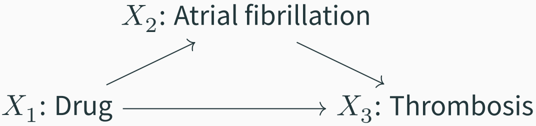
Higher Moment Varieties of Non-Gaussian Graphical Models

Carlos Améndola Mathias Drton **Alexandros Grosdos**
Roser Homs Elina Robeva

MEGA 21

Graphical models

Graphical models capture causal relations between random variables



translating to equations

$$
\begin{aligned}
X_1 &= & & & & & & \varepsilon_1 \\
X_2 &= &\lambda_{12}X_1 & & & &+ &\varepsilon_2 \\
X_3 &= &\lambda_{13}X_1 &+ &\lambda_{23}X_2 & &+ &\varepsilon_3
\end{aligned}
$$

A graph G gives rise to structural equations

$$X_i = \sum_{j \in \mathrm{pa}(i)} \lambda_{ji} X_j + \varepsilon_i, \quad i \in V,$$

where

- ε_i represent stochastic errors with $\mathbb{E}[\varepsilon_i] = 0$,
- λ_{ji} are unknown parameters forming a matrix $\Lambda = (\lambda_{ji})$.

The corresponding model is

$$\begin{aligned}\mathcal{M}^{(2,3)}(G) = \{(S = (I-\Lambda)^{-T}\Omega^{(2)}(I-\Lambda)^{-1},\\ T = \Omega^{(3)} \bullet (I-\Lambda)^{-1} \bullet (I-\Lambda)^{-1} \bullet (I-\Lambda)^{-1}) :\\ \Omega^{(2)} \text{ is } n \times n \text{ positive definite diagonal matrix,}\\ \Omega^{(3)} \text{ is } n \times n \times n \text{ diagonal 3-way tensor, and } \Lambda \in \mathbb{R}^E\}.\end{aligned}$$

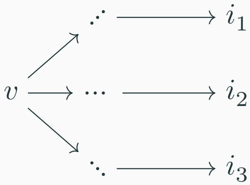
This makes (statistical) sense for Non-Gaussian random variables.

Treks on graphs

A *trek* τ with top v between i and j is formed by two paths sharing a source v

$$i \leftarrow i_l \leftarrow \cdots \leftarrow i_1 \leftarrow v \rightarrow j_1 \rightarrow \ldots \rightarrow j_r \rightarrow j.$$

An n-*trek* between n vertices $i_1, \ldots, i_n$ is an ordered collection of n directed paths $T = (P_1, \ldots, P_n)$, where P_r has sink i_r and they all share the same top vertex as source $v = \mathrm{top}(T)$.



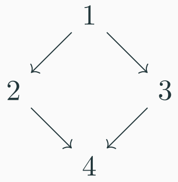
For a graph G, let $T(i_1, \dots, i_n)$ denote all minimal n-treks between $i_1, \dots, i_n$.

Consider the ring morphism ϕ_G:

$$\begin{aligned} \mathbb{C}[s_{ij}, t_{ijk} \mid 1 \le i \le j \le k \le n] &\to \mathbb{C}[a_i, b_i, \lambda_{ij} \mid i \to j \in E] \\ s_{ij} &\mapsto \sum_{T \in T(i,j)} a_{\text{top}(T)} \prod_{k \to l \in T} \lambda_{kl}, \\ t_{ijk} &\mapsto \sum_{T \in T(i,j,k)} b_{\text{top}(T)} \prod_{m \to l \in T} \lambda_{ml}. \end{aligned}$$

Example

$$\begin{aligned} s_{ii} &\mapsto a_i \\ t_{iii} &\mapsto b_i \\ s_{13} &\mapsto a_1 \lambda_{13} \\ s_{14} &\mapsto a_1 \lambda_{12} \lambda_{24} + a_1 \lambda_{13} \lambda_{34} \\ t_{123} &\mapsto b_1 \lambda_{12} \lambda_{13} \end{aligned}$$



Proposition [Sullivant 08; Améndola, Drton, G, Homs & Robeva 21+] Let G be a DAG (directed acyclic graph) and ϕ_G given by the simple trek rule. Then the vanishing ideal $I^{(2,3)}(G) := \mathcal{J}(\mathcal{M}^{(2,3)}(G))$ of the model is

$$I^{(2,3)}(G) = \ker \phi_G.$$

Corollary [Améndola, Drton, G, Homs & Robeva 21+] If G is a tree, $I^{(2,3)}(G)$ is a toric ideal.

Let $i, j \in V$ be two vertices such that a 2-trek between i and j exists. Define

$$A_{ij} := \begin{bmatrix} s_{ik_1} & \cdots & s_{ik_r} & t_{i\ell_1 m_1} & \cdots & t_{i\ell_q m_q} \\ s_{jk_1} & \cdots & s_{jk_r} & t_{j\ell_1 m_1} & \cdots & t_{j\ell_q m_q} \end{bmatrix},$$

where

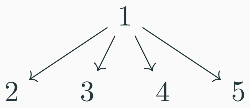
- $k_1, \ldots, k_r$ are all vertices such that $\mathrm{top}(i, k_a) = \mathrm{top}(j, k_a)$ and
- (l_1, m_1),... ,(l_q, m_q) are all pairs of vertices such that $\mathrm{top}(i, l_b, m_b) = \mathrm{top}(j, l_b, m_b)$.

Proposition [Améndola, Drton, G, Homs & Robeva 21+] For a tree G, the following polynomials are in $I^{(2,3)}(G)$:

- s_{ij} such that there is no 2-trek between i and j,
- t_{ijk} such that there is no 3-trek between i, j and k,
- the 2-minors of A_{ij}, for all (i, j) with a 2-trek between them.

Proposition [Améndola, Drton, G, Homs & Robeva 21+] All quadratic binomials in $I^{(2,3)}(G)$ are linear combinations of 2-minors of matrices A_{ij}.

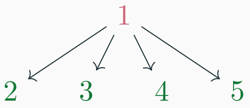
Example The binomial $f = s_{23}t_{145} - s_{45}t_{123}$ lies in $I^{(2,3)}(G)$. It is the sum of the minors from A_{13}, A_{14} and A_{15}.



Theorem [Améndola, Drton, G, Homs & Robeva 21+] All binomials in $I^{(2,3)}(G)$ are generated by quadratic binomials, i.e. $I^{(2,3)}(G)$ is generated by the matrices A_{ij} (plus vanishing indeterminates).

Proof A distance reduction argument for binomials in the ideal, showing that matrix minors are a Markov basis.

Let $H \cup O$ be a partition of the nodes of the DAG G. The hidden nodes H are said to be *upstream* from the observed nodes O in G if there are no edges $o \to h$ in G with $o \in O$ and $h \in H$.



Lemma The ideal $I^{(2,3)}(G)$ is homogeneous w.r.t. the grading:

$$\begin{aligned} \deg s_{ij} &= (1, 1 + \text{number of elements in the multiset } \{i,j\} \text{ in } O) \\ \deg t_{ijk} &= (1, \text{number of elements in the multiset } \{i,j,k\} \text{ in } O). \end{aligned}$$

Proposition For a tree G, $I_O^{(2,3)}(G)$ is generated by the minors of the submatrices of A_{ij} with i, j both in O, with columns indexed by k and (l, m) where k, l, m are all in O.

Variety admits more compact description

Theorem [Améndola, Drton, G, Homs & Robeva 21+] Let J be the ideal generated by the linear generators of $I^{(2,3)}(G)$ and matrices A_{ij} such that there is a directed path between i and j. Then

$$\mathcal{M}^{(2,3)}(G) = V(J) \cap PD(n).$$

In particular, pick $(S,T) \in \mathcal{M}^{(2,3)}(G)$. For $i \to j \in E$, let $\lambda_{ij} = \frac{s_{ij}}{s_{ii}}$, coming from A_{ij}. Then one can show

$$S' = (I-\Lambda)^T S(I-\Lambda) \quad \text{and} \quad T' = T \bullet (I-\Lambda) \bullet (I-\Lambda) \bullet (I-\Lambda)$$

are diagonal.

Example Let G be $1 \to 2, 1 \to 3, 1 \to 4, 1 \to 5$. Computations show

$$I^{(2,3)}(G) = (J : s_{11}^{\infty})$$

and

$$\mathcal{M}^{(2,3)}(G) = V(I^{(2,3)}(G)) \cap PD(5) = V(J) \cap PD(5).$$

Summary

- Graphical models are richer in the Non-Gaussian setting, it is meaningful to study covariance matrices plus higher-order moment tensors.
- The trek rules can be extended for h.o.m. and one can obtain binomial (matrix minors) descriptions of the corresponding ideals.
- The hidden variable ideals and the varieties only need a subset of the polynomials.

For more information have a look at the extended abstract and stay tuned for the preprint.

Thank you!