The complexity of Gaussian conditional independence models

Tobias Boege

Max-Planck Institute for Mathematics in the Sciences, Leipzig

Oberseminar Statistics and Data Science
München, 19 July 2022

Conditional independence $X \perp\!\!\!\perp Y \mid Z$

"When does knowing Z make X irrelevant for Y?"

Conditional independence $X \perp\!\!\!\perp Y \mid Z$

"When does knowing Z make X irrelevant for Y?"

Example: Two independent fair coins c_1 and c_2 are wired to a bell b which rings if and only if $c_1 = c_2$.

Conditional independence $X \perp\!\!\!\perp Y \mid Z$

"When does knowing Z make X irrelevant for Y?"

Example: Two independent fair coins c_1 and c_2 are wired to a bell b which rings if and only if $c_1 = c_2$.

- $c_1 \perp\!\!\!\perp c_2$

Conditional independence $X \perp\!\!\!\perp Y \mid Z$

"When does knowing Z make X irrelevant for Y?"

Example: Two independent fair coins c_1 and c_2 are wired to a bell b which rings if and only if $c_1 = c_2$.

- $c_1 \perp\!\!\!\perp c_2$
- $\neg(c_1 \perp\!\!\!\perp c_2 \mid b)$ ...

Conditional independence $X \perp\!\!\!\perp Y \mid Z$

"When does knowing Z make X irrelevant for Y?"

Example: Two independent fair coins c_1 and c_2 are wired to a bell b which rings if and only if $c_1 = c_2$.

- $c_1 \perp\!\!\!\perp c_2$
- $\neg(c_1 \perp\!\!\!\perp c_2 \mid b)$ ...

Question: When can we conclude from some independences other independences? E.g., is it possible that $c_1 \perp\!\!\!\perp b$?

Graphical models

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2017WR020412

Key Points:

- We develop a statistical graphical model to characterize the statewide California reservoir system
- We quantify the influence of external physical and economic factors (e.g., statewide PDSI and consumer price index) on the reservoir network
- Further analysis gives a system-wide health diagnosis as a function of PDSI, indicating when heavy management practices may be needed

Supporting Information:

- Supporting Information S1
- Supporting Information S2

A Statistical Graphical Model of the California Reservoir System

A. Taeb[1], J. T. Reager[2], M. Turmon[2], and V. Chandrasekaran[3]

[1]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [3]Department of Computing and Mathematical Sciences and Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA

Abstract The recent California drought has highlighted the potential vulnerability of the state's water management infrastructure to multiyear dry intervals. Due to the high complexity of the network, dynamic storage changes in California reservoirs on a state-wide scale have previously been difficult to model using either traditional statistical or physical approaches. Indeed, although there is a significant line of research on exploring models for single (or a small number of) reservoirs, these approaches are not amenable to a system-wide modeling of the California reservoir network due to the spatial and hydrological heterogeneities of the system. In this work, we develop a state-wide statistical graphical model to characterize the dependencies among a collection of 55 major California reservoirs across the state; this model is defined with respect to a graph in which the nodes index reservoirs and the edges specify the

Gaussian conditional independence

Assume $\xi = (\xi_i : i \in N)$ are jointly Gaussian with covariance matrix $\Sigma \in \mathrm{PD}_N$.

Definition

The polynomial $\Sigma[K] \coloneqq \det \Sigma_{K,K}$ is a *principal minor* of Σ and $\Sigma[ij \mid K] \coloneqq \det \Sigma_{iK,jK}$ is an *almost-principal minor*.

- Σ is PD if and only if $\Sigma[K] > 0$ for all $K \subseteq N$.
- $[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ holds if and only if $\Sigma[ij \mid K] = 0$.
- $\mathbb{E}[\xi] = \mu$ is irrelevant.

Very special polynomials

$$\begin{aligned}
\Sigma[\,ij\,|\,\,] &= x_{ij}\\
\Sigma[\,ij\,|\,k\,] &= x_{ij}x_{kk} - x_{ik}x_{jk}\\
\Sigma[\,ij\,|\,kl\,] &= x_{ij}x_{kk}x_{ll} - x_{il}x_{jl}x_{kk} + x_{il}x_{jk}x_{kl} + x_{ik}x_{jl}x_{kl} - x_{ij}x_{kl}^2 - x_{ik}x_{jk}x_{ll}\\
\Sigma[\,ij\,|\,klm\,] &= x_{ij}x_{kk}x_{ll}x_{mm} + x_{im}x_{jm}x_{kl}^2 - x_{im}x_{jl}x_{kl}x_{km} - x_{il}x_{jm}x_{kl}x_{km} +\\
&\quad x_{il}x_{jl}x_{km}^2 - x_{im}x_{jm}x_{kk}x_{ll} + x_{im}x_{jk}x_{km}x_{ll} + x_{ik}x_{jm}x_{km}x_{ll} -\\
&\quad x_{ij}x_{km}^2x_{ll} + x_{im}x_{jl}x_{kk}x_{lm} + x_{il}x_{jm}x_{kk}x_{lm} - x_{im}x_{jk}x_{kl}x_{lm} -\\
&\quad x_{ik}x_{jm}x_{kl}x_{lm} - x_{il}x_{jk}x_{km}x_{lm} - x_{ik}x_{jl}x_{km}x_{lm} + 2x_{ij}x_{kl}x_{km}x_{lm} +\\
&\quad x_{ik}x_{jk}x_{lm}^2 - x_{ij}x_{kk}x_{lm}^2 - x_{il}x_{jl}x_{kk}x_{mm} + x_{il}x_{jk}x_{kl}x_{mm} +\\
&\quad x_{ik}x_{jl}x_{kl}x_{mm} - x_{ij}x_{kl}^2x_{mm} - x_{ik}x_{jk}x_{ll}x_{mm}\\
&\;\vdots
\end{aligned}$$

Gaussian CI models

Definition

A *CI constraint* is a CI statement $[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ or its negation $\neg[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$.
The *model* of a set of CI constraints is the set of all PD matrices which satisfy them.

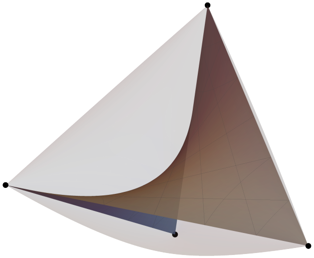


Figure: Model of $\Sigma[12 \mid 3] = a - bc = 0$ in the space of 3×3 correlation matrices.

Basic questions

- How hard is it to decide if the model specification is inconsistent?
- How hard is it to *certify* consistency by showing a point in the model?
- What is the geometric structure of the models?

Basic questions

- How hard is it to decide if the model specification is inconsistent?
- How hard is it to *certify* consistency by showing a point in the model?
- What is the geometric structure of the models?

What is the model of $[X \perp\!\!\!\perp Y] \wedge [X \perp\!\!\!\perp Z \mid Y] \wedge \neg[X \perp\!\!\!\perp Y \mid Z]$?

Models and inference

Consider two sets of CI statements $\mathcal{P}$ and $\mathcal{Q}$:

$$\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q}$$

Models and inference

Consider two sets of CI statements $\mathcal{P}$ and $\mathcal{Q}$:

$$\underbrace{\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q}}_{\text{is not valid}} \quad \Longleftrightarrow \quad \underbrace{\mathcal{P} \cup \neg\mathcal{Q}}_{\text{has a point}}$$

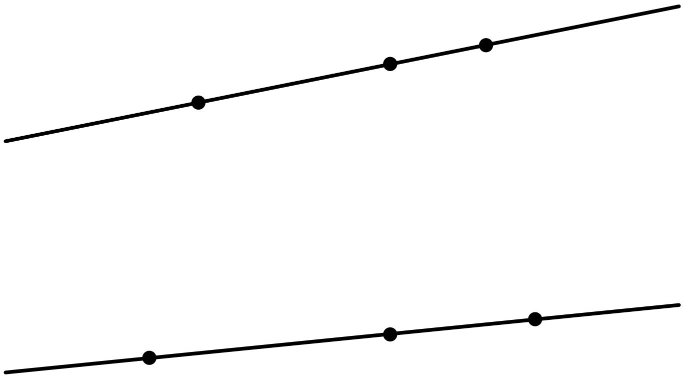
Models and inference

Consider two sets of CI statements $\mathcal{P}$ and $\mathcal{Q}$:

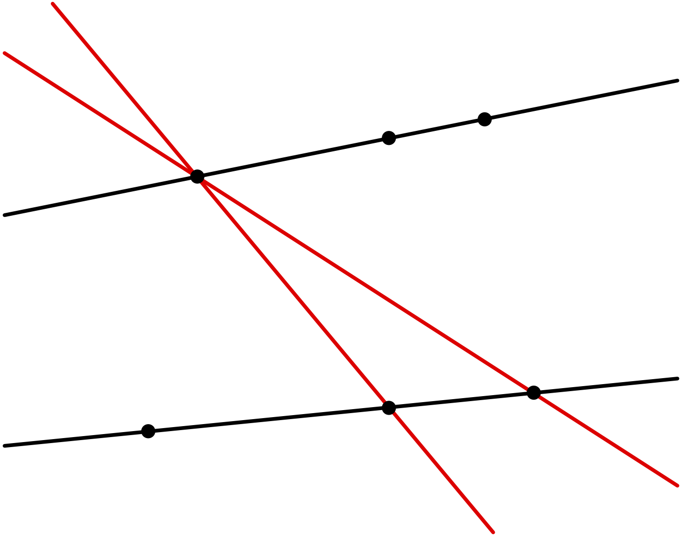
$$\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q} \text{ is not valid} \quad \iff \quad \mathcal{P} \cup \neg\mathcal{Q} \text{ has a point}$$

Reasoning about CI statements in normally distributed random variables is the same as reasoning about the vanishing of very special kinds of determinants on very special kinds of varieties inside the positive definite matrices.

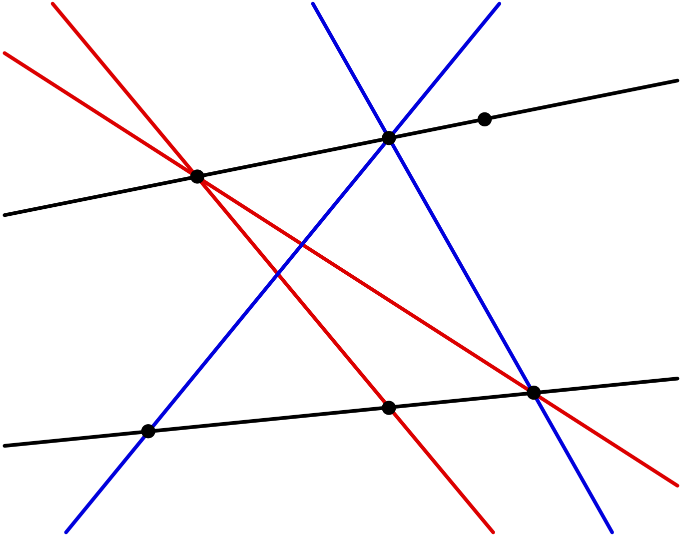
For ancient geometers: conditional independence ≈ collinearity



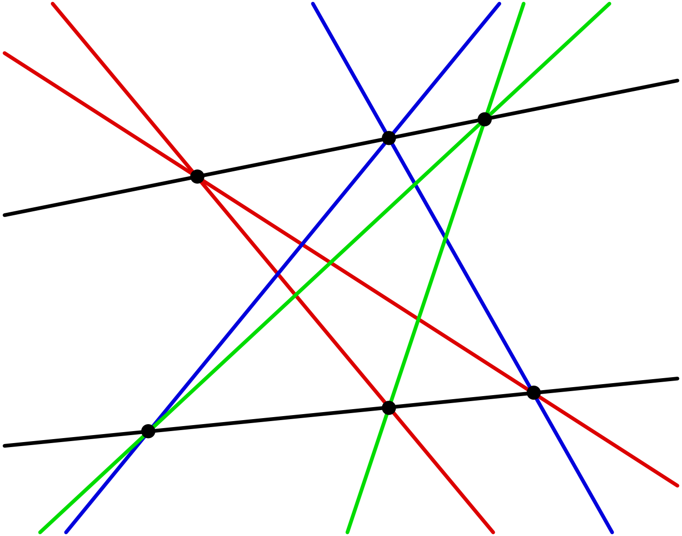
For ancient geometers: conditional independence ≈ collinearity



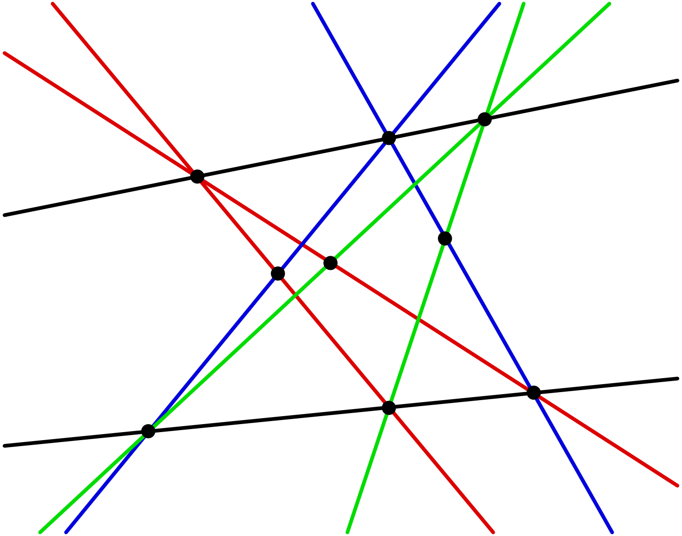
For ancient geometers: conditional independence ≈ collinearity



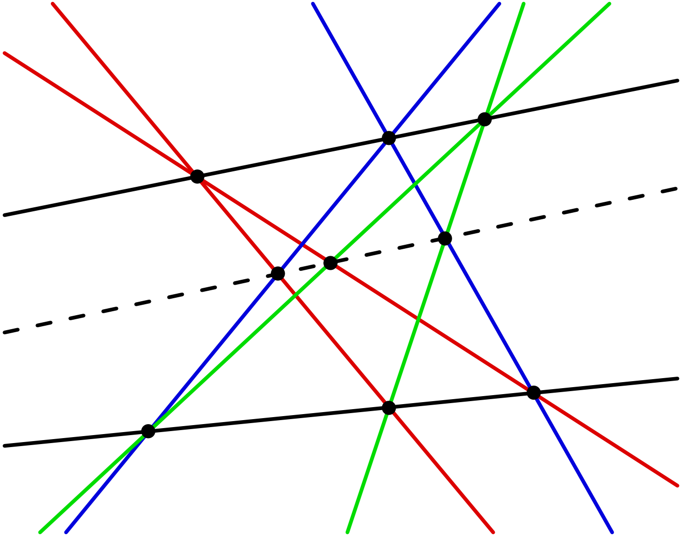
For ancient geometers: conditional independence ≈ collinearity



For ancient geometers: conditional independence ≈ collinearity



For ancient geometers: conditional independence ≈ collinearity



Normal form for proofs and refutations

Let $f_i \in \mathbb{Z}[t_1, \ldots, t_k]$ be integer polynomials in finitely many variables.

Theorem (Tarski's transfer principle)

If a polynomial system $\{f_i \bowtie_i 0\}$, where $\bowtie_i \in \{=, \neq, <, \leq, \geq, >\}$, has a solution over $\mathbb{R}$, then it has a solution in a finite real extension of $\mathbb{Q}$.

Normal form for proofs and refutations

Let $f_i \in \mathbb{Z}[t_1, \ldots, t_k]$ be integer polynomials in finitely many variables.

Theorem (Tarski's transfer principle)

If a polynomial system $\{f_i \bowtie_i 0\}$, where $\bowtie_i \in \{=, \neq, <, \leq, \geq, >\}$, has a solution over $\mathbb{R}$, then it has a solution in a finite real extension of $\mathbb{Q}$.

→ If $\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q}$ is false, there exists a counterexample matrix Σ with algebraic entries.

$[12||] \wedge [12|3] \Rightarrow [13||]$ is false and a counterexample is

$$\begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1 \end{pmatrix}.$$

Normal form for proofs and refutations

Let $f_i, g_j, h_k \in \mathbb{Z}[t_1, \ldots, t_k]$ be integer polynomials in finitely many variables.

Theorem (Positivstellensatz)

A polynomial system $\{f_i = 0, g_j \geq 0, h_k \neq 0\}$ is infeasible if and only if there exist $f \in \mathrm{ideal}(f_i)$, $g \in \mathrm{cone}(g_j)$ and $h \in \mathrm{monoid}(h_k)$ such that $g + h^2 = f$.

Normal form for proofs and refutations

Let $f_i, g_j, h_k \in \mathbb{Z}[t_1, \ldots, t_k]$ be integer polynomials in finitely many variables.

Theorem (Positivstellensatz)

A polynomial system $\{f_i = 0, g_j \geq 0, h_k \neq 0\}$ *is infeasible if and only if there exist* $f \in \mathrm{ideal}(f_i)$, $g \in \mathrm{cone}(g_j)$ *and* $h \in \mathrm{monoid}(h_k)$ *such that* $g + h^2 = f$.

→ If $\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q}$ is true, there exists an algebraic proof for it with integer coefficients.

$[12||] \wedge [12|3] \Rightarrow [13||] \vee [23||]$ is true and a proof is the final polynomial

$$\Sigma[13||] \cdot \Sigma[23||] = \Sigma[3] \cdot \Sigma[12||] - \Sigma[12|3].$$

Computer algebra proves laws of probabilistic reasoning

The following inference rule is valid for all positive definite 5×5 matrices:

$$[12\,|\,] \wedge [14\,|\,5] \wedge [23\,|\,5] \wedge [35\,|\,1] \wedge [45\,|\,2] \wedge [15\,|\,23] \wedge [34\,|\,12] \wedge [24\,|\,135] \;\Rightarrow\; [25\,|\,] \vee [34\,|\,].$$

Computer algebra proves laws of probabilistic reasoning

The following inference rule is valid for all positive definite 5×5 matrices:

$$[12\,|\,] \wedge [14\,|\,5] \wedge [23\,|\,5] \wedge [35\,|\,1] \wedge [45\,|\,2] \wedge [15\,|\,23] \wedge [34\,|\,12] \wedge [24\,|\,135] \;\Rightarrow\; [25\,|\,] \vee [34\,|\,].$$

$$\begin{gathered}
[25\,|\,][34\,|\,] \cdot [1][2][3][15] = \\
\big(cd^2egr + bd^2fgr - ad^2grh - 2cd^2e^2i - 2bd^2efi - 2pdfgri + 2ad^2ehi + 2pdefi^2 - 2pdqhi^2 + 2pcqi^3 + \\
2pdqrij - 2pbqi^2j - pcegrt + pbfgrt + pagrht + 2pce^2it - 2pcqrit + 2pbqhit - 2paehit\big) \cdot [12\,|\,] + \\
\big(pdqer + pbqgr - 2pbqei\big) \cdot [14\,|\,5] - \big(pcdqr + p^2fgr - 2pbcqi + 2pb^2qj - 2p^2qrj\big) \cdot [23\,|\,5] + \\
\big(cdqgr - 2cdqei + 2pqghi - 2pqfi^2 - pqgrj + 2pqeij - 2pe^2ft + 2pqfrt\big) \cdot [35\,|\,1] + \\
\big(pd^2er - 2pbdei + p^2gri + 2pb^2et - 2p^2ert\big) \cdot [45\,|\,2] - \big(2pdfi - 2pbft\big) \cdot [15\,|\,23] - \\
\big(d^2gr - 2d^2ei - pgrt + 2peit\big) \cdot [34\,|\,12] - 2pqi \cdot [24\,|\,135].
\end{gathered}$$

Computer algebra proves laws of probabilistic reasoning

```
R = QQ[p,a,b,c,d, q,e,f,g, r,h,i, s,j, t];
X = genericSymmetricMatrix(R,p,5);
I = ideal(
  det X_{0}^{1}, det X_{0,3}^{2,3}, det X_{0,4}^{3,4},
  det X_{1,4}^{2,4}, det X_{2,0}^{4,0}, det X_{3,1}^{4,1},
  det X_{0,1,2}^{4,1,2}, det X_{2,0,1}^{3,0,1},
  det X_{1,0,2,4}^{3,0,2,4}
);
U = g*h*p*q*r*(p*t-d^2); -- [25||][34||]·[1][2][3][15] ∈ monoid(𝒱)
U % I --> 0, meaning monoid(𝒱) ∩ ideal(𝒱) ≠ ∅ in ℚ[X]
-- Get a proof that U is in I:
G = gens I; -- the equations generating ideal(𝒱)
H = U // G; -- linear combinators for U from G
U == G*H --> true
```

Consistency checking is hard

The complexity class $\exists\mathbb{R}$ contains all decision problems which can be reduced in polynomial time to the feasibility of a semialgebraic set:

- polynomial optimization
- computational geometry
- algebraic statistics ...

Consistency checking is hard

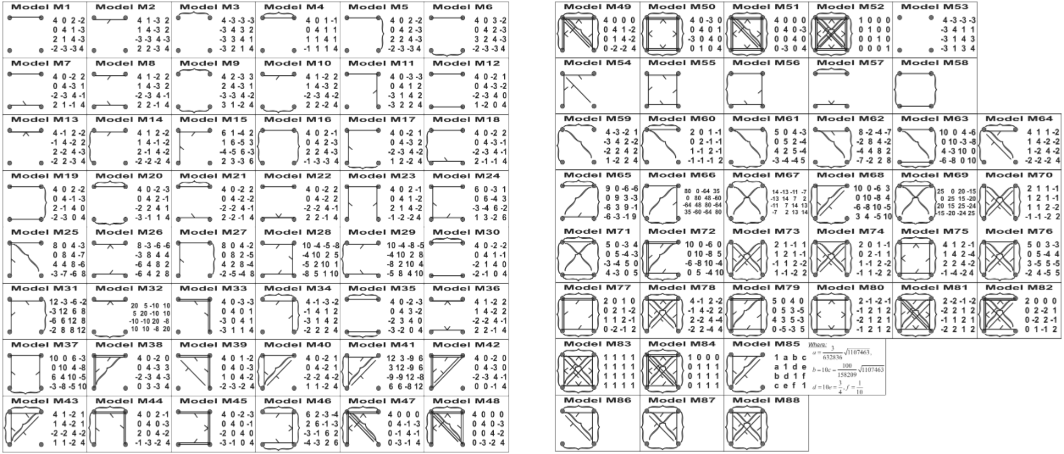
The complexity class $\exists\mathbb{R}$ contains all decision problems which can be reduced in polynomial time to the feasibility of a semialgebraic set:

- polynomial optimization
- computational geometry
- algebraic statistics ...

Theorem

The problem of deciding whether a general CI model is non-empty is complete for $\exists\mathbb{R}$.

(Graphical models are always consistent.)



Petr Šimeček. "Gaussian representation of independence models over four random variables". In: *COMPSTAT conference*. 2006

Consistency certification is hard

Šimeček's Question (2006)

Does every non-empty Gaussian CI model contain a rational point?

Or: can every wrong inference rule be refuted over $\mathbb{Q}$?

Consistency certification is hard

Šimeček's Question (2006)

Does every non-empty Gaussian CI model contain a rational point?

Or: can every wrong inference rule be refuted over $\mathbb{Q}$?

Model M85

1	a	b	c
a	1	d	e
b	d	1	f
c	e	f	1

Where:

$$a = \frac{3}{632836}\sqrt{1107463},$$

$$b = 10c = \frac{100}{158209}\sqrt{1107463}$$

$$d = 10e = \frac{3}{4}, f = \frac{1}{10}$$

$$\begin{pmatrix} 1 & -1/17 & -49/51 & -7/17 \\ -1/17 & 1 & 1/3 & 1/7 \\ -49/51 & 1/3 & 1 & 3/7 \\ -7/17 & 1/7 & 3/7 & 1 \end{pmatrix}$$

Consistency certification is hard

Šimeček's Question (2006)

Does every non-empty Gaussian CI model contain a rational point?

Or: can every wrong inference rule be refuted over $\mathbb{Q}$?

Theorem

For every finite real extension $\mathbb{K}$ of $\mathbb{Q}$ there exists a CI model $\mathcal{M}$ such that $\mathcal{M} \cap \mathrm{PD}_N(\mathbb{K}) \neq \varnothing$ but $\mathcal{M} \cap \mathrm{PD}_N(\mathbb{L}) = \varnothing$ for all proper subfields $\mathbb{L} \subsetneq \mathbb{K}$.

(Graphical models always have rational points.)

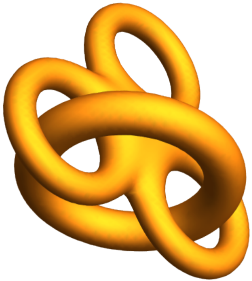
Model topology can be bad

An oriented CI model is specified by sign constraints on partial correlations.

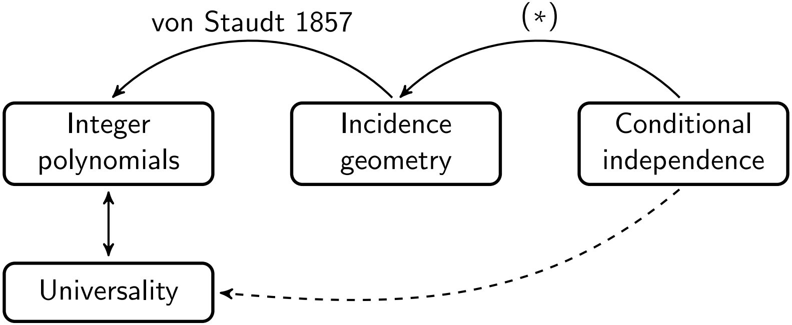
Theorem

For every primary basic semialgebraic set Z there exists an oriented CI model $\mathcal{M}$ which is homotopy-equivalent to Z.

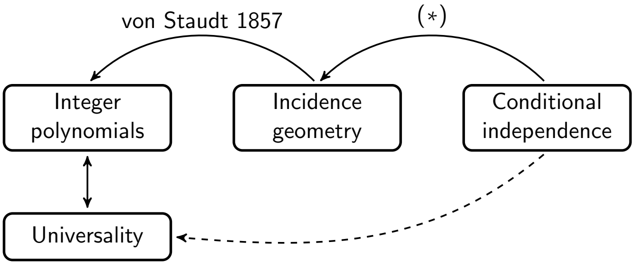
(Graphical models are always contractible.)



Universality theorems



Universality theorems



- Realization spaces of rank-3 matroids
- Realization spaces of 4-polytopes
- Nash equilibria of 3-person games
- Gaussian CI models with conditioning sets of size up to 3 ...

References

Tobias Boege. “The Gaussian conditional independence inference problem”. PhD thesis. OvGU Magdeburg, 2022.

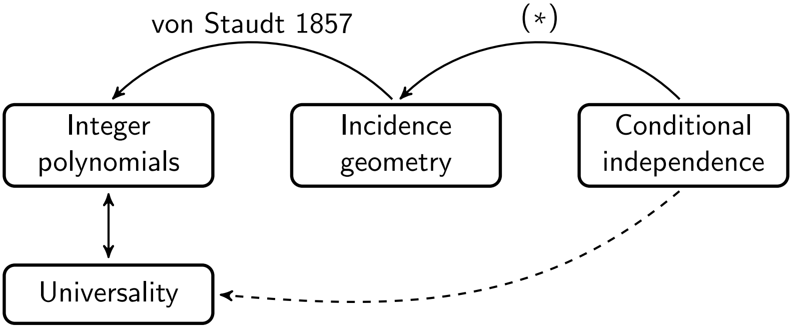
Jürgen Bokowski and Bernd Sturmfels. *Computational synthetic geometry*. Vol. 1355. Lecture Notes in Mathematics. Springer, 1989.

Ruchira S. Datta. “Universality of Nash equilibria”. In: *Math. Oper. Res.* 28.3 (2003), pp. 424–432. ISSN: 0364-765X.

Jürgen Richter-Gebert. *Perspectives on projective geometry. A guided tour through real and complex geometry*. Springer, 2011, pp. xxii + 571.

Petr Šimeček. “Gaussian representation of independence models over four random variables”. In: *COMPSTAT conference*. 2006.

Universality theorems: Background



Theorem

To every polynomial system $\{f_i \bowtie 0\}$ there is a set of CI constraints which has a model over a field $\mathbb{K}/\mathbb{Q}$ if and only if the polynomial system has a solution in $\mathbb{K}$.

Very special polynomials

$\Sigma[\,ij\,|\,] = x_{ij} \to$ impose $x_{kl} = x_{km} = x_{lm} = 0$ on a correlation matrix, then:

$$
\begin{aligned}
\Sigma[\,ij\,|\,klm] = {} & x_{ij}x_{kk}x_{ll}x_{mm} + x_{im}x_{jm}\underline{x_{kl}^2} - x_{im}x_{jl}\underline{x_{kl}x_{km}} - x_{il}x_{jm}\underline{x_{kl}x_{km}} + x_{il}x_{jl}\underline{x_{km}^2} \\
& - x_{im}x_{jm}x_{kk}x_{ll} + x_{im}x_{jk}\underline{x_{km}}x_{ll} + x_{ik}x_{jm}\underline{x_{km}}x_{ll} - x_{ij}\underline{x_{km}^2}x_{ll} \\
& + x_{im}x_{jl}x_{kk}\underline{x_{lm}} + x_{il}x_{jm}x_{kk}\underline{x_{lm}} - x_{im}x_{jk}\underline{x_{kl}x_{lm}} - x_{ik}x_{jm}\underline{x_{kl}x_{lm}} \\
& - x_{il}x_{jk}\underline{x_{km}x_{lm}} - x_{ik}x_{jl}\underline{x_{km}x_{lm}} + 2x_{ij}\underline{x_{kl}x_{km}x_{lm}} + x_{ik}x_{jk}\underline{x_{lm}^2} \\
& - x_{ij}x_{kk}\underline{x_{lm}^2} - x_{il}x_{jl}x_{kk}x_{mm} + x_{il}x_{jk}\underline{x_{kl}}x_{mm} + x_{ik}x_{jl}\underline{x_{kl}}x_{mm} \\
& - x_{ij}\underline{x_{kl}^2}x_{mm} - x_{ik}x_{jk}x_{ll}x_{mm} \\
= {} & x_{ij} - \sum_{k=k,l,m} x_{ik}x_{jk} \quad = \quad x_{ij} - \left\langle \begin{pmatrix} x_{ik} \\ x_{il} \\ x_{im} \end{pmatrix}, \begin{pmatrix} x_{jk} \\ x_{jl} \\ x_{jm} \end{pmatrix} \right\rangle .
\end{aligned}
$$

The rest is 19$^{\text{th}}$ century projective geometry. Keyword: *von Staudt constructions*.

Covariance matrix simulating a projective plane

$$
\begin{array}{c|ccc|ccc|ccc|}
 & \mathrm{p}_1 & \cdots & \mathrm{p}_n & \mathrm{l}_1 & \cdots & \mathrm{l}_m & \mathrm{x} & \mathrm{y} & \mathrm{z} \\
\hline
\mathrm{p}_1 & p_1^* & & \langle p, p' \rangle & & & & p_1^x & p_1^y & p_1^z \\
\vdots & & \ddots & & & \langle p, \ell \rangle & & & \vdots & \\
\mathrm{p}_n & \langle p', p \rangle & & p_n^* & & & & p_n^x & p_n^y & p_n^z \\
\hline
\mathrm{l}_1 & & & & \ell_1^* & & \langle \ell, \ell' \rangle & \ell_1^x & \ell_1^y & \ell_1^z \\
\vdots & & \langle \ell, p \rangle & & & \ddots & & & \vdots & \\
\mathrm{l}_m & & & & \langle \ell', \ell \rangle & & \ell_m^* & \ell_m^x & \ell_m^y & \ell_m^z \\
\hline
\mathrm{x} & p_1^x & & p_n^x & \ell_1^x & & \ell_m^x & x^* & 0 & 0 \\
\mathrm{y} & p_1^y & \cdots & p_n^y & \ell_1^y & \cdots & \ell_m^y & 0 & y^* & 0 \\
\mathrm{z} & p_1^z & & p_n^z & \ell_1^z & & \ell_m^z & 0 & 0 & z^*
\end{array}
$$