Tobias Boege
joint work with Garett Cunningham

Computational methods in Gaussian conditional independence inference

MathCoRe seminar
4 Nov 2020

Institut für Algebra und Geometrie
Otto-von-Guericke-Universität Magdeburg

DFG-Graduiertenkolleg
MATHEMATISCHE
KOMPLEXITÄTSREDUKTION

Gaussian conditional independence

- Finite ground set $[n] = \{1, \dots, n\}$ indexing some objects ξ_i.
- CI statement $(ij|K)$ abbreviating $\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K$.
- All CI statements $\mathcal{A}_n = \{(ij|K) : ij \in \binom{[n]}{2}, K \subseteq [n] \setminus ij\}$.

Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ be a vector of n random variables with a joint regular Gaussian distribution ($\Sigma \in \mathrm{PD}_n$), then

$$\begin{array}{ccc} \xi_i \perp\!\!\!\perp \xi_j \mid \xi_K & \Leftrightarrow & \det \Sigma_{iK,jK} = 0. \\ \Updownarrow & & \vdots\vdots \\ (ij|K) & & \det \Sigma_{ij|K} \end{array}$$

Realizability and the Implication problem

$$[\![\Sigma]\!] = \{(ij|K) \in \mathcal{A}_n : \det \Sigma_{ij|K} = 0\}.$$

The sets $[\![\Sigma]\!]$ for $\Sigma \in \mathrm{PD}_n$ have a special combinatorial structure.
For example:

$$(ij|L)_{\in[\![\Sigma]\!]} \wedge (ij|kL)_{\in[\![\Sigma]\!]} \Rightarrow (ik|L)_{\in[\![\Sigma]\!]} \vee (jk|L)_{\in[\![\Sigma]\!]} \qquad (*)$$

$$(ij|L) \wedge (ij|kL) \Rightarrow (ik|L) \vee (jk|L) \qquad (*)$$

for all distinct $i, j, k \in [n]$ and $L \subseteq [n] \setminus ijk$.

Goal. Describe all the sets $[\![\Sigma]\!]$, the CI structures *realizable* by regular Gaussian distributions.

Observation. Knowing all realizable structures is equivalent to knowing all valid inference rules of the form $(*)$.

All valid inference rules for Gaussians can be "symmetrized":

$$(12|) \wedge (12|3) \Rightarrow (13|) \vee (23|) \text{ for } n = 3 \curvearrowright$$
$$(ij|L) \wedge (ij|kL) \Rightarrow (ik|L) \vee (jk|L) \;\; \forall i, j, k, L \text{ for every } n \geq 3.$$

Theorem (Šimeček / Sullivant). No finite set of inference rules exactly describes regular Gaussian CI structures.

Theorem. Every finite set of inference rules which allows all regular Gaussian CI structures allows asymptotically $2^{2^{\Omega(n)}}$ structures. The number of Gaussian CI structures is bounded by $2^{\mathcal{O}(n^3)}$.

Theme: properties of realizable structures give valid inference rules. Derive valid rules from approximations of realizable structures.

1. Combinatorial compatibility & SAT solvers
2. Information inequalities & LP solvers
3. The selfadhesivity phenomenon
4. Real algebra & Semidefinite programming

Combinatorial compatibility

The *Gaussian CI configuration space* is the image of

$$\Sigma \mapsto (p_I(\Sigma) : I \subseteq [n]) \cup \left(a_{ij|K}(\Sigma) : (ij|K) \in \mathcal{A}_n\right).$$

Study the zero patterns of configuration vectors.
There exist polynomial relations on them, for example

$$p_{kL} a_{ij|L} - p_L a_{ij|kL} - a_{ik|L} a_{jk|L} = 0. \qquad \text{(ET)}$$

Combinatorial compatibility means "vanishing under uncertainty":
What if we only knew that $p_I \neq 0$ and whether $a_{ij|K} = 0$ or $a_{ij|K} \neq 0$?

$$(ij|L) \wedge (ij|kL) \Rightarrow (ik|L) \vee (jk|L)$$
$$(ij|L) \wedge (ik|L) \Rightarrow (ij|kL)$$
$$\vdots$$

Gaussoids

$$\begin{aligned}
(ij|L) \wedge (ik|jL) &\Rightarrow (ik|L) \wedge (ij|kL) \\
(ij|kL) \wedge (ik|jL) &\Rightarrow (ij|L) \wedge (ik|L) \\
(ij|L) \wedge (ik|L) &\Rightarrow (ij|kL) \wedge (ik|jL) \\
(ij|L) \wedge (ij|kL) &\Rightarrow (ik|L) \vee (jk|L)
\end{aligned}$$

For $n = 4$:

- All subsets of $\mathcal{A}_4$: 16,777,216
- Gaussoids: 679
- Realizable gaussoids: 629

For $n = 5$:

- All subsets of $\mathcal{A}_5$: 1,208,925,819,614,629,174,706,176
- Gaussoids: 60,212,776
- Realizable gaussoids: ???

Tools: AllSAT solver to enumerate satisfying assignments.

Breakout break

Oriented gaussoids

$$p_{kL} a_{ij|L} - p_L a_{ij|kL} - a_{ik|L} a_{jk|L} = 0 \tag{ET}$$

What if we only knew that $\operatorname{sgn} p_I = +1$ and the value of $\operatorname{sgn} a_{ij|K}$?

$$\begin{aligned} +(ij|L) \wedge -(ij|kL) \; &\Rightarrow [+(ik|L) \wedge +(jk|L)] \\ &\vee [-(ik|L) \wedge -(jk|L)] \end{aligned}$$

For $n = 4$:

- Gaussoids: 679
- Orientable gaussoids: **629**

For $n = 5$:

- Gaussoids: 60,212,776
- Orientable gaussoids: 20,584,290

Tools: AllSAT and incremental SAT solver to compute projected satisfying assignments.

The multiinformation region

$$a_{ij|L}^2 = p_{iL} p_{jL} - p_L p_{ijL} \tag{FT}$$

The *Gaussian multiinformation region* is the image of

$$\Sigma \mapsto (\log p_I(\Sigma) : I \subseteq [n]) \, .$$

By (FT), the Gaussian multiinformation functions $m(I) = \log \det \Sigma_I$ satisfy the following *linear information inequalities*

$$\Delta_{ij|K} m := m(iK) + m(jK) - m(ijK) - m(K) \geq 0, \quad \text{(Submodularity)}$$

and submodularity at $\Delta_{ij|K}$ is tight if and only if $(ij|K) \in [\![\Sigma]\!]$.

Information inequalities

Linear information inequalities at the multiinformation region of the form

$$\sum_{\beta \in \mathcal{M}} c_\beta \Delta_\beta m \le \sum_{\alpha \in \mathcal{L}} c_\alpha \Delta_\alpha m, \quad c_\alpha, c_\beta > 0, \ \forall m$$

with $\mathcal{L}, \mathcal{M} \subseteq \mathcal{A}_n$ encode CI inference rules

$$\bigwedge_{\alpha \in \mathcal{L}} \alpha \Rightarrow \bigwedge_{\beta \in \mathcal{M}} \beta.$$

CI implication: does tightness of some inequalities imply tightness of others? Study the *face lattice* of polyhedral cones spanned by valid linear information inequalities.

Semimatroids

The (balanced) Shannon information inequalities define a rational polyhedral cone in $\mathbb{R}^{2^n}$ whose elements are valid inequalities at the Gaussian multiinformation region. The CI inference rules encoded in its face lattice define *semimatroids*.

For $n = 4$:

- Gaussoids: 679
- Semimatroidal gaussoids: **629**

For $n = 5$:

- Gaussoids: 60,212,776
- Semimatroidal gaussoids: 39,807,192
- Semimatroidal orientable gaussoids: 20,576,142, 8,148 less than just orientable gaussoids.

Tools: Exact rational LP solver to find an inner point on a face of the cone or, alternatively, a Farkas certificate.

Selfadhesivity

Theorem. Positive definite matrices are *selfadhesive* in the sense that for every $\Sigma \in \mathrm{PD}_N$ and every $I \subseteq N$ there exist a copy M of N with $N \cap M = I$ and $\Phi \in \mathrm{PD}_{NM}$ such that:

- $\Phi_N = \Sigma$ and $\Phi_M = \Sigma$,
- $\Phi_{N,M}$ has rank exactly $|N \cap M|$.

Such a Φ is an *adhesive extension of* Σ *at* I.

Realizable CI structures inherit *structural* selfadhesivity properties, which are necessary for realizability:

- can be glued as gaussoids (apparently trivial?),
- can be glued as orientable gaussoids (19,723,980/20,584,290),
- can be glued as semimatroids (39,595,332/39,807,192),
- ...

Outlook: Formal realizations

Transfer principle. Let $\mathbb{K}$ be an ordered field extension of $\mathbb{R}$. Then for every matrix Σ over $\mathbb{K}$ which satisfies

$$\det \Sigma_I > 0 \ \text{ in } \mathbb{K} \text{ for all } I \subseteq [n],$$

the formal CI structure $[\![\Sigma]\!]$ is realizable over $\mathbb{R}$.

$$\begin{pmatrix} 1 & g & \varepsilon f & \varepsilon \\ g & 1 & 0 & \varepsilon g \\ \varepsilon f & 0 & 1 & f \\ \varepsilon & \varepsilon g & f & 1 \end{pmatrix}, \quad f = \frac{2\varepsilon}{1+\varepsilon^2}, \ g = \frac{1-\varepsilon^2}{1+\varepsilon^2}$$

over $\mathbb{R}(\varepsilon)$, $\varepsilon > 0$ infinitesimal, proves realizability of

$$\{(13|4), (14|23), (23|), (24|1)\}.$$

Outlook: Optimization

Real algebraic geometry has an exact algebraic solution for the implication / realizability problem (*the real deal, not a relaxation!*):

Theorem. The inference rule $\bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ is valid for regular Gaussians if and only if $\prod_{(ij|K)\in\mathcal{M}} a_{ij|K}$ vanishes on the semialgebraic set of matrices Σ given by

$$\{\det \Sigma_{ij|K} = 0 \;\; \forall (ij|K) \in \mathcal{L}\} \cap \{\det \Sigma_I > 0 \;\; \forall I \subseteq [n]\}.$$

This happens if and only if

$$\prod_{(ij|K)\in\mathcal{M}} a_{ij|K} \;\in\; \sqrt[\mathbb{R}]{\mathcal{I}(a_{ij|K} : (ij|K) \in \mathcal{L}) + \mathcal{I}(1 - y_I^2 p_I : I \subseteq [n])}$$

in the polynomial ring $\mathbb{R}[\Sigma, y_I : I \subseteq [n]]$.

Tools: Polynomial optimization and SDP relaxations.

Tobias Boege, Alessio D'Alì, Thomas Kahle, and Bernd Sturmfels.
The Geometry of Gaussoids.
Found. Comput. Math., 19(4):775–812, 2019.

Terence H. Chan.
Balanced information inequalities.
IEEE Trans. Inform. Theory, 49(12):3261–3267, 2003.

František Matúš.
Conditional independence structures examined via minors.
Ann. Math. Artif. Intell., 21(1):99–30, 1997.

František Matúš.
Adhesivity of polymatroids.
Discrete Math., 307(21):2464–2477, 2007.