximum likelihood estimation

Goal: Find Σ that best explains data

Observations: $Y_1, \ldots, Y_n$
Sample covariance matrix: $S = \frac{1}{n}\Sigma_{i=1}^{n} Y_i Y_i^T$

If the MLE exists, it is the unique positive definite matrix Σ that satisfies:

$$\Sigma_{ij} = S_{ij} \text{ for } ij \in E \text{ and } i = j$$

$$(\Sigma)_{ij}^{-1} = 0 \text{ for } ij \notin E \text{ and } i \neq j$$

When $n \geq m$, the MLE exists with probability one. What about the case when $m >> n$?

Question (Lauritzen)

For a given graph G what is the smallest n such that the MLE exists with probability one?

Maximum likelihood threshold

Definition

We call the smallest n such that the MLE exists with probably one (i.e. for generic data) the **maximum likelihood threshold**, or, **mlt**.

Proposition (Buhl 1993)

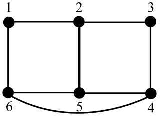
$$\textit{clique number of } G \leq \textit{mlt}(G) \leq \textit{tree width of } G + 1$$

◦ **Clique number:** $\omega(G)$ = size of a largest clique of G

◦ **Chordal graph:** A graph with no induced cycle of length ≥ 4.

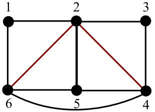
◦ **Chordal cover of $G = (V, E)$:** A graph $H = (V, E')$ such that H is chordal and $E \subseteq E'$.

◦ **Tree width:** $\tau(G) = \min\{\omega(H) - 1 \ : \ H \text{ is a chordal cover of } G\}$.



$\omega(G) = 3$, $\tau(G) = 3$

Chordal cover of G:



Maximum likelihood threshold

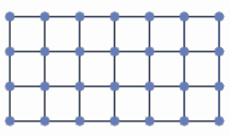
Definition

We call the smallest n such that the MLE exists with probably one (i.e. for generic data) the **maximum likelihood threshold**, or, **mlt**.

Proposition (Buhl 1993)

$$\textit{clique number of } G \leq \textit{mlt}(G) \leq \textit{tree width of } G + 1$$

Notice that these bounds can be far away from each other. Consider for example, $G = Gr_{k_1,k_2}$, the $k_1 \times k_2$ grid graph:



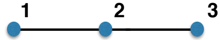
$\omega(G) = \text{size of largest clique} = 2$
$\tau(G) = \text{tree width} = \min(k_1, k_2)$

Geometry of Gaussian graphical models

◦ $|V| = m$ ◦ $\mathbb{S}^m :=$ $m \times m$ symmetric real matrices
◦ $\mathbb{S}^m_{>0} :=$ pos. def. matrices in $\mathbb{S}^m$ ◦ $\mathbb{S}^m_{\geq 0} :=$ psd matrices in $\mathbb{S}^m$

- Let π_G be the projection map that extracts the entries of Σ corresponding to the vertices and edges of G:

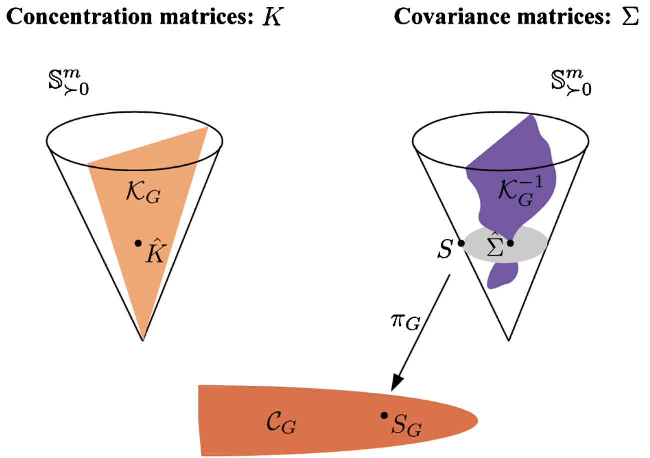
$$\pi_G : \mathbb{S}^m \to \mathbb{R}^{V+E}$$
$$\phi_G(\Sigma) = (\sigma_{ii})_{i \in V} \oplus (\sigma_{ij})_{ij \in E}$$

G: 

$$\phi_G\left(\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}\right)$$
$$= (1, 1, 1, 2, 2)^T$$

- **Cone of sufficient statistics**: $\mathcal{C}_G := \phi_G(\mathbb{S}^m_{>0})$.
- For a given $S \in \mathbb{S}^m_{\geq 0}$, the MLE exists if and only if $\phi_G(S) \in \mathrm{int}(\mathcal{C}_G)$.
- $\mathcal{C}_G$ is the convex dual to the cone of concentration matrices $\mathcal{K}_G$

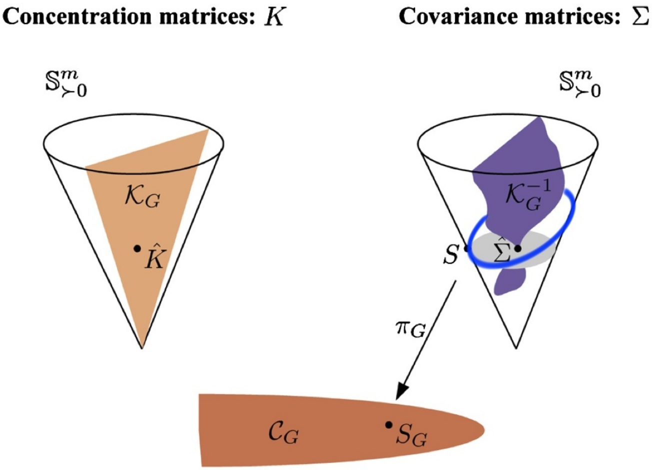
Geometry of maximum likelihood estimation



Uhler, *Geometry of maximum likelihood estimation in Gaussian graphical models* (2012)

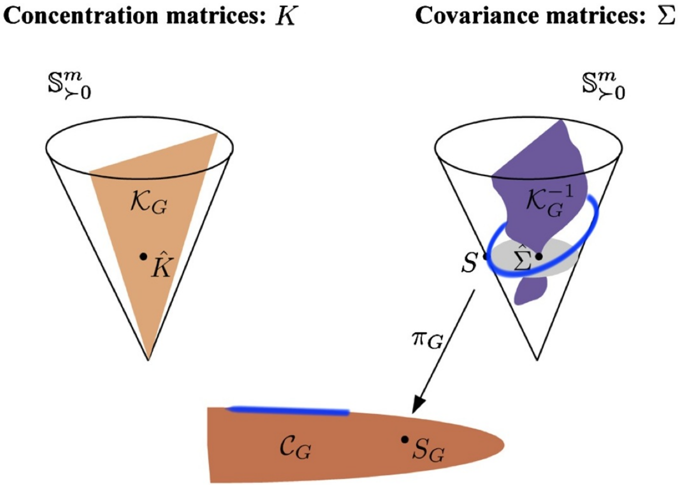
Light orange: $\mathcal{K}_G$, cone of concentration matrices, **Purple**: $\mathcal{K}_G^{-1}$, cone of covariance matrices, **Gray**: Set of positive definite completions of S, **Dark orange**: $\mathcal{C}_G$, Cone of sufficient statistics

Geometry of maximum likelihood estimation



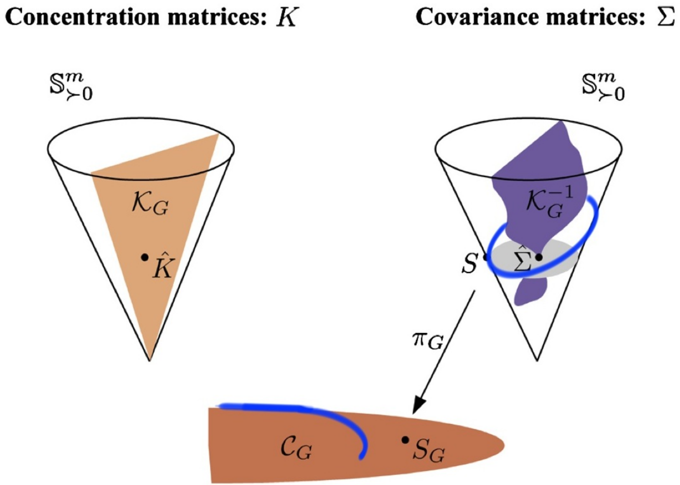
Blue: $S(m,n) \cap \mathbb{S}_{\geq 0}$, Set of $m \times m$ positive semi-definite symmetric matrices of rank $\leq n$

Geometry of maximum likelihood estimation



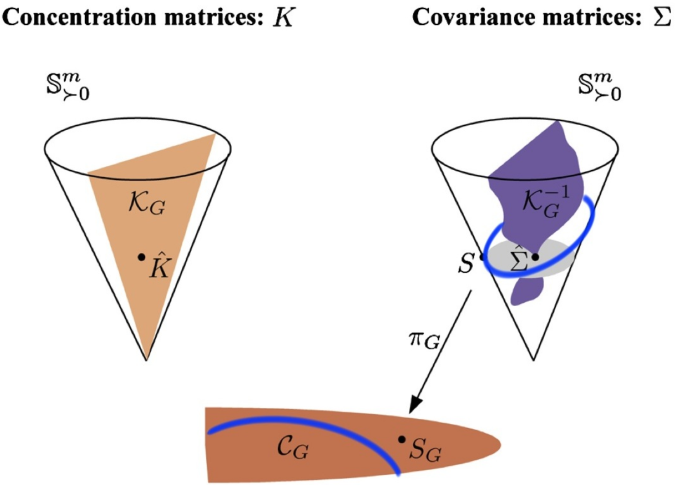
Blue: $S(m,n) \cap \mathbb{S}_{\geq 0}$, Set of $m \times m$ positive semi-definite symmetric matrices of rank $\leq n$

Geometry of maximum likelihood estimation



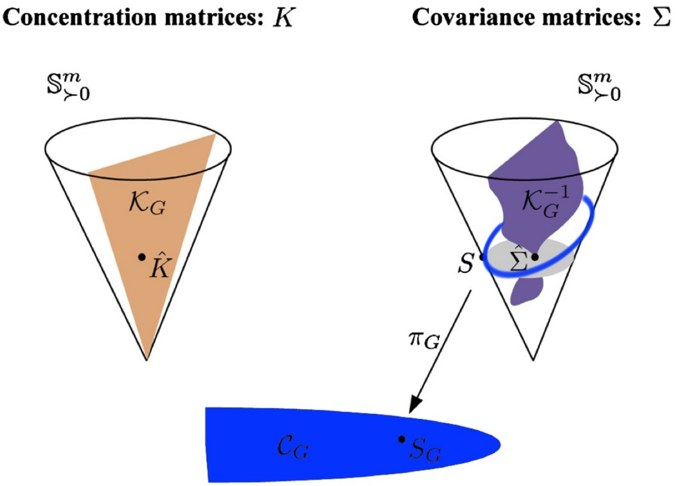
Blue: $S(m, n) \cap \mathbb{S}_{\geq 0}$, Set of $m \times m$ positive semi-definite symmetric matrices of rank $\leq n$

Geometry of maximum likelihood estimation



Blue: $S(m,n) \cap \mathbb{S}_{\geq 0}$, Set of $m \times m$ positive semi-definite symmetric matrices of rank $\leq n$

Geometry of maximum likelihood estimation



Blue: $S(m,n) \cap \mathbb{S}_{\geq 0}$, Set of $m \times m$ positive semi-definite symmetric matrices of rank $\leq n$

Rank of a graph

Definition

Let $S(m,n) := \{\Sigma \in \mathbb{S}^m \ : \ \text{rank}(\Sigma) \leq n\}$.

The **rank of a graph** G is the minimal n such that $\dim \phi_G(S(m,n)) = \dim \mathcal{C}_G = |V| + |E|$

Proposition (Uhler 2012)

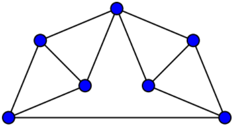
$$mlt(G) \leq rank(G)$$

Goal: Connect the rank of a graph to combinatorial rigidity theory.

Method: Use algebraic matroids, in particular compare the rigidity matroid and the symmetric minor matroid.

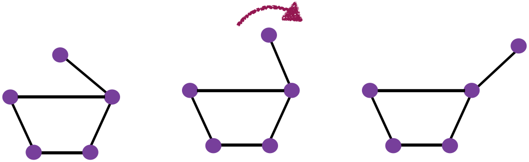
Combinatorial Rigidity Theory

The study of rigidity deals with with **bar and joint frameworks**. A **framework** is a graph G embedded in $\mathbb{R}^n$.



Rigid in $\mathbb{R}^2$

A graph G is called **rigid** if, for generic points $\mathbf{p}_1, \ldots, \mathbf{p}_m \in \mathbb{R}^n$, the only continuous deformations that preserve the distances $||\mathbf{p}_i - \mathbf{p}_j||_2$ for $ij \in E$ are rotations and translations.



Not rigid in $\mathbb{R}^2$

The rigidity matroid

A **matroid** M is a pair (E, I) where E is a finite set of elements, called the **ground set** and I is a collection of subsets of E, called the **independent sets**.

Consider the map $\psi_n \ : \mathbb{R}^{n\times m} \to \mathbb{R}^{m(m-1)/2}$

$$(\mathbf{p}_1, \ldots, \mathbf{p}_m) \mapsto (||\mathbf{p}_i - \mathbf{p}_j||_2^2 \ : \ 1 \leq i < j \leq m).$$

This is polynomial map with an associated matroid:

- $E =$ columns of the Jacobian at a generic point
- $I =$ all collections of independent columns

This matroid is called the **n - dimensional generic rigidity matroid**, denoted $\mathcal{A}(n)$.

- Spanning sets in the matroid are called (generically infinitesimally) rigid graphs.
- Bases in the matroid are called (generically) isostatic graphs.

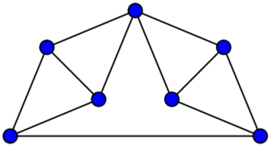
Laman's Theorem

- To be a spanning set in $\mathcal{A}(n)$ must have $nm - \binom{n+1}{2} \leq |E|$.
- Conversely, to be independent must have $nm - \binom{n+1}{2} \geq |E|$.

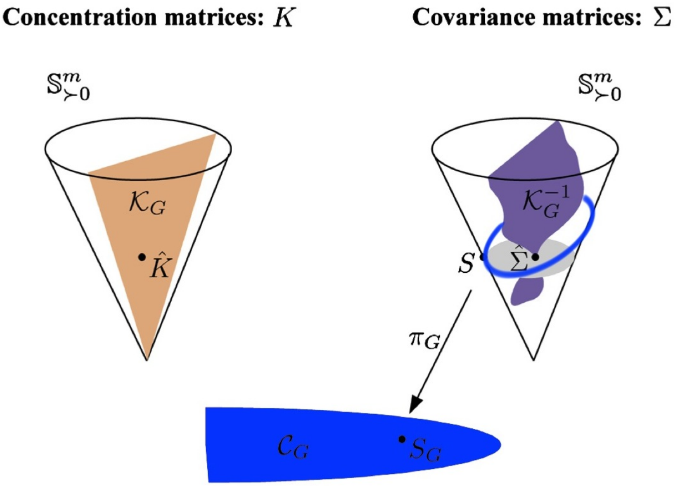
Theorem (Laman 1970)

A graph $G = (V, E)$ is a basis (isostatic) in the rigidity matroid $\mathcal{A}(2)$ if and only if

- $|E| = 2m - 3$, *and*
- *For every induced subgraph* $G_W = (W, E_W), \; |E_W| \leq 2|W| - 3$.



Back to the rank of a graph



Blue: $S(m, n) \cap \mathbb{S}_{\geq 0}$, Set of $m \times m$ positive semi-definite symmetric matrices of rank $\leq n$

Symmetric minor matroid

Definition

Let $I \subset \mathbb{K}[x_1, \ldots, x_r]$ be a prime ideal. This defines an **algebraic matroid** with ground set $\{x_1, \ldots, x_r\}$, and $K \subseteq \{x_1, \ldots, x_r\}$ an independent set if and only if $I \cap \mathbb{K}[K] = \langle 0 \rangle$.

$S(m, n)$ is an algebraic set whose defining ideal I_n is generated by the $(n+1) \times (n+1)$-minors of a $m \times m$ symmetric matrix $\Sigma = (\sigma_{ij})$.

The set $S(m, n)$ has an associated algebraic matroid:

- ground set $= \{\sigma_{ij} \; : \; i < j\}$,
- independent sets= graphs G such that $\pi_G(S(m, n)) = \mathbb{R}^{V+E}$

This matroid is called the **rank n symmetric minor matroid**

Remark

If G is an independent set in the rank n symmetric minor matroid then $\mathrm{rank}(G) \leq n$, and consequently, $\mathrm{mlt}(G) \leq n$. (Uhler 2012)

Rigidity Matroid $\cong$ Symmetric Minor Matroid

Theorem (Gross-Sullivant)

- *A graph G has $rank(G) = n$ if and only if G is an independent set in the $(n-1)$-dimensional rigidity matroid $\mathcal{A}(n-1)$ and not an independent set in $\mathcal{A}(n-2)$.*
- *The $(n-1)$-dimensional rigidity matroid $\mathcal{A}(n-1)$ is isomorphic to the rank n symmetric minor matroid.*

Proof.

Compare the Jacobian of the map

$$(\mathbf{p}_1, \dots, \mathbf{p}_m) \mapsto (||\mathbf{p}_i - \mathbf{p}_j||_2^2 \ : \ 1 \leq i < j \leq m)$$

to the Jacobian of the map

$$(\mathbf{p}_1, \dots, \mathbf{p}_m) \mapsto (\mathbf{p}_i \cdot \mathbf{p}_j \ : \ 1 \leq i < j \leq m)$$

$\square$

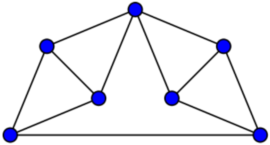
This means that we can bound the mlt of a graph by checking whether G is an independent set in $\mathcal{A}(n-1)$.

Corollary (Laman's Theorem)

Let $G = (V, E)$ be a graph, if for all subgraphs $G' = (V', E')$ of G

$$\#E' \leq 2(\#V') - 3,$$

then $mlt(G) \leq 3$.



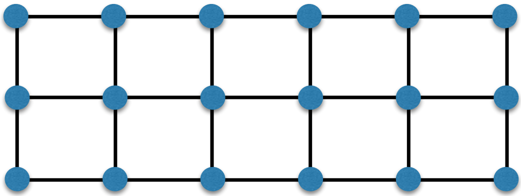
Definition

Let G be a graph and $r \in \mathbb{N}$. The r-**core** of G is the graph obtained by successively removing vertices of G of degree$< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then $\mathrm{rank}(G) \leq n$.

$\Rightarrow$ $\mathbf{mlt}(Gr_{k_1,k_2}) = 3$



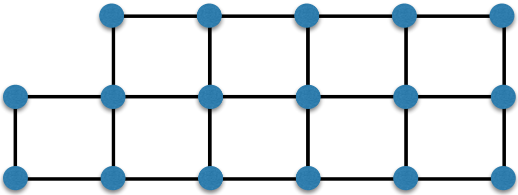
Definition

Let G be a graph and $r \in \mathbb{N}$. The ***r*-core** of G is the graph obtained by successively removing vertices of G of degree$< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then $rank(G) \leq n$.

$\Rightarrow$ $\mathbf{mlt}(Gr_{k_1,k_2}) = 3$



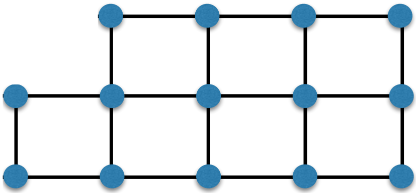
Definition

Let G be a graph and $r \in \mathbb{N}$. The r-**core** of G is the graph obtained by successively removing vertices of G of degree$< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then $rank(G) \leq n$.

$\Rightarrow \mathbf{mlt}(Gr_{k_1,k_2}) = 3$



r-cores

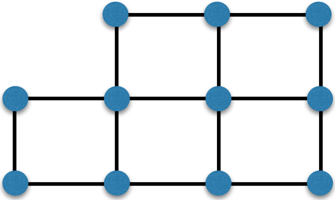
Definition

Let G be a graph and $r \in \mathbb{N}$. The r**-core** of G is the graph obtained by successively removing vertices of G of degree $< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then $rank(G) \leq n$.

$\Rightarrow \mathbf{mlt}(Gr_{k_1,k_2}) = 3$



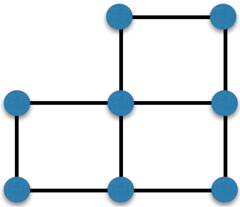
Definition

Let G be a graph and $r \in \mathbb{N}$. The r-**core** of G is the graph obtained by successively removing vertices of G of degree$< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then $rank(G) \leq n$.

$\Rightarrow \mathbf{mlt}(Gr_{k_1,k_2}) = 3$



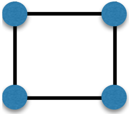
Definition

Let G be a graph and $r \in \mathbb{N}$. The ***r*-core** of G is the graph obtained by successively removing vertices of G of degree$< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then $rank(G) \leq n$.

$\Rightarrow$ $\mathbf{mlt}(Gr_{k_1,k_2}) = 3$



r-cores

Definition

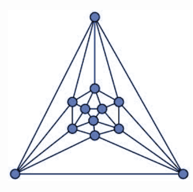
Let G be a graph and $r \in \mathbb{N}$. The r-**core** of G is the graph obtained by successively removing vertices of G of degree$< r$.

Theorem (Gross-Sullivant, Ben-David)

Let G have an empty n-core, then rank$(G) \leq n$.

$\Rightarrow \mathbf{mlt}(Gr_{k_1,k_2}) = 3$

Planar graphs



Theorem (Gross-Sullivant)

If G is a planar graph then $mlt(G) \leq 4$.

Proof.

- Cauchy's theorem implies that every edge graph of simplicial 3-polytope is rigid.
- Edge count $\rightarrow$ G isostatic $\rightarrow$ rank$(G) \leq 4$
- Every planar graph is a subgraph of a graph of a simplicial 3-polytope.

□

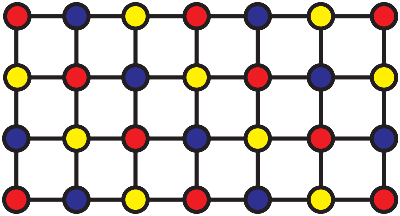
Cycle-Free Coloring

Theorem (Gross-Sullivant 2018)

Let G be a graph and $V_1, \ldots, V_k$ a partition of the vertices of G such that

1. *for all i, V_i is an independent set of G and*
2. *for all $i \neq j$, $G(V_i, V_j)$ has no cycles.*

Then $\text{rank}(G) \leq k$.



Score matching estimator

The **score matching estimator** is a computationally efficient and consistent estimator for Gaussian graphical models (Hyvärinen 2005, Forbes–Lauritzen 2014) .

Definition

We call the smallest n such that the scoring matching exists with probably one (i.e. for generic data) the **scoring matching threshold**, or, **smt**.

Theorem (Gross-Sullivant)

Let G be a graph. Then

$$smt(G) = rank(G).$$

Some questions

- How are the boundary components of $\mathcal{C}_G$ related to the circuits in the rigidity matroid?
- Maximum likelihood threshold has a natural rigidity theory analogue: are they equivalent?
- Determine the score matching threshold for Gaussian graphical models with symmetries. Can the same methods be used here?
- How different can the maximum likelihood threshold be from the weak maximum likelihood threshold?

Thank you

- G. Blekherman and R. Sinn. Maximum likelihood threshold and generic completion rank of graphs. *Discrete & Computational Geometry* 61, no. 2 (2019): 303-324.

- D. I. Bernstein, G. Blekherman, and R. Sinn. Typical and generic ranks in matrix completion. *Linear Algebra and its Applications* 585 (2020): 71-104.

- S. Buhl. On the existence of maximum likelihood estimators for graphical Gaussian models. *Scand. J. Statist.* **20** (1993), no. 3, 263–270.

- **E. Gross and S. Sullivant. The maximum likelihood threshold of a graph. Bernoulli, 24 no. 1 (2018) 386-407.**

- S. Lauritzen. *Graphical models*. Oxford Statistical Science Series **17**, Oxford University Press, New York, 1996.

- C. Uhler. Geometry of maximum likelihood estimation in Gaussian graphical models. *Ann. Statist.* **40** (2012), no. 1, 238–261.