Maximum likelihood thresholds for colored Gaussian graphical models

Roser Homs Pons, Centre de Recerca Matemàtica

work *in progress* with Olga Kuznetsova, Bernadette Stolz, Danai Deligeorgaki, Joe Johnson, Bryson Kagy and Aida Maraj

New Directions in Algebraic Statistics
IMSI, July 25, 2025

How much data do we need? Maximum likelihood thresholds

Question What is the minimum amount of data required to do statistical inference?

How much data do we need? Maximum likelihood thresholds

Question What is the minimum amount of data required to do statistical inference?
More precisely, given a statistical model, the minimum amount of data required for the existence of the maximum likelihood estimate (MLE) of its parameters...

How much data do we need? Maximum likelihood thresholds

Question What is the minimum amount of data required to do statistical inference?
More precisely, given a statistical model, the minimum amount of data required for the existence of the maximum likelihood estimate (MLE) of its parameters...

- ...with probability 1 is called the *maximum likelihood threshold* of the model.

How much data do we need? Maximum likelihood thresholds

Question What is the minimum amount of data required to do statistical inference?

More precisely, given a statistical model, the minimum amount of data required for the existence of the maximum likelihood estimate (MLE) of its parameters...

- ...with probability 1 is called the *maximum likelihood threshold* of the model.
- ...with positive probability is called the *weak maximum likelihood threshold* of the model.

How much data do we need? Maximum likelihood thresholds

Question What is the minimum amount of data required to do statistical inference?
More precisely, given a statistical model, the minimum amount of data required for the existence of the maximum likelihood estimate (MLE) of its parameters...

- ...with probability 1 is called the *maximum likelihood threshold* of the model.
- ...with positive probability is called the *weak maximum likelihood threshold* of the model.

For graphical models, we use the notation $\mathrm{mlt}(G)$ and $\mathrm{wmlt}(G)$.

How much data do we need? Maximum likelihood thresholds

Question What is the minimum amount of data required to do statistical inference?
More precisely, given a statistical model, the minimum amount of data required for the existence of the maximum likelihood estimate (MLE) of its parameters...

- ...with probability 1 is called the *maximum likelihood threshold* of the model.
- ...with positive probability is called the *weak maximum likelihood threshold* of the model.

For graphical models, we use the notation $\mathrm{mlt}(G)$ and $\mathrm{wmlt}(G)$.

Our models: Gaussian graphical models with additional vertex and edge symmetries

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

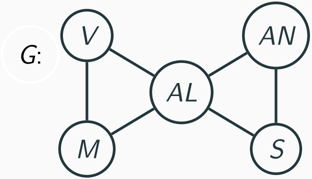
$$K_G = \begin{pmatrix} 5.24 & -2.44 & -2.74 & 0.01 & 0.14 \\ -2.44 & 10.43 & -4.71 & -0.79 & -0.17 \\ -2.74 & -4.71 & 26.95 & -7.05 & -4.70 \\ 0.01 & -0.79 & -7.05 & 9.88 & -2.02 \\ 0.14 & -0.17 & -4.70 & -2.02 & 6.45 \end{pmatrix}$$

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

$$K_G = \begin{pmatrix} 5.24 & -2.44 & -2.74 & 0.01 & 0.14 \\ -2.44 & 10.43 & -4.71 & -0.79 & -0.17 \\ -2.74 & -4.71 & 26.95 & -7.05 & -4.70 \\ 0.01 & -0.79 & -7.05 & 9.88 & -2.02 \\ 0.14 & -0.17 & -4.70 & -2.02 & 6.45 \end{pmatrix}$$



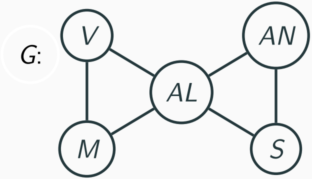
Whittaker, 1990; Edwards, 2000

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

$$K_G = \begin{pmatrix} 5.24 & -2.44 & -2.74 & 0.01 & 0.14 \\ -2.44 & 10.43 & -4.71 & -0.79 & -0.17 \\ -2.74 & -4.71 & 26.95 & -7.05 & -4.70 \\ 0.01 & -0.79 & -7.05 & 9.88 & -2.02 \\ 0.14 & -0.17 & -4.70 & -2.02 & 6.45 \end{pmatrix}$$



Whittaker, 1990; Edwards, 2000

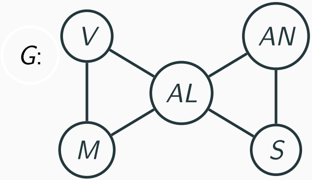
Theorem Uncolored chordal graphs:
$\mathrm{wmlt}(G) = \mathrm{mlt}(G) =$ maximal clique size

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

$$K_G = \begin{pmatrix} 5.24 & -2.44 & -2.74 & 0.01 & 0.14 \\ -2.44 & 10.43 & -4.71 & -0.79 & -0.17 \\ -2.74 & -4.71 & 26.95 & -7.05 & -4.70 \\ 0.01 & -0.79 & -7.05 & 9.88 & -2.02 \\ 0.14 & -0.17 & -4.70 & -2.02 & 6.45 \end{pmatrix}$$



Whittaker, 1990; Edwards, 2000

Theorem Uncolored chordal graphs:
$\mathrm{wmlt}(G) = \mathrm{mlt}(G) =$ maximal clique size

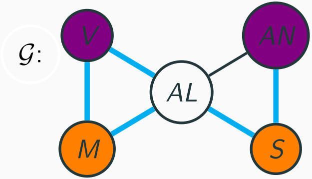
$$\mathrm{wmlt}(G) = \mathrm{mlt}(G) = 3$$

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

$$K_G = \begin{pmatrix} 5.24 & -2.44 & -2.74 & 0.01 & 0.14 \\ -2.44 & 10.43 & -4.71 & -0.79 & -0.17 \\ -2.74 & -4.71 & 26.95 & -7.05 & -4.70 \\ 0.01 & -0.79 & -7.05 & 9.88 & -2.02 \\ 0.14 & -0.17 & -4.70 & -2.02 & 6.45 \end{pmatrix}$$



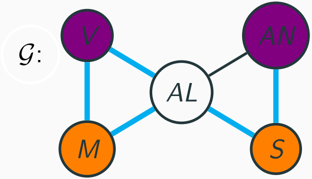
Hogsgaard-Lauritzen, 2008

Start with an example

Table 1. Empirical concentrations (on or above the diagonal) and partial correlations (below the diagonal) for the examination marks in five mathematical subjects

Subject	*Concentrations (×1000) and partial correlations*				
	Mechanics	*Vectors*	*Algebra*	*Analysis*	*Statistics*
Mechanics	5.24	−2.44	−2.74	0.01	−0.14
Vectors	0.33	10.43	−4.71	−0.79	−0.17
Algebra	0.23	0.28	26.95	−7.05	−4.70
Analysis	−0.00	0.08	0.43	9.88	−2.02
Statistics	0.02	0.02	0.36	0.25	6.45

$$K_G = \begin{pmatrix} 5.24 & -2.44 & -2.74 & 0.01 & 0.14 \\ -2.44 & 10.43 & -4.71 & -0.79 & -0.17 \\ -2.74 & -4.71 & 26.95 & -7.05 & -4.70 \\ 0.01 & -0.79 & -7.05 & 9.88 & -2.02 \\ 0.14 & -0.17 & -4.70 & -2.02 & 6.45 \end{pmatrix}$$



Hogsgaard-Lauritzen, 2008

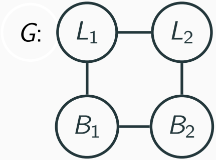
$$\mathrm{wmlt}(G) = \mathrm{mlt}(G) = 3$$
$$\mathrm{wmlt}(\mathcal{G}) = \mathrm{mlt}(\mathcal{G}) = 1$$

And yet another example: Frets' heads

Study of heredity of heads dimensions (Freds, 1921): length and breath of 25 pairs of first and second sons.

And yet another example: Frets' heads

Study of heredity of heads dimensions (Freds, 1921): length and breath of 25 pairs of first and second sons.

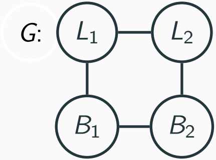


$$K_G = \begin{pmatrix} \lambda_1 & \lambda_5 & 0 & \lambda_8 \\ \lambda_5 & \lambda_2 & \lambda_6 & 0 \\ 0 & \lambda_6 & \lambda_3 & \lambda_7 \\ \lambda_8 & 0 & \lambda_7 & \lambda_4 \end{pmatrix}$$

$\mathrm{wmlt}(G) = 2,$
$\mathrm{mlt}(G) = 3$
Buhl, 1993

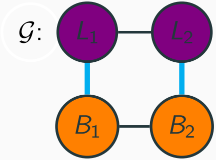
And yet another example: Frets' heads

Study of heredity of heads dimensions (Freds, 1921): length and breath of 25 pairs of first and second sons.



$$K_G = \begin{pmatrix} \lambda_1 & \lambda_5 & 0 & \lambda_8 \\ \lambda_5 & \lambda_2 & \lambda_6 & 0 \\ 0 & \lambda_6 & \lambda_3 & \lambda_7 \\ \lambda_8 & 0 & \lambda_7 & \lambda_4 \end{pmatrix}$$

$\mathrm{wmlt}(G) = 2$,
$\mathrm{mlt}(G) = 3$
Buhl, 1993



$$K_{\mathcal{G}} = \begin{pmatrix} \lambda_1 & \lambda_5 & 0 & \lambda_8 \\ \lambda_5 & \lambda_2 & \lambda_6 & 0 \\ 0 & \lambda_6 & \lambda_3 & \lambda_7 \\ \lambda_8 & 0 & \lambda_7 & \lambda_4 \end{pmatrix}$$

$\mathrm{wmlt}(\mathcal{G}) = 1$,
$\mathrm{mlt}(\mathcal{G}) = 2$
Uhler, 2012

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.
The *colored Gaussian graphical model* $\mathcal{K}_{\mathcal{G}}$ associated to $\mathcal{G}$ is the set of all concentration matrices $K \in \mathrm{PD}_n$ such that

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.
The *colored Gaussian graphical model* $\mathcal{K}_{\mathcal{G}}$ associated to $\mathcal{G}$ is the set of all concentration matrices $K \in \mathrm{PD}_n$ such that

- $K_{ij} = 0$ for all $\{i, j\} \notin E(\mathcal{G})$ with $i \neq j$,

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.
The *colored Gaussian graphical model* $\mathcal{K}_{\mathcal{G}}$ associated to $\mathcal{G}$ is the set of all concentration matrices $K \in \mathrm{PD}_n$ such that

- $K_{ij} = 0$ for all $\{i, j\} \notin E(\mathcal{G})$ with $i \neq j$,
- $K_{ii} = K_{jj}$ whenever $\lambda(i) = \lambda(j)$, and

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.
The *colored Gaussian graphical model* $\mathcal{K}_{\mathcal{G}}$ associated to $\mathcal{G}$ is the set of all concentration matrices $K \in \mathrm{PD}_n$ such that

- $K_{ij} = 0$ for all $\{i, j\} \notin E(\mathcal{G})$ with $i \neq j$,
- $K_{ii} = K_{jj}$ whenever $\lambda(i) = \lambda(j)$, and
- $K_{ij} = K_{k\ell}$ whenever $\{i, j\}, \{k, \ell\} \in E(\mathcal{G})$ and $\lambda(\{i, j\}) = \lambda(\{k, \ell\})$.

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.
The *colored Gaussian graphical model* $\mathcal{K}_{\mathcal{G}}$ associated to $\mathcal{G}$ is the set of all concentration matrices $K \in \mathrm{PD}_n$ such that

- $K_{ij} = 0$ for all $\{i, j\} \notin E(\mathcal{G})$ with $i \neq j$,
- $K_{ii} = K_{jj}$ whenever $\lambda(i) = \lambda(j)$, and
- $K_{ij} = K_{k\ell}$ whenever $\{i, j\}, \{k, \ell\} \in E(\mathcal{G})$ and $\lambda(\{i, j\}) = \lambda(\{k, \ell\})$.

$\mathcal{L}_{\mathcal{G}} := \overline{\mathcal{K}_{\mathcal{G}}}$ d-dimensional linear space of symmetric matrices.

Colored Gaussian graphical models

$G = (V, E)$ graph with vertex set $V = [m]$,
$\lambda : V \sqcup E \twoheadrightarrow [d]$ coloring of the vertices and edges of G for some $d \in \mathbb{N}$,
$\Lambda(V)$ and $\Lambda(E)$ set of vertex colors and edge colors, resp.
We call $\mathcal{G} = (G, \lambda)$ a *colored graph* with underlying graph G and coloring λ.
The *colored Gaussian graphical model* $\mathcal{K}_{\mathcal{G}}$ associated to $\mathcal{G}$ is the set of all concentration matrices $K \in \mathrm{PD}_n$ such that

- $K_{ij} = 0$ for all $\{i, j\} \notin E(\mathcal{G})$ with $i \neq j$,
- $K_{ii} = K_{jj}$ whenever $\lambda(i) = \lambda(j)$, and
- $K_{ij} = K_{k\ell}$ whenever $\{i, j\}, \{k, \ell\} \in E(\mathcal{G})$ and $\lambda(\{i, j\}) = \lambda(\{k, \ell\})$.

$\mathcal{L}_{\mathcal{G}} := \overline{\mathcal{K}_{\mathcal{G}}}$ d-dimensional linear space of symmetric matrices.

$\mathcal{G}$:

$$\mathcal{L}_{\mathcal{G}} = \left\{ \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_1 \end{pmatrix} \in \mathrm{Sym}(2) \mid \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

Existence of the maximum likelihood estimate

$S = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^t$ sample covariance matrix, $x_1, \ldots, x_n \in \mathbb{R}^m$.

MLE for the covariance matrix $\hat{\Sigma} = \left(\hat{K}\right)^{-1}$

$$\hat{K} = \begin{array}{cc} \text{arg max}_K & \log\det K - \langle S, K\rangle, \\ \text{subject to} & K \in \mathcal{K}_G \end{array} \qquad \langle S, K\rangle := \mathrm{tr}(SK) = \sum_{1\le i,j\le m} s_{ij}k_{ij}$$

Existence of the maximum likelihood estimate

$S = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^t$ sample covariance matrix, $x_1, \ldots, x_n \in \mathbb{R}^m$.

MLE for the covariance matrix $\hat{\Sigma} = \left(\hat{K}\right)^{-1}$

$$\hat{K} = \begin{array}{cc} \arg\max_K & \log\det K - \langle S, K\rangle, \\ \text{subject to} & K \in \mathcal{K}_G \end{array} \qquad \langle S, K\rangle := \mathrm{tr}(SK) = \sum_{1\le i,j\le m} s_{ij}k_{ij}$$

Sufficient statistics:

$$\begin{array}{rccc} \pi_{\mathcal{G}}: & \mathrm{Sym}(m) & \longrightarrow & \mathbb{R}^d \\ & S & \longmapsto & (\langle S, K_1\rangle, \ldots, \langle S, K_d\rangle) \end{array}$$

$K_1, \ldots, K_d$ basis of $\mathcal{L}_G$

Existence of the maximum likelihood estimate

$S = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^t$ sample covariance matrix, $x_1, \ldots, x_n \in \mathbb{R}^m$.

MLE for the covariance matrix $\hat{\Sigma} = \left(\hat{K}\right)^{-1}$

$$\hat{K} = \begin{array}{cc} \text{arg max}_K & \log\det K - \langle S, K\rangle, \\ \text{subject to} & K \in \mathcal{K}_G \end{array} \qquad \langle S, K\rangle := \mathrm{tr}(SK) = \sum_{1\le i,j\le m} s_{ij}k_{ij}$$

Sufficient statistics:

$$\begin{array}{rrcl} \pi_{\mathcal{G}}: & \mathrm{Sym}(2) & \longrightarrow & \mathbb{R}^2 \\ & S & \longmapsto & (s_{11}+s_{22}, 2s_{12}) \end{array}$$

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ basis of $\mathcal{L}_G$

Existence of the maximum likelihood estimate

$S = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^t$ sample covariance matrix, $x_1, \ldots, x_n \in \mathbb{R}^m$.

MLE for the covariance matrix $\hat{\Sigma} = \left(\hat{K}\right)^{-1}$

$$\hat{K} = \begin{array}{cc} \text{arg max}_K & \log\det K - \langle S, K\rangle, \\ \text{subject to} & K \in \mathcal{K}_G \end{array} \qquad \langle S, K\rangle := \mathrm{tr}(SK) = \sum_{1\le i,j\le m} s_{ij}k_{ij}$$

Sufficient statistics:

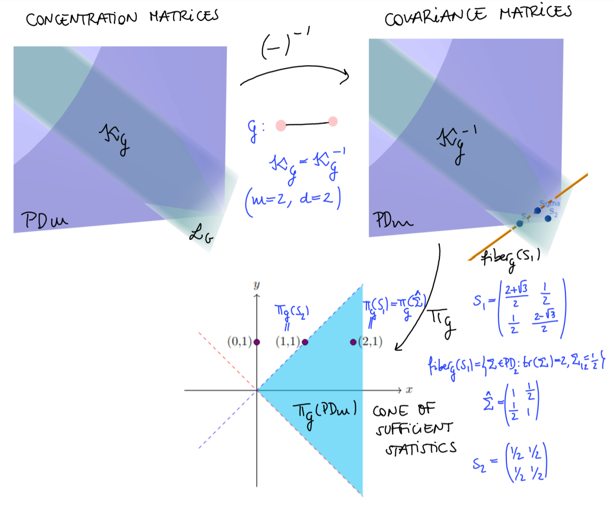
$$\begin{array}{rccc} \pi_{\mathcal{G}}: & \mathrm{Sym}(m) & \longrightarrow & \mathbb{R}^d \\ & S & \longmapsto & (\langle S, K_1\rangle, \ldots, \langle S, K_d\rangle) \end{array}$$

$K_1, \ldots, K_d$ basis of $\mathcal{L}_G$

Theorem The MLE $\hat{\Sigma}$ exists for a sample covariance matrix S iff

$$\text{fiber}_{\mathcal{G}}(S) := \{\Sigma \in \mathrm{PD}_m \mid \pi_{\mathcal{G}}(\Sigma) = \pi_{\mathcal{G}}(S)\} \neq \emptyset.$$

Then the MLE $\hat{\Sigma}$ is the unique matrix in $\text{fiber}_{\mathcal{G}}(S)$ such that $\hat{\Sigma}^{-1} \in \mathcal{K}_G$.


CONCENTRATION MATRICES
COVARIANCE MATRICES
PDm
PDm
(0,1)
(1,1)
(2,1)
CONE OF SUFFICIENT STATISTICS

An algebraic relaxation: the generic completion rank

Let $\mathcal{G}$ be a colored graph on $[m]$. The *generic completion rank* $\mathrm{gcr}(\mathcal{G})$ is the smallest n such that $\dim \pi_{\mathcal{G}}(\mathrm{Sym(m, n)}) = d$.

An algebraic relaxation: the generic completion rank

Let $\mathcal{G}$ be a colored graph on $[m]$. The *generic completion rank* $\mathrm{gcr}(\mathcal{G})$ is the smallest n such that $\dim \pi_{\mathcal{G}}(\mathrm{Sym(m, n)}) = d$.

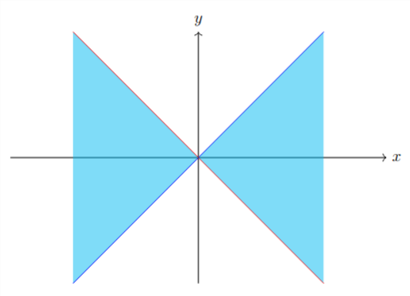


Figure 1: $\pi_{\mathcal{G}}(\mathrm{Sym}(2, 1))$ for the 2-cycle with a single vertex color.

An algebraic relaxation: the generic completion rank

Let $\mathcal{G}$ be a colored graph on $[m]$. The *generic completion rank* $\mathrm{gcr}(\mathcal{G})$ is the smallest n such that $\dim \pi_{\mathcal{G}}(\mathrm{Sym(m,n)}) = d$.

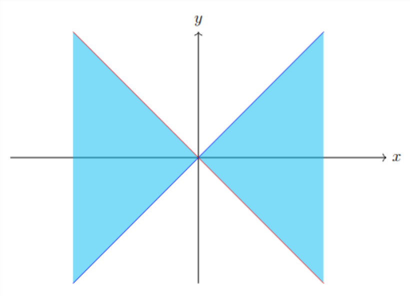


Figure 1: $\pi_{\mathcal{G}}(\mathrm{Sym}(2,1))$ for the 2-cycle with a single vertex color.

Theorem (Uhler, 2012) $\mathrm{mlt}(\mathcal{G}) \leq \mathrm{gcr}(\mathcal{G})$

Matrix completion problems

For $S = \begin{pmatrix} s_{11} & 1/2 \\ 1/2 & s_{22} \end{pmatrix}$, the MLE is $\hat{\Sigma} = \begin{pmatrix} \mathrm{tr}(S)/2 & 1/2 \\ 1/2 & \mathrm{tr}(S)/2 \end{pmatrix}$ if it exists.

Matrix completion problems

For $S = \begin{pmatrix} s_{11} & 1/2 \\ 1/2 & s_{22} \end{pmatrix}$, the MLE is $\hat{\Sigma} = \begin{pmatrix} \mathrm{tr}(S)/2 & 1/2 \\ 1/2 & \mathrm{tr}(S)/2 \end{pmatrix}$ if it exists.

Positive definite completions:

- $\mathrm{tr}(S) = 2$: $\text{fiber}_{\mathcal{G}}(S) = \left\{ \begin{pmatrix} 1 \pm a & 1/2 \\ 1/2 & 1 \mp a \end{pmatrix} : 0 \leq a < \sqrt{3}/2 \right\}$
- $\mathrm{tr}(S) = 0, 1$: $\text{fiber}_{\mathcal{G}}(S) = \emptyset$

Matrix completion problems

For $S = \begin{pmatrix} s_{11} & 1/2 \\ 1/2 & s_{22} \end{pmatrix}$, the MLE is $\hat{\Sigma} = \begin{pmatrix} \mathrm{tr}(S)/2 & 1/2 \\ 1/2 & \mathrm{tr}(S)/2 \end{pmatrix}$ if it exists.

Positive definite completions:

- $\mathrm{tr}(S) = 2$: $\text{fiber}_{\mathcal{G}}(S) = \left\{ \begin{pmatrix} 1 \pm a & 1/2 \\ 1/2 & 1 \mp a \end{pmatrix} : 0 \leq a < \sqrt{3}/2 \right\}$
- $\mathrm{tr}(S) = 0, 1$: $\text{fiber}_{\mathcal{G}}(S) = \emptyset$

Low rank completions (rank 1):

- $\mathrm{tr}(S) = 2$: $\begin{pmatrix} (2 \pm \sqrt{3})/2 & 1/2 \\ 1/2 & (2 \mp \sqrt{3})/2 \end{pmatrix}$
- $\mathrm{tr}(S) = 1$: $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$
- $\mathrm{tr}(S) = 0$: no real rank 1 completion but $\begin{pmatrix} \pm i/2 & 1/2 \\ 1/2 & \mp i/2 \end{pmatrix}$

Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with $\mathrm{mlt}(\mathcal{G}) < \mathrm{gcr}(\mathcal{G})$?

Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with $\mathrm{mlt}(\mathcal{G}) < \mathrm{gcr}(\mathcal{G})$?
  Question solved in the uncolored setting by Bleckherman-Sinn: $K_{5,5}$

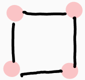
Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with $\mathrm{mlt}(\mathcal{G}) < \mathrm{gcr}(\mathcal{G})$?

  Question solved in the uncolored setting by Bleckherman-Sinn: $K_{5,5}$

  If $\mathcal{G}$ is the 4-cycle with $|\Lambda(V)| = 1$ and $|\Lambda(E)| = |E|$,

  - $\mathrm{gcr}(\mathcal{G}) = 2$,
  - $\mathrm{mlt}(\mathcal{G}) = 1$.



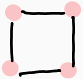
Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with $\mathrm{mlt}(\mathcal{G}) < \mathrm{gcr}(\mathcal{G})$?

  Question solved in the uncolored setting by Bleckherman-Sinn: $K_{5,5}$

  If $\mathcal{G}$ is the 4-cycle with $|\Lambda(V)| = 1$ and $|\Lambda(E)| = |E|$,

  - $\mathrm{gcr}(\mathcal{G}) = 2$,
  - $\mathrm{mlt}(\mathcal{G}) = 1$.



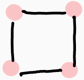
- Are there graphs with bounded $\mathrm{mlt}(\mathcal{G})$ and arbitrarily high $\mathrm{gcr}(\mathcal{G})$?

Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with $\mathrm{mlt}(\mathcal{G}) < \mathrm{gcr}(\mathcal{G})$?

  Question solved in the uncolored setting by Bleckherman-Sinn: $K_{5,5}$

  If $\mathcal{G}$ is the 4-cycle with $|\Lambda(V)| = 1$ and $|\Lambda(E)| = |E|$,

  - $\mathrm{gcr}(\mathcal{G}) = 2$,
  - $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with bounded $\mathrm{mlt}(\mathcal{G})$ and arbitrarily high $\mathrm{gcr}(\mathcal{G})$?

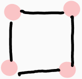
  Question posed in the uncolored setting by Bleckherman-Sinn

Is gcr a good bound for mlt?

- Trivial case: If $\mathrm{gcr}(\mathcal{G}) = 1$, then $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with $\mathrm{mlt}(\mathcal{G}) < \mathrm{gcr}(\mathcal{G})$?

  Question solved in the uncolored setting by Bleckherman-Sinn: $K_{5,5}$

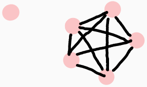
  If $\mathcal{G}$ is the 4-cycle with $|\Lambda(V)| = 1$ and $|\Lambda(E)| = |E|$,

  - $\mathrm{gcr}(\mathcal{G}) = 2$,
  - $\mathrm{mlt}(\mathcal{G}) = 1$.
- Are there graphs with bounded $\mathrm{mlt}(\mathcal{G})$ and arbitrarily high $\mathrm{gcr}(\mathcal{G})$?

  Question posed in the uncolored setting by Bleckherman-Sinn

  If $\mathcal{G}$ is the disjoint union of a vertex and an m-complete graph with $|\Lambda(V)| = 1$ and $|\Lambda(E)| = |E|$,

  - $\mathrm{gcr}(\mathcal{G}) \geq \lceil (m-2)/2 \rceil$,
  - $\mathrm{mlt}(\mathcal{G}) = 1$.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i, c) := \{j \in V \mid \lambda(i, j) = c\} \subset N(i)$.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i,c) := \{j \in V \mid \lambda(i,j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)
If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

(3) for $c \in \Lambda(E)$, $N(i,c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \setminus \{\varphi(c)\}, \end{cases}$

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i, c) := \{j \in V \mid \lambda(i, j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)
If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

(3) for $c \in \Lambda(E)$, $N(i, c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \backslash \{\varphi(c)\}, \end{cases}$

injectivity necessary,
(1)&(2) not sufficient,
(3) not necessary

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i,c) := \{j \in V \mid \lambda(i,j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)
If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

injectivity necessary,
(1)&(2) not sufficient,
(3) not necessary

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

(3) for $c \in \Lambda(E)$, $N(i,c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \backslash \{\varphi(c)\}, \end{cases}$

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Idea of the proof: Show that the Jacobian of $\pi_{\mathcal{G}} \mid_{Sym(m,1)}$ is full rank.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i,c) := \{j \in V \mid \lambda(i,j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)

If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

injectivity necessary,
(1)&(2) not sufficient,
(3) not necessary

(3) for $c \in \Lambda(E)$, $N(i,c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \setminus \{\varphi(c)\}, \end{cases}$

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Idea of the proof: Show that the Jacobian of $\pi_{\mathcal{G}} \mid_{Sym(m,1)}$ is full rank.

Corollary

- For $|\Lambda(E)| = 1$: $\mathrm{gcr}(\mathcal{G}) = 1$ if $|\Lambda(V)| < m$.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i,c) := \{j \in V \mid \lambda(i,j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)
If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

injectivity necessary,
(1)&(2) not sufficient,
(3) not necessary

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

(3) for $c \in \Lambda(E)$, $N(i,c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \setminus \{\varphi(c)\}, \end{cases}$

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Idea of the proof: Show that the Jacobian of $\pi_{\mathcal{G}} |_{Sym(m,1)}$ is full rank.

Corollary

- For $|\Lambda(E)| = 1$: $\mathrm{gcr}(\mathcal{G}) = 1$ if $|\Lambda(V)| < m$. if and only if

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i,c) := \{j \in V \mid \lambda(i,j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)
If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

injectivity necessary,
(1)&(2) not sufficient,
(3) not necessary

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

(3) for $c \in \Lambda(E)$, $N(i,c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \setminus \{\varphi(c)\}, \end{cases}$

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Idea of the proof: Show that the Jacobian of $\pi_{\mathcal{G}}\,|_{Sym(m,1)}$ is full rank.

Corollary

- For $|\Lambda(E)| = 1$: $\mathrm{gcr}(\mathcal{G}) = 1$ if $|\Lambda(V)| < m$. if and only if
- For $|\Lambda(V)| = 1$: $\mathrm{gcr}(\mathcal{G}) = 1$ if there is a vertex adjacent to all edge colors.

Which colorings give $\mathrm{gcr}(\mathcal{G}) = 1$?

For $i \in V$ and $c \in \Lambda(E)$, $N(i,c) := \{j \in V \mid \lambda(i,j) = c\} \subset N(i)$.

Proposition (Deligeorgaki, H., Johnson, Kagy, Maraj 25+)
If there exists an injective map $\varphi : \quad \Lambda(V) \sqcup \Lambda(E) \quad \longrightarrow \quad V$ s.t.

injectivity necessary,
(1)&(2) not sufficient,
(3) not necessary

(1) for $c \in \Lambda(V)$, $\lambda(\varphi(c)) = c$,

(2) for $c \in \Lambda(E)$, $N(\varphi(c), c) \neq \emptyset$,

(3) for $c \in \Lambda(E)$, $N(i,c) \cap \varphi(\Lambda(V)) = \begin{cases} \neq \emptyset, & \text{if } \varphi(c) = i, \\ = \emptyset, & \text{if } i \in \varphi(\Lambda(E)) \setminus \{\varphi(c)\}, \end{cases}$

then $\mathrm{gcr}(\mathcal{G}) = 1$.

Idea of the proof: Show that the Jacobian of $\pi_{\mathcal{G}} \mid_{Sym(m,1)}$ is full rank.

Corollary

- For $|\Lambda(E)| = 1$: $\mathrm{gcr}(\mathcal{G}) = 1$ if $|\Lambda(V)| < m$. if and only if
- For $|\Lambda(V)| = 1$: $\mathrm{gcr}(\mathcal{G}) = 1$ if there is a vertex adjacent to all edge colors. not necessary

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{t_i - \langle S, K_i \rangle\}_{1 \leq i \leq d} \right\rangle \right) \cap \mathbb{R}[t_1, \ldots, t_d].$$

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{ t_i - \langle S, K_i \rangle \}_{1 \leq i \leq d} \right\rangle \right) \cap \mathbb{R}[t_1, \ldots, t_d].$$

If $I_{\mathcal{G},n} = (0)$, the MLE exists for n observations with probability 1

$$\mathrm{wmlt}(\mathcal{G}) \leq \mathrm{mlt}(\mathcal{G}) \leq \mathrm{gcr}(\mathcal{G}) \leq n$$

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{ t_i - \langle S, K_i \rangle \}_{1 \leq i \leq d} \right\rangle \right) \cap \mathbb{R}[t_1, \ldots, t_d].$$

If $I_{\mathcal{G},n} = (0)$, the MLE exists for n observations with probability 1

$$\mathrm{wmlt}(\mathcal{G}) \leq \mathrm{mlt}(\mathcal{G}) \leq \mathrm{gcr}(\mathcal{G}) \leq n$$

If $I_{\mathcal{G},n} \neq (0)$, the MLE for n observations... $\mathrm{gcr}(\mathcal{G}) > n$

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{t_i - \langle S, K_i \rangle\}_{1\leq i\leq d} \right\rangle\right) \cap \mathbb{R}[t_1, \ldots, t_d].$$

If $I_{\mathcal{G},n} = (0)$, the MLE exists for n observations with probability 1

$$\mathrm{wmlt}(\mathcal{G}) \leq \mathrm{mlt}(\mathcal{G}) \leq \mathrm{gcr}(\mathcal{G}) \leq n$$

If $I_{\mathcal{G},n} \neq (0)$, the MLE for n observations... $\mathrm{gcr}(\mathcal{G}) > n$

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{ t_i - \langle S, K_i \rangle \}_{1 \le i \le d} \right\rangle \right) \cap \mathbb{R}[t_1, \ldots, t_d].$$

If $I_{\mathcal{G},n} = (0)$, the MLE exists for n observations with probability 1

$$\mathrm{wmlt}(\mathcal{G}) \le \mathrm{mlt}(\mathcal{G}) \le \mathrm{gcr}(\mathcal{G}) \le n$$

If $I_{\mathcal{G},n} \neq (0)$, the MLE for n observations... $\mathrm{gcr}(\mathcal{G}) > n$

- ...exists with probability 1 $\mathrm{wmlt}(\mathcal{G}) \le \mathrm{mlt}(\mathcal{G}) \le n$

4-cycle with single vertex color for $n = 1$

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{ t_i - \langle S, K_i \rangle \}_{1 \le i \le d} \right\rangle \right) \cap \mathbb{R}[t_1, \dots, t_d].$$

If $I_{\mathcal{G},n} = (0)$, the MLE exists for n observations with probability 1

$\mathrm{wmlt}(\mathcal{G}) \le \mathrm{mlt}(\mathcal{G}) \le \mathrm{gcr}(\mathcal{G}) \le n$

If $I_{\mathcal{G},n} \neq (0)$, the MLE for n observations... $\mathrm{gcr}(\mathcal{G}) > n$

- ...exists with probability 1 $\mathrm{wmlt}(\mathcal{G}) \le \mathrm{mlt}(\mathcal{G}) \le n$

  4-cycle with single vertex color for $n = 1$
- ...exists with probability strictly between 0 and 1

  $\mathrm{mlt}(\mathcal{G}) > n$, $\mathrm{wmlt}(\mathcal{G}) \le n$

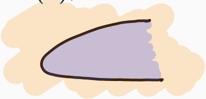
  3-cycle with single vertex color for $n = 1$

MLE existence via the ideal of the projection

Ideal of the projection via $\pi_{\mathcal{G}}$ of rank n matrices:

$$I_{\mathcal{G},n} := \left(I_{n+1}(S) + \left\langle \{t_i - \langle S, K_i \rangle\}_{1 \le i \le d} \right\rangle \right) \cap \mathbb{R}[t_1, \dots, t_d].$$

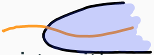
If $I_{\mathcal{G},n} = (0)$, the MLE exists for n observations with probability 1



$\mathrm{wmlt}(\mathcal{G}) \le \mathrm{mlt}(\mathcal{G}) \le \mathrm{gcr}(\mathcal{G}) \le n$

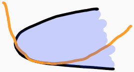
If $I_{\mathcal{G},n} \neq (0)$, the MLE for n observations... $\mathrm{gcr}(\mathcal{G}) > n$

- ...exists with probability 1 $\mathrm{wmlt}(\mathcal{G}) \le \mathrm{mlt}(\mathcal{G}) \le n$



4-cycle with single vertex color for $n = 1$

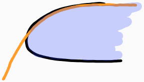
- ...exists with probability strictly between 0 and 1



$\mathrm{mlt}(\mathcal{G}) > n$, $\mathrm{wmlt}(\mathcal{G}) \le n$

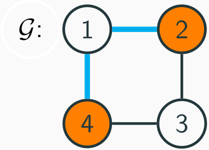
3-cycle with single vertex color for $n = 1$

- ...never exists $\mathrm{mlt}(\mathcal{G}) \ge \mathrm{wmlt}(\mathcal{G}) > n$



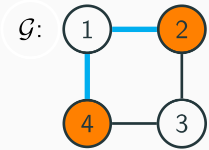
3-cycle with two vertex color for $n = 1$

Computing MLT via elimination ideal - an example



Sufficient statistics: $t_1 = s_{1,1}$,
$t_2 = s_{2,2} + s_{4,4}$, $t_3 = s_{3,3}$, $t_4 = 2s_{1,2} + 2s_{1,4}$,
$t_5 = 2s_{2,3}$, $t_6 = 2s_{3,4}$.

Computing MLT via elimination ideal - an example

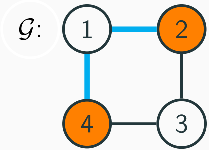


Sufficient statistics: $t_1 = s_{1,1}$, $t_2 = s_{2,2} + s_{4,4}$, $t_3 = s_{3,3}$, $t_4 = 2s_{1,2} + 2s_{1,4}$, $t_5 = 2s_{2,3}$, $t_6 = 2s_{3,4}$.

Ideals of the projection $\pi_{\mathcal{G}}$: $I_{\mathcal{G},2} = I_{\mathcal{G},3} = 0$, and

$$I_{\mathcal{G},1} = \langle t_3 t_4^2 - t_1 t_5^2 - 2t_1 t_5 t_6 - t_1 t_6^2, 4t_2 t_3 - t_5^2 - t_6^2 \rangle.$$

Computing MLT via elimination ideal - an example



Sufficient statistics: $t_1 = s_{1,1}$, $t_2 = s_{2,2} + s_{4,4}$, $t_3 = s_{3,3}$, $t_4 = 2s_{1,2} + 2s_{1,4}$, $t_5 = 2s_{2,3}$, $t_6 = 2s_{3,4}$.

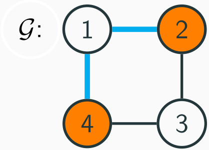
Ideals of the projection $\pi_{\mathcal{G}}$: $I_{\mathcal{G},2} = I_{\mathcal{G},3} = 0$, and

$$I_{\mathcal{G},1} = \langle t_3 t_4^2 - t_1 t_5^2 - 2t_1 t_5 t_6 - t_1 t_6^2, 4t_2 t_3 - t_5^2 - t_6^2 \rangle.$$

Cholesky decomposition $S = LL^t$:

$$\begin{aligned} f_2(l_{ij}) = 4(-l_{13}l_{22} + l_{12}l_{23})^2 + 4(l_{12}l_{33})^2 + 4(-l_{14}l_{23} + l_{13}l_{24})^2 \\ + 4(-l_{14}l_{33} + l_{13}l_{34})^2 + 4(l_{13}l_{44})^2 + 4(l_{22}l_{33})^2 \\ + 4(-l_{24}l_{33} + l_{23}l_{34})^2 + 4(l_{23}l_{44})^2 + 4(l_{33}l_{44})^2 > 0. \end{aligned}$$

Computing MLT via elimination ideal - an example



Sufficient statistics: $t_1 = s_{1,1}$, $t_2 = s_{2,2} + s_{4,4}$, $t_3 = s_{3,3}$, $t_4 = 2s_{1,2} + 2s_{1,4}$, $t_5 = 2s_{2,3}$, $t_6 = 2s_{3,4}$.

Ideals of the projection $\pi_{\mathcal{G}}$: $I_{\mathcal{G},2} = I_{\mathcal{G},3} = 0$, and

$$I_{\mathcal{G},1} = \langle t_3 t_4^2 - t_1 t_5^2 - 2t_1 t_5 t_6 - t_1 t_6^2, 4t_2 t_3 - t_5^2 - t_6^2 \rangle.$$

Cholesky decomposition $S = LL^t$:

$$\begin{aligned} f_2(l_{ij}) = {} & 4(-l_{13}l_{22} + l_{12}l_{23})^2 + 4(l_{12}l_{33})^2 + 4(-l_{14}l_{23} + l_{13}l_{24})^2 \\ & + 4(-l_{14}l_{33} + l_{13}l_{34})^2 + 4(l_{13}l_{44})^2 + 4(l_{22}l_{33})^2 \\ & + 4(-l_{24}l_{33} + l_{23}l_{34})^2 + 4(l_{23}l_{44})^2 + 4(l_{33}l_{44})^2 > 0. \end{aligned}$$

$$\mathrm{wmlt}(\mathcal{G}) = \mathrm{mlt}(\mathcal{G}) = \mathrm{gcr}(\mathcal{G}) = 2$$

Reinterpretation of thresholds: generalizing Bleckherman-Sinn

generic completion rank

Theorem The maximum likelihood threshold of $\mathcal{G}$ is the smallest

weak maximum likelihood threshold

$\mathcal{L}_{\mathcal{G}}$

n for which there is no matrix $K \neq 0$ in $\mathcal{L}_{\mathcal{G}} \cap \mathrm{PSD}_m$ such that

$\mathcal{L}_{\mathcal{G}} \cap \mathrm{PSD}_m$

some l.i.

$x_i \in \ker(K)$ for generic observations $x_1, \ldots, x_n \in \mathbb{R}^m$.

some l.i.

Reinterpretation of thresholds: generalizing Bleckherman-Sinn

generic completion rank

Theorem The maximum likelihood threshold of $\mathcal{G}$ is the smallest

weak maximum likelihood threshold

is equal to Sym_m

n for which the vector space $\langle x_1, \ldots, x_n \rangle_2 + \mathcal{L}_G^\perp$ contains a PD matrix

contains a PD matrix

some

for generic observations $x_1, \ldots, x_n \in \mathbb{R}^m$.

some

Reinterpretation of thresholds: generalizing Bleckherman-Sinn

generic completion rank

Theorem The maximum likelihood threshold of $\mathcal{G}$ is the smallest

weak maximum likelihood threshold

is equal to Sym_m

n for which the vector space $\langle x_1, \ldots, x_n \rangle_2 + \mathcal{L}_G^\perp$ contains a PD matrix

contains a PD matrix

some

for generic observations $x_1, \ldots, x_n \in \mathbb{R}^m$.

some

$$I_{\mathcal{G}} := \langle \{(y_1 \; \ldots \; y_m) K_i (y_1 \; \ldots \; y_m)^t\}_{d+1<i\leq m(m-1)/2} \rangle, \; R := \mathbb{R}[y_1, \ldots, y_m]/I_{\mathcal{G}},$$

$K_{d+1}, \ldots, K_{m(m-1)/2}$ basis of $\mathcal{L}_G^\perp$

$\ell_i = x_i^{(1)} y_1 + \cdots + x_i^{(m)} y_m \in R_1$ for some $x_i^{(j)} \in \mathbb{R}$

$\langle \ell_1, \ldots, \ell_n \rangle_2 \in R_2 \simeq \mathrm{Sym}(m)/\mathcal{L}_{\mathcal{G}}^\perp$,

$x_i = \left(x_i^{(1)}, \ldots, x_i^{(m)} \right)^t \in \mathbb{R}^m$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored *m*-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)

If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Sketch of the proof:

- $\mathrm{gcr}(\mathcal{G}) = 2$:
    - $\mathrm{gcr}(\mathcal{G}) \geq 2$: $\dim\langle \ell \rangle_2 < \dim R_2$ for any $l \in R_1$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Sketch of the proof:

- $\mathrm{gcr}(\mathcal{G}) = 2$:
  - $\mathrm{gcr}(\mathcal{G}) \geq 2$: $\dim\langle \ell \rangle_2 < \dim R_2$ for any $l \in R_1$.
  - $\mathrm{gcr}(\mathcal{G}) \leq 2$: $\langle \ell_1, \ell_2 \rangle_2 = R_2$ for $\ell_1 = \sum_{i=1}^{\lfloor m/2 \rfloor} y_i$, $\ell_2 = y_1 + \sum_{\lfloor m/2 \rfloor}^{m-1} y_i$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)|= 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Sketch of the proof:

- $\mathrm{gcr}(\mathcal{G}) = 2$:
  - $\mathrm{gcr}(\mathcal{G}) \geq 2$: $\dim\langle\ell\rangle_2 < \dim R_2$ for any $l \in R_1$.
  - $\mathrm{gcr}(\mathcal{G}) \leq 2$: $\langle\ell_1, \ell_2\rangle_2 = R_2$ for $\ell_1 = \sum_{i=1}^{\lfloor m/2\rfloor} y_i$, $\ell_2 = y_1 + \sum_{\lfloor m/2\rfloor}^{m-1} y_i$.
- $\mathrm{mlt}(\mathcal{G}) = 1$:
  - Equivalent to proving that there is no $0 \neq K \in \mathcal{L}_G \cap \mathrm{PSD}_m$ that contains a generic column vector $(a_1 \ldots a_m)^t$ in its kernel.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Sketch of the proof:

- $\mathrm{gcr}(\mathcal{G}) = 2$:
  - $\mathrm{gcr}(\mathcal{G}) \geq 2$: $\dim\langle\ell\rangle_2 < \dim R_2$ for any $l \in R_1$.
  - $\mathrm{gcr}(\mathcal{G}) \leq 2$: $\langle \ell_1, \ell_2 \rangle_2 = R_2$ for $\ell_1 = \sum_{i=1}^{\lfloor m/2 \rfloor} y_i$, $\ell_2 = y_1 + \sum_{\lfloor m/2 \rfloor}^{m-1} y_i$.
- $\mathrm{mlt}(\mathcal{G}) = 1$:
  - Equivalent to proving that there is no $0 \neq K \in \mathcal{L}_G \cap \mathrm{PSD}_m$ that contains a generic column vector $(a_1 \dots a_m)^t$ in its kernel.
  - Any PSD matrix $K \neq 0$ in $\mathcal{L}_\mathcal{G} \cap \mathrm{PSD}_m$ satisfies $\lambda_1 \geq |\lambda_i|$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Sketch of the proof:

- $\mathrm{gcr}(\mathcal{G}) = 2$:
  - $\mathrm{gcr}(\mathcal{G}) \geq 2$: $\dim\langle \ell \rangle_2 < \dim R_2$ for any $l \in R_1$.
  - $\mathrm{gcr}(\mathcal{G}) \leq 2$: $\langle \ell_1, \ell_2 \rangle_2 = R_2$ for $\ell_1 = \sum_{i=1}^{\lfloor m/2 \rfloor} y_i$, $\ell_2 = y_1 + \sum_{\lfloor m/2 \rfloor}^{m-1} y_i$.
- $\mathrm{mlt}(\mathcal{G}) = 1$:
  - Equivalent to proving that there is no $0 \neq K \in \mathcal{L}_G \cap \mathrm{PSD}_m$ that contains a generic column vector $(a_1 \ldots a_m)^t$ in its kernel.
  - Any PSD matrix $K \neq 0$ in $\mathcal{L}_{\mathcal{G}} \cap \mathrm{PSD}_m$ satisfies $\lambda_1 \geq |\lambda_i|$.
  - Rewrite $K(a_1 \ldots a_m)^t = 0$ as $A\,(\lambda_1 \ldots \lambda_{m+1})^t = 0$.

4-cycle with a single vertex color

Proposition (H., Kuznetsova, Stolz 25+)
If $\mathcal{G}$ is a colored m-cycle with $|\Lambda(V)| = 1$, then $\mathrm{gcr}(\mathcal{G}) = 2$.
If m is even, then $\mathrm{mlt}(\mathcal{G}) = 1$.

Sketch of the proof:

- $\mathrm{gcr}(\mathcal{G}) = 2$:
  - $\mathrm{gcr}(\mathcal{G}) \geq 2$: $\dim\langle\ell\rangle_2 < \dim R_2$ for any $l \in R_1$.
  - $\mathrm{gcr}(\mathcal{G}) \leq 2$: $\langle \ell_1, \ell_2\rangle_2 = R_2$ for $\ell_1 = \sum_{i=1}^{\lfloor m/2\rfloor} y_i$, $\ell_2 = y_1 + \sum_{\lfloor m/2\rfloor}^{m-1} y_i$.
- $\mathrm{mlt}(\mathcal{G}) = 1$:
  - Equivalent to proving that there is no $0 \neq K \in \mathcal{L}_G \cap \mathrm{PSD}_m$ that contains a generic column vector $(a_1 \dots a_m)^t$ in its kernel.
  - Any PSD matrix $K \neq 0$ in $\mathcal{L}_\mathcal{G} \cap \mathrm{PSD}_m$ satisfies $\lambda_1 \geq |\lambda_i|$.
  - Rewrite $K(a_1 \dots a_m)^t = 0$ as $A\left(\lambda_1 \dots \lambda_{m+1}\right)^t = 0$.
  - If $a_1 \neq 0$, $\left(\sum_i^m (-1)^{i+1} a_i^2\right)\lambda_1 = 0$ for m even: $\lambda_1 = 0$.

Summary of techniques

For gcr:

- Compute ideal of the projection.

For mlt or wmlt:

- Studying the sign of the generators of the ideal of the projection.
- Studying the algebraic boundary of the cone of sufficient statistics.

Summary of techniques

For gcr:

- Compute ideal of the projection.
- Studying existence (or non-existence) of matrices in $\mathcal{L}_{\mathcal{G}}$ with n observations in its kernel.
- Studying existence (or non-existence) of n observations that span $\mathrm{Sym}(m)/\mathcal{L}_{\mathcal{G}}^{\perp}$ in degree 2.

For mlt or wmlt:

- Studying the sign of the generators of the ideal of the projection.
- Studying the algebraic boundary of the cone of sufficient statistics.
- Studying existence (or non-existence) of PSD matrices in $\mathcal{L}_{\mathcal{G}}$ with n observations in its kernel.
- Studying existence (or non-existence) of n observations whose degree 2 span contains a PD matrix.

Summary of techniques

For gcr:

- Compute ideal of the projection.
- Studying existence (or non-existence) of matrices in $\mathcal{L}_{\mathcal{G}}$ with n observations in its kernel.
- Studying existence (or non-existence) of n observations that span $\mathrm{Sym}(m)/\mathcal{L}_{\mathcal{G}}^{\perp}$ in degree 2.

For mlt or wmlt:

- Studying the sign of the generators of the ideal of the projection.
- Studying the algebraic boundary of the cone of sufficient statistics.
- Studying existence (or non-existence) of PSD matrices in $\mathcal{L}_{\mathcal{G}}$ with n observations in its kernel.
- Studying existence (or non-existence) of n observations whose degree 2 span contains a PD matrix.
- Perturbing matrices while preserving their sufficient statistics.

Summary of techniques

For gcr:

- Compute ideal of the projection.
- Studying existence (or non-existence) of matrices in $\mathcal{L}_{\mathcal{G}}$ with n observations in its kernel.
- Studying existence (or non-existence) of n observations that span $\mathrm{Sym}(m)/\mathcal{L}_{\mathcal{G}}^{\perp}$ in degree 2.
- Compute rank of the jacobian of $\pi_{\mathcal{G}}$ restricted to $\mathrm{Sym}(m, n)$.

For mlt or wmlt:

- Studying the sign of the generators of the ideal of the projection.
- Studying the algebraic boundary of the cone of sufficient statistics.
- Studying existence (or non-existence) of PSD matrices in $\mathcal{L}_{\mathcal{G}}$ with n observations in its kernel.
- Studying existence (or non-existence) of n observations whose degree 2 span contains a PD matrix.
- Perturbing matrices while preserving their sufficient statistics.

Local persistence homology

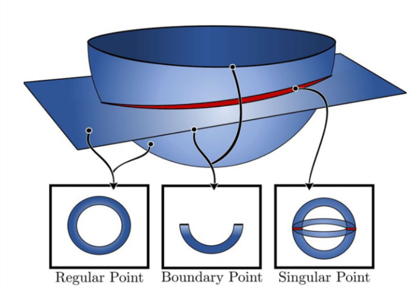
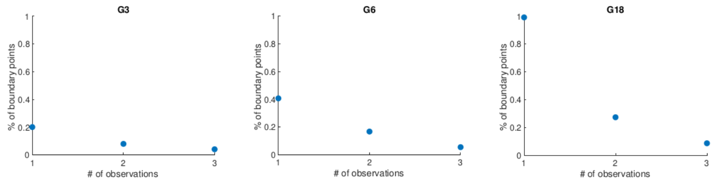


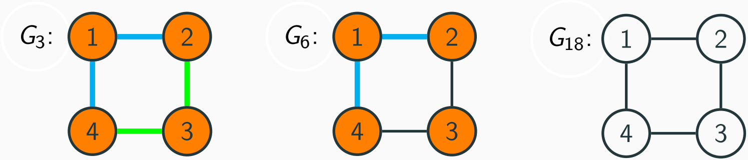
Figure 2: *Geometric anomaly detection in data*, Bernadette Stolz et al. PNAS (2020)

Performance



(A) Dim 1 homology (B) Dim 1&2 homology (C) Dim 2 homology

Figure 3: Percentage of boundary points in two colored 4-cycles and the uncolored 4-cycle. Points with less than 4 neighbours in their annular neighbourhoods are excluded.



$\text{wmlt}(G_3) = \text{mlt}(G_3) = 1$ $\text{wmlt}(G_6) = \text{mlt}(G_6) = 1$ $\text{wmlt}(G_{18}) = 2, \text{mlt}(G_{18}) = 3$

Thanks a lot!
Moltes gràcies!