Computing Maximum Likelihood Estimates for Gaussian Graphical Models with Macaulay2

joint work with Carlos Améndola, Luis David García Puente, Roser Homs, Olga Kuznetsova

Harshit J Motwani

Ghent University
Department of Mathematics: Algebra and Geometry

MEGA 2021, June 7-11, 2021
Overview

1. Graphical Models
2. Maximum Likelihood Estimation
3. Our Software Package
4. Examples
1. Graphical Models

- Graphical models are statistical models associated to graphs $G$. 
1. Graphical Models

- Graphical models are statistical models associated to graphs $G$.
- The nodes of $G$ represent random variables.
- The edges of $G$ encode dependencies among random variables.
1. Graphical Models

- Graphical models are statistical models associated to graphs $G$.
- The nodes of $G$ represent random variables.
- The edges of $G$ encode dependencies among random variables.
- $G$ is a mixed graph with undirected, directed and bidirected edges.

Figure: Food discovery in Uber Eats
• We denote mixed graphs by $G = (V, E)$ with undirected edges $i - j$, directed edges $i \rightarrow j$ and bidirected edges $i \leftrightarrow j$. 
We denote mixed graphs by $G = (V, E)$ with undirected edges $i - j$, directed edges $i \rightarrow j$ and bidirected edges $i \leftrightarrow j$. 

![Diagram of graphs](image)
Loopless mixed graph (LMG)

- An LMG is a mixed graph without loops or directed cycles.
2. Maximum Likelihood Estimation (MLE)

- MLE is one of the major estimation tools in statistics.

Let $X_1, X_2, \ldots, X_n \sim P_\theta$ be i.i.d r.v’s according to some probability distribution $P_\theta \in \mathcal{P}_\theta$.

Goal
Find $\theta$ that best describes the data.
2. Maximum Likelihood Estimation (MLE)

- MLE is one of the major estimation tools in statistics.
- Let $X^1, X^2, \ldots, X^n \sim P_\theta$ be i.i.d r.v’s according to some probability distribution $P_\theta \in \mathcal{P}_\theta$.

**Goal**

Find $\theta$ that best describes the data.
Multivariate Gaussian distribution

- A r.v. $X = (X_1, X_2, \cdots, X_m) \sim \mathcal{N}(\mu, \Sigma)$, if it follows this distribution:

$$p_{\mu, \Sigma}(x) = \frac{\exp \left( -\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu) \right)}{\sqrt{(2\pi)^m \det \Sigma}}, \quad x \in \mathbb{R}^m,$$

where $\mu \in \mathbb{R}^m$ and $\Sigma$ is assumed to be p.d.
Let $\mathcal{N}(\mu, \Sigma)$ be a multivariate Gaussian distribution.

Let $S$ be the sample covariance matrix.
Optimisation problem

- Let $\mathcal{N}(\mu, \Sigma)$ be a multivariate Gaussian distribution.
- Let $S$ be the sample covariance matrix.
- Computing MLE reduces to solving the following optimization problem:
• Let $\mathcal{N}(\mu, \Sigma)$ be a multivariate Gaussian distribution.
• Let $S$ be the sample covariance matrix.
• Computing MLE reduces to solving the following optimization problem:

$$\max_{\Sigma \in \mathbb{R}^{m \times m}} \ell(\Sigma) = - \log \det \Sigma - \text{tr}(S\Sigma^{-1})$$

subject to: $\Sigma \succ 0$
Gaussian Graphical Models

- Graphical models where r.v. $X = (X_v | v \in V) \sim \mathcal{N}(\mu, \Sigma)$
Gaussian Graphical Models

- Graphical models where r.v. $X = (X_{v} | v \in V) \sim \mathcal{N}(\mu, \Sigma)$
- Gaussian graphical models impose constraints on $\Sigma$
Gaussian Graphical Models

• Graphical models where r.v. $X = (X_v | v \in V) \sim \mathcal{N}(\mu, \Sigma)$
• Gaussian graphical models impose constraints on $\Sigma$

$$\Sigma = (I - \Lambda)^{-T} \begin{bmatrix} K^{-1} & 0 \\ 0 & \Psi \end{bmatrix} (I - \Lambda)^{-1},$$
Gaussian Graphical Models

- Graphical models where r.v. $X = (X_v | v \in V) \sim \mathcal{N}(\mu, \Sigma)$
- Gaussian graphical models impose constraints on $\Sigma$

$$\Sigma = (I - \Lambda)^{-1} \begin{bmatrix} K^{-1} & 0 \\ 0 & \Psi \end{bmatrix} (I - \Lambda)^{-1}$$

where

- $\Lambda = [\lambda_{ij}]$ is such that $\lambda_{ij} = 0 \iff i \rightarrow j \notin E$
- $K = [k_{ij}]$ is p.d. such that $k_{ij} = 0 \iff i - j \notin E$
- $\Psi = [\psi_{ij}]$ is p.d. such that $\psi_{ij} = 0 \iff i \leftrightarrow j \notin E$

- Example: $\lambda_{1,4} = k_{1,4} = \psi_{1,4} = \lambda_{1,2} = \psi_{1,2} = 0$
How to compute MLE?

- We have to maximize the following objective function:

\[
\max_{\Sigma \in \mathbb{R}^{m \times m}} \ell(\Sigma) = - \log \det \Sigma - \text{tr}(SS^{-1})
\]
How to compute MLE?

- We have to maximize the following objective function:

\[
\max_{\Sigma \in \mathbb{R}^{m \times m}} \ell(\Sigma) = -\log \det \Sigma - \text{tr}(S\Sigma^{-1})
\]

- To find critical points we first compute the following partial derivatives:

\[
- \frac{\partial}{\partial (\cdot)} \det \Sigma - \det \Sigma \frac{\partial}{\partial (\cdot)} \text{tr}(S\Sigma^{-1}).
\]

- Let \( I \) be the ideal generated by the above equations.
How to compute MLE?

- We have to maximize the following objective function:

$$
\max_{\Sigma \in \mathbb{R}^{m \times m}} \ell(\Sigma) = -\log \det \Sigma - \text{tr}(S\Sigma^{-1})
$$

- To find critical points we first compute the following partial derivatives:

$$
- \frac{\partial}{\partial (\cdot)} \det \Sigma - \det \Sigma \frac{\partial}{\partial (\cdot)} \text{tr}(S\Sigma^{-1}).
$$

- Let $I$ be the ideal generated by the above equations.
- Let $J$ be the saturation of $I$ with respect to $\det \Sigma$. 

How to compute MLE?

• We have to maximize the following objective function:

\[
\max_{\Sigma \in \mathbb{R}^{m \times m}} \ell(\Sigma) = -\log \det \Sigma - \text{tr}(S\Sigma^{-1})
\]

• To find critical points we first compute the following partial derivatives:

\[
-\frac{\partial}{\partial (\cdot)} \det \Sigma - \det \Sigma \frac{\partial}{\partial (\cdot)} \text{tr}(S\Sigma^{-1}).
\]

• Let I be the ideal generated by the above equations.
• Let J be the saturation of I with respect to \( \det \Sigma \).
• The ideal J is called the ideal of score equations.
How to compute MLE?

• We have to maximize the following objective function:

\[
\max_{\Sigma \in \mathbb{R}^{m \times m}} \ell(\Sigma) = -\log \det \Sigma - \text{tr}(S\Sigma^{-1})
\]

• To find critical points we first compute the following partial derivatives:

\[
-\frac{\partial}{\partial(\cdot)} \det \Sigma - \det \Sigma \frac{\partial}{\partial(\cdot)} \text{tr}(S\Sigma^{-1}).
\]

• Let \(I\) be the ideal generated by the above equations.
• Let \(J\) be the saturation of \(I\) with respect to \(\det \Sigma\).
• The ideal \(J\) is called the ideal of score equations.
• \(V(J)\) is the set of critical points of the objective function.
3. Our Software Package

Goal

To develop a software package for computing MLE of Gaussian models.
3. Our Software Package

**Goal**
To develop a software package for computing MLE of Gaussian models.

**GraphicalModelsMLE [AGHKM21]**
GraphicalModelsMLE computes MLE of Gaussian graphical models.

- We compute MLE, the ideal $J$ of score equations, critical points of the objective function, and ML-degree.
- Our running examples:
  - [1 2 3 4]
  - LMG

Harshit J Motwani (Ghent University)
3. Our Software Package

Goal
To develop a software package for computing MLE of Gaussian models.

GraphicalModelsMLE [AGHKM21]
GraphicalModelsMLE computes MLE of Gaussian graphical models.

- We compute MLE, the ideal $J$ of score equations, critical points of the objective function, and ML-degree.
- Our running examples:
Example 1 (MLE of undirected 4-cycle)

```plaintext
i1 : loadPackage "GraphicalModelsMLE";

i2 : G=graph{{1,2},{2,3},{3,4},{4,1}};

i3 : S=matrix {{.105409, -.0745495, -.0186132, .0621907},
               {-.0745495, .0783734,-.00844503,-.0872842},
               {-.0186132, -.00844503, .128307, .0230245},
               {.0621907, -.0872842, .0230245,.109849}};

i4 : solverMLE(G,S,SampleData=>false)

o4 = (6.62005, | .105409 -.0745495 .0124099 .0621907 |, 5)
     | -.0745495 .0783734 -.00844503 -.0439427 |
     | .0124099 -.00844503 .128307 .0230245 |
     | .0621907 -.0439427 .0230245 .109849 |
```

undirected 4-cycle

1——2
  |   |
  |   |
3——4
Example 1 (Ideal of score equations for 4-cycle)

```plaintext
i5 : U=matrix{{3,5,9,5},{1,6,1,5},{2,9,6,6},{2,5,0,4}};
i6 : J=scoreEquations(gaussianRing G,U);
o6 : Ideal of QQ[k_{1,1}, k_{1,2}, k_{1,4}, k_{2,2}, k_{2,3}, k_{2,4}, k_{3,3}, k_{3,4}, k_{4,4}, k_{4,1}] 
   1,1 2,2 3,3 4,4 1,2 1,4 2,3 3,4
i7 : dim J
o7 = 0
```

- The score ideal is generated by 4 linear polynomials and 10 quadratic polynomials such as

\[ 1312002k_{3,4}^2 - 387081k_{1,2} + 109860k_{1,4} + 1972025k_{2,3} - 898518k_{3,4} - 291556. \]

![Undirected 4-cycle diagram](image-url)
Example 1 (ML-degree of 4-cycle)

i2 : G=graph{[1,2],[2,3],[3,4],[4,1]};
i3 : MLdegree(gaussianRing G)
o3 = 5

ML-degree of 4-cycle

```
i2 : G=graph{[1,2],[2,3],[3,4],[4,1]};
i3 : MLdegree(gaussianRing G)
o3 = 5
```

undirected 4-cycle
Example 2 (MLE of LMG on 4 vertices)

\[
\begin{aligned}
&i2 : G = \text{mixedGraph}(\text{graph}\{\{1,2\}\}, \text{digraph}\{\{1,3\}, \{2,4\}\}, \text{bigraph}\{\{3,4\}\}); \\
&i3 : S = \text{matrix}\{\{34183/50000, 716539/10000000, 204869/250000, 12213/25000\}, \\
&\quad \{716539/10000000, 112191/500000, 309413/1000000, 1803/4000\}, \\
&\quad \{204869/250000, 309413/1000000, 3849/3125, 15172/15625\}, \\
&\quad \{12213/25000, 1803/4000, 15172/15625, 4487/4000\}\}; \\
&i4 : \text{solverMLE}(G, S, \text{SampleData}=>\text{false}) \\
o4 = (9.36624, \{| .68366 .0716539 1.00282 .234375 |}, 5) \\
&\quad | .0716539 .224382 .105105 .733937 | \\
&\quad | 1.00282 .105105 1.76955 -.0700599 | \\
&\quad | .234375 .733937 -.0700599 2.97432 | \\
\end{aligned}
\]
Example 2 (Critical points for LMG on 4 vertices)

i5 : R = gaussianRing G;
i6 : (J,Sigma)=scoreEquations(R,S,SampleData=>false,CovarianceMatrix=>true);
i7 : dim J, degree J
   o7 = (0, 5)
i8 : sols=zeroDimSolve(J);
i9 : checkPD(apply(sols,i->sub(Sigma,matrix{coordinates(i)})))
| .68366  .0716539  1.00282   234375  |, | .68366  .0716539  467724  .070198  |
| .0716539  .224382  .105105   .733937  | | .0716539  .224382  .0490218  .219823  |
| 1.00282  .105105   1.76955  -.0700599 | | .467724  .0490218   75038  .429714  |
| .234375   .733937  -.0700599  2.97432  | | .070198  .219823  .429714  .66928  |
| .68366  .0716539   .675787   .117978  |
| .0716539  .224382  .0708287  .369443  |
| .675787  .0708287   .947611  .211905  |
| .117978  .369443   .211905   .854009  |

LMG

1

2

3

4
Example 3 (Non-identifiable model)

i2: G=mixedGraph(digraph\{\{1,3\},\{1,2\},\{2,4\},\{3,4\}\},
           graph\{\{1,2\}\});
i3: MLdegree(gaussianRing G)
error: the ideal of score equations has dimension 1 > 0,
so ML degree is not well-defined. The degree of this ideal is 2.
Software packages in Macaulay2

- **Our contributions:**
  
  - Created StatGraphs and GraphicalModelsMLE.
  - Updated GraphicalModels.
Software packages in Macaulay2

• **Our contributions:**
  • We created StatGraphs and GraphicalModelsMLE.

Software dependencies:

- Graphs
- StatGraphs
- GraphicalModels
- EigenSolver
- GraphicalModelsMLE
Our contributions:

- We created StatGraphs and GraphicalModelsMLE.
- We also updated GraphicalModels.

Software dependencies:
Software packages in Macaulay2

- **Our contributions:**
  - We created StatGraphs and GraphicalModelsMLE.
  - We also updated GraphicalModels.

- **Software dependencies:**

```plaintext
Graphs
   ↓
StatGraphs
   ↓
GraphicalModels
       ↓
GraphicalModelsMLE
       ↓
EigenSolver
```
The work presented here is based on:


Additional references:


- Sullivant S. Algebraic statistics. American Mathematical Soc; 2018