Gröbner Bases of Gaussian Graphical Models

Alex Fink, Jenna Rajchgot, Seth Sullivant

Queen Mary University, University of Michigan, North Carolina State University

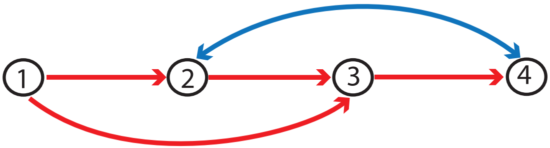
October 4, 2015

Gaussian Graphical Models

- Graphical models are a flexible framework for building statistical models on large collections of random variables.
- Edges of different types represent different types of interactions between neighboring random variables.
  - directed edges: $i \rightarrow j$
  - bidirected edges: $i \leftrightarrow j$
  - undirected edges: $i - j$
- Graph is used to express both
  - conditional independence structures
  - parametric representations of the model.
- For jointly normal random variables, graph structure relates variables to their neighbors via linear relationships with possibly correlated error terms.

Mixed Graphs

- $G = (V, B, D)$ graph with directed edges D ($i \to j$) and bidirected edges B ($i \leftrightarrow j$)
- Vertex set $V = [m] := \{1, 2, \dots m\}$
- G is acyclic: $i \to j \in D$ implies $i < j$.



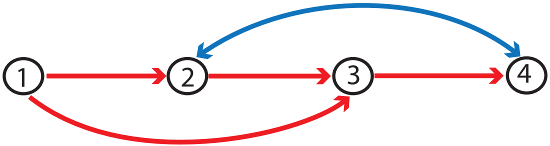
Gaussian Graphical Models

Gaussian graphical model is a statistical model that associates a family of normal distributions to the graph $G = (V, B, D)$.

- $PD_m =$ cone of symmetric positive definite matrices
- Let $PD(B) := \{M \in PD_m : M_{ij} = 0 \text{ if } i \neq j \text{ and } i \leftrightarrow j \notin B\}$
- $\epsilon \in \mathbb{R}^m$, $\epsilon \sim \mathcal{N}(0, \Omega)$ with $\Omega \in PD(B)$
- For $i \to j \in D$, let $\lambda_{ij} \in \mathbb{R}$
- Define $X \in \mathbb{R}^m$ recursively by

$$X_j = \sum_{i : i \to j \in D} \lambda_{ij} X_i + \epsilon_j.$$

Example



$$\epsilon \sim \mathcal{N}(0, \Omega) \qquad \Omega = \begin{pmatrix} \omega_{11} & 0 & 0 & 0 \\ 0 & \omega_{22} & 0 & \omega_{24} \\ 0 & 0 & \omega_{33} & 0 \\ 0 & \omega_{42} & 0 & \omega_{44} \end{pmatrix}$$

$$X_1 = \epsilon_1, \ X_2 = \lambda_{12} X_1 + \epsilon_2, \ X_3 = \lambda_{13} X_1 + \lambda_{23} X_2 + \epsilon_3, \ X_4 = \lambda_{34} X_3 + \epsilon_4$$

Matrix Factorization

Let Λ $m \times m$ upper triangular matrix such that

$$\Lambda_{ij} = \begin{cases} \lambda_{ij} & \text{if } i \to j \in D \\ 0 & \text{otherwise.} \end{cases}$$

Proposition

X from the graph $G = (V, B, D)$ is distributed $\mathcal{N}(0, \Sigma)$ where

$$\Sigma = (I - \Lambda)^{-T} \Omega (I - \Lambda)^{-1}.$$

Note that $(I - \Lambda)^{-1} = I + \Lambda + \Lambda^2 + \cdots + \Lambda^{m-1}$.

The Algebraic Perspective

- Let $\mathbb{R}^D = \{\Lambda \in \mathbb{R}^{m\times m} : \lambda_{ij} = 0 \text{ if } i \to j \notin D\}$
- $PD(B) := \{M \in PD_m : M_{ij} = 0 \text{ if } i \neq j \text{ and } i \leftrightarrow j \notin B\}$

Definition

The Gaussian graphical model $\mathcal{M}_G \subseteq PD_m$ consists of all covariance matrices Σ, that arise for some choice of $\Lambda \in \mathbb{R}^D$ and $\Omega \in PD(B)$.

More algebraically, we have a polynomial map

$$\phi_G : \mathbb{R}^D \times PD(B) \to PD_m, \qquad \phi_G(\Lambda, \Omega) = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}.$$

$$\mathcal{M}_G = \mathrm{im}\phi_G.$$

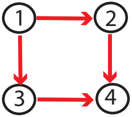
We would like to "understand" ϕ_G and $\mathcal{M}_G$.

Constraints on Gaussian Graphical Models

Problem

Find a generating set or Gröbner basis of

$$I_G = \mathcal{I}(\mathcal{M}_G) = \{f \in \mathbb{R}[\sigma_{ij} : i, j \in [m]] : f(\Sigma) = 0 \text{ for all } \Sigma \in \mathcal{M}_G\}.$$



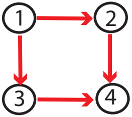
$$I_G = \langle |\Sigma_{12,13}|, |\Sigma_{123,234}| \rangle$$

Constraints on Gaussian Graphical Models

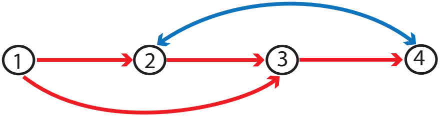
Problem

Find a generating set or Gröbner basis of

$$I_G = \mathcal{I}(\mathcal{M}_G) = \{f \in \mathbb{R}[\sigma_{ij} : i, j \in [m]] : f(\Sigma) = 0 \text{ for all } \Sigma \in \mathcal{M}_G\}.$$



$$I_G = \langle |\Sigma_{12,13}|, |\Sigma_{123,234}| \rangle$$



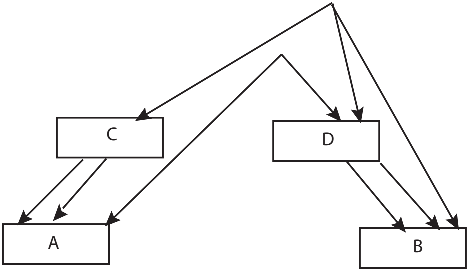
$$\begin{aligned} I_G \quad = \quad \langle\, & \sigma_{12}\sigma_{13}\sigma_{14}\sigma_{23} - \sigma_{11}\sigma_{14}\sigma_{23}^2 - \sigma_{12}\sigma_{13}^2\sigma_{24} + \sigma_{11}\sigma_{13}\sigma_{23}\sigma_{24} \\ & -\sigma_{12}^2\sigma_{14}\sigma_{33} + \sigma_{11}\sigma_{14}\sigma_{22}\sigma_{33} + \sigma_{12}^2\sigma_{13}\sigma_{34} - \sigma_{11}\sigma_{13}\sigma_{22}\sigma_{34} \,\rangle \end{aligned}$$

Trek Separation

- A trek from i to j is a path in G from i to j with no sequence of edges $k \rightarrow l \leftarrow m$, $k \leftrightarrow l \leftarrow m$, $k \rightarrow l \leftrightarrow m$, or $k \leftrightarrow l \leftrightarrow m$.

Definition

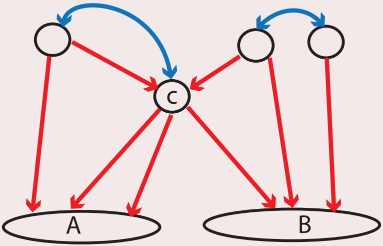
Let A, B, C, and D be four subsets of $V(G)$ (not necessarily disjoint). We say that (C, D) t-separates A from B if every trek from A to B passes through either a vertex in C on the A-side of the trek, or a vertex in D on the B-side of the trek.



Theorem (S-Talaska-Draisma)

The matrix $\Sigma_{A,B}$ has rank $\leq d$ if and only if there are $C, D \subset [n]$ with $\#C + \#D \leq d$ such that (C, D) t-separate A from B in G. Then all $(d+1) \times (d+1)$ minors of $\Sigma_{A,B}$ belong to I_G.

Example



$(\{c\}, \{c\})$ *t*-separates A from B.
$\Sigma_{A,B}$ has rank at most 2. All 3×3 minors of $\Sigma_{A,B}$ belong to I_G.

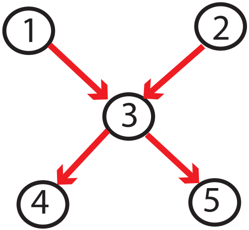
When do Determinantal Constraints Generate I_G?

Question

What conditions on the graph G guarantee that the t-separation determinantal constraints generate I_G?

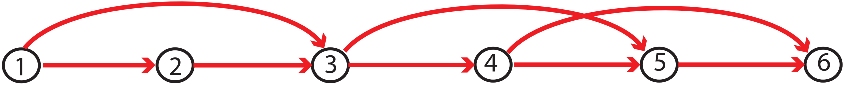
Theorem (Sullivant 2008)

If G is a tree, then I_G is generated in degree 1 and 2 by conditional independence constraints. In particular, I_G generated by the t-separation determinantal constraints.



Proof Sketch.

For trees I_G is a toric ideal. Do binomial manipulations. □



Definition

A generalized Markov chain is a mixed graph $G = (V, B, D)$ such that:

- If $i \rightarrow j \in D$ then $i < j$,
- If $i \rightarrow j \in D$ and $i \leq k < l \leq j$ then $k \rightarrow l \in D$, and
- If $i \leftrightarrow j \in B$ and $i \leq k < l \leq j$ then $k \leftrightarrow l \in B$.

Theorem (Fink-Rajchgot-S (2015))

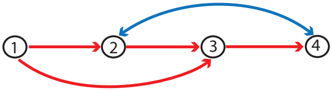
If G is a generalized Markov chain then I_G is generated by the t-separation determinantal constraints implied by G, and they form a Gröbner basis in a suitable lexicographic term order.

Proof Sketch.

Relate generalized Markov chain parametrization to parametrization of type B matrix Schubert varieties. □

Further Results and Open Problems

- We have extended t-separation characterization of determinantal constraints to ancestral graphs and AMP chain graphs.
- Connections to toric varieties and matrix Schubert varieties allow us to characterize the vanishing ideals for some graphs.
- How to determine general combinatorial descriptions of other hidden variable constraints?



$$\begin{vmatrix} \sigma_{11} & \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{12} & \sigma_{22} & \sigma_{23} \\ 0 & \sigma_{13} & \sigma_{23} & \sigma_{33} \\ 0 & \sigma_{14} & \sigma_{24} & \sigma_{34} \end{vmatrix} = 0$$

References

A. Fink, J. Rajchgot, S. Sullivant. Symmetric matrix Schubert varieties and Gaussian graphical models. (2015) In preparation.

Knutson, Allen. Frobenius splitting, point-counting, and degeneration. (2009) :0911.4941

Knutson, Allen; Miller, Ezra. Gröbner geometry of Schubert polynomials. *Ann. of Math.* (2) 161 (2005), no. 3, 1245–1318.

S. Sullivant. Algebraic geometry of Gaussian Bayesian networks. *Adv. in Appl. Math.* **40** (2008), no. 4, 482–513. 0704.0918

S. Sullivant, K. Talaska and J. Draisma. Trek separation for Gaussian graphical models. *Annals of Statistics* **38** no.3 (2010) 1665-1685 0812.1938