# The complexity of Gaussian conditional independence inference

Tobias Boege

Aalto University
School of Science

Algebraic Structures in Statistical Methodology
Mathematisches Forschungsinstitut Oberwolfach,
4–10 December 2022

# Conditional independence models

### Definition

A *CI constraint* is a CI statement $[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ or its negation $\neg[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ constraining a random vector $\xi$.

# Conditional independence models

**Definition**

A *CI constraint* is a CI statement $[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ or its negation $\neg[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ constraining a random vector $\xi$.

- How hard is it to decide if a set of constraints is consistent?
- How hard is it to certify consistency by exhibiting a distribution?

# Conditional independence models

## Definition

A *CI constraint* is a CI statement $[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ or its negation $\neg[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ constraining a random vector $\xi$.

- How hard is it to decide if a set of constraints is consistent?
- How hard is it to certify consistency by exhibiting a distribution?

Dual view: Conditional independence inference

## Definition

A *CI inference formula* is a Boolean formula in implication form whose variables are CI statements: $\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q}$, where $\mathcal{P}$ and $\mathcal{Q}$ are sets of CI statements.

# Conditional independence models

**Definition**

A *CI constraint* is a CI statement $[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ or its negation $\neg[\xi_i \perp\!\!\!\perp \xi_j \mid \xi_K]$ constraining a random vector $\xi$.

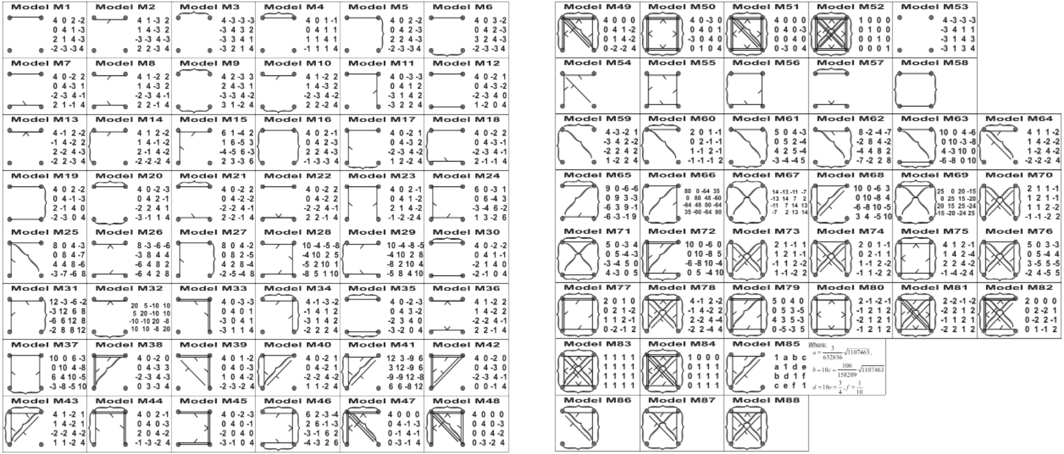
- How hard is it to decide if a set of constraints is consistent?
- How hard is it to certify consistency by exhibiting a distribution?

Dual view: Conditional independence inference

**Definition**

A *CI inference formula* is a Boolean formula in implication form whose variables are CI statements: $\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q}$, where $\mathcal{P}$ and $\mathcal{Q}$ are sets of CI statements.

- How hard is it to decide if an implication is valid?
- How hard is it to certify validity and with what data?




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

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

- $\Sigma$ 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

The *model* of a set of CI constraints is the set of all PD matrices which satisfy them.

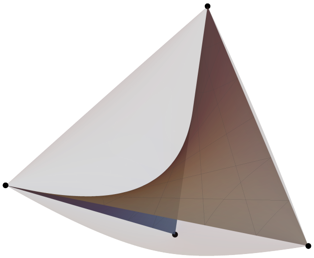


Figure: Model of $\Sigma[12\,|\,3] = a - bc = 0$ in the $3 \times 3$ correlation matrices.

# Gaussian CI models

## Definition

The *model* of a set of CI constraints is the set of all PD matrices which satisfy them.

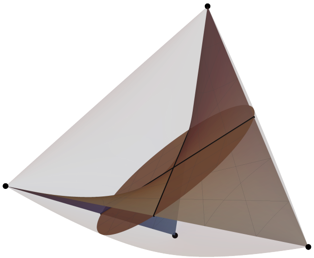


Figure: Model of $\Sigma[12|\,] = a = 0$ and $\Sigma[12|3] = a - bc = 0$ in the $3 \times 3$ correlation matrices.

# 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 $\mathcal{Q}$:

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

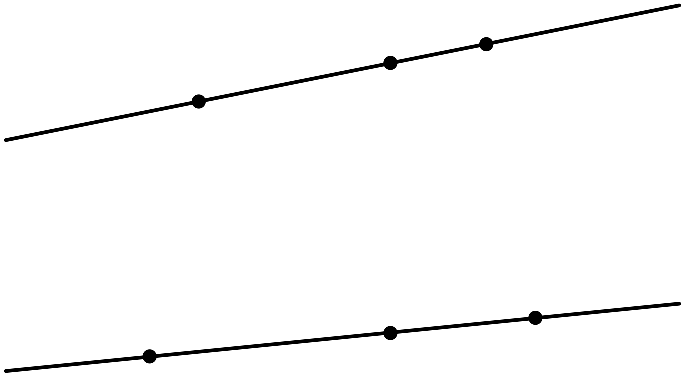
# Models and inference

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

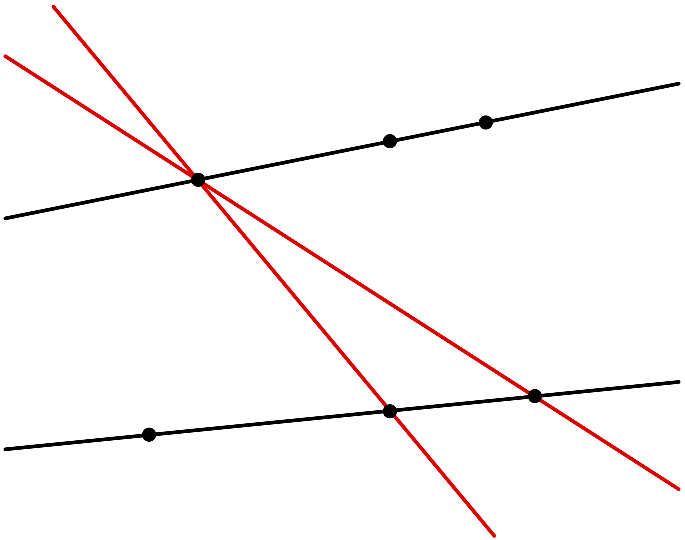
$$\bigwedge \mathcal{P} \Rightarrow \bigvee \mathcal{Q} \text{ is not valid} \iff \mathcal{P} \cup \neg\mathcal{Q} \text{ has a point}$$

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.

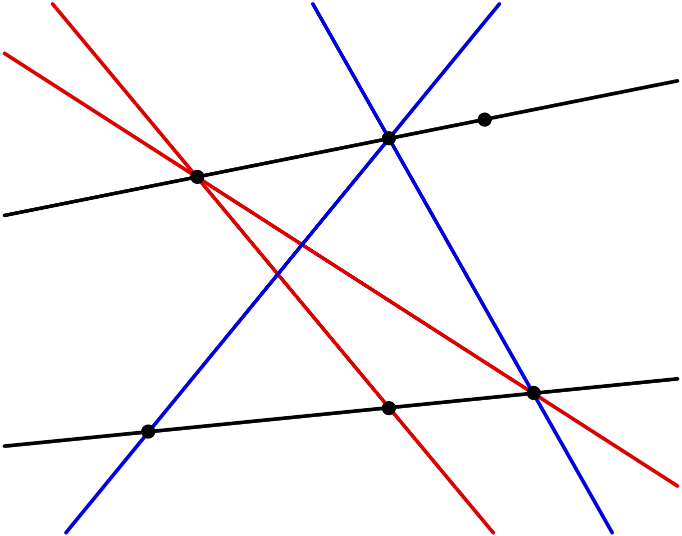
# For ancient geometers: conditional independence ≈ collinearity



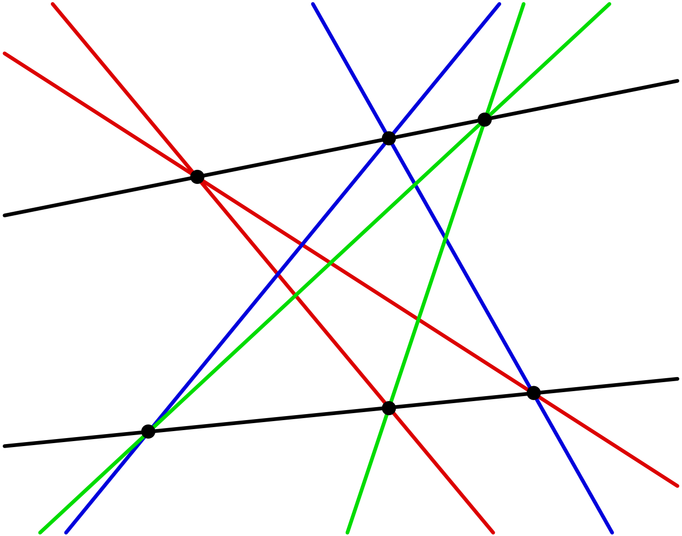
# For ancient geometers: conditional independence ≈ collinearity



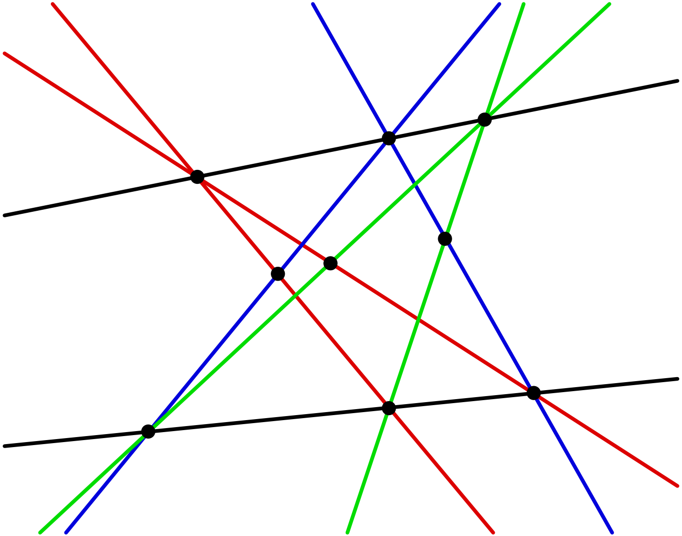
# For ancient geometers: conditional independence ≈ collinearity



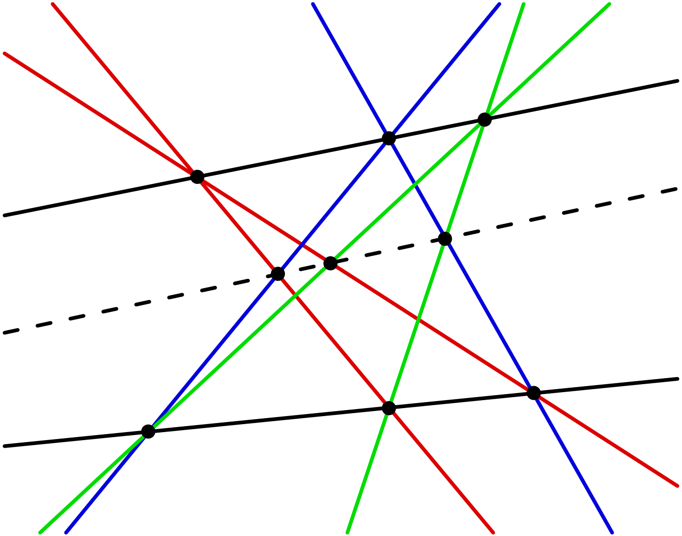
# For ancient geometers: conditional independence ≈ collinearity



# 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 $\Sigma$ 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 $f_i, g_j, h_k \in \mathbb{Z}[t_1, \dots, t_k]$ 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* $f \in \text{ideal}(f_i)$, $g \in \text{cone}(g_j)$ *and* $h \in \text{monoid}(h_k)$ *such that* $g + h^2 = f$.

# Normal form for proofs and refutations

Let $f_i, g_j, h_k \in \mathbb{Z}[t_1, \dots, t_k]$ 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 $5 \times 5$ 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 $5 \times 5$ 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 Gaussian CI model is non-empty is $\exists\mathbb{R}$-complete.*

# 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 $\mathcal{M} \neq \varnothing$ but $\mathcal{M} \cap \mathrm{PD}_N(\mathbb{K}) = \varnothing$.*

# Very special polynomials revisited

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

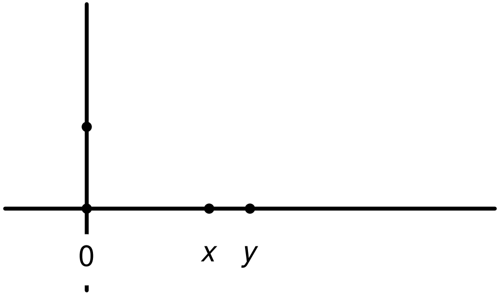
$$\begin{aligned}
\Sigma[\,ij\,|\,klm] = x_{ij}&x_{kk}x_{ll}x_{mm} + x_{im}x_{jm}\underline{x_{kl}^2} - x_{im}x_{jl}\underline{x_{kl}x_{km}} - x_{il}x_{jm}\underline{x_{kl}x_{km}} + x_{il}x_{jl}\underline{x_{km}^2}\\
&- x_{im}x_{jm}x_{kk}x_{ll} + x_{im}x_{jk}\underline{x_{km}}x_{ll} + x_{ik}x_{jm}\underline{x_{km}}x_{ll}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

# Covariance matrix simulating a projective plane

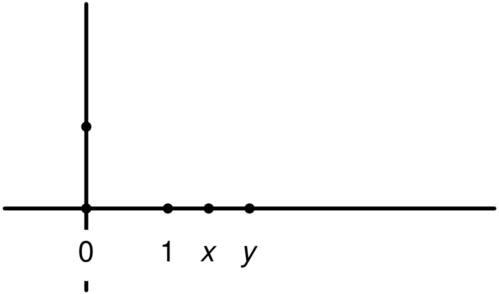
$$
\begin{array}{c} \\ \mathrm{p}_1 \\ \vdots \\ \mathrm{p}_n \\ \mathrm{l}_1 \\ \vdots \\ \mathrm{l}_m \\ \mathrm{x} \\ \mathrm{y} \\ \mathrm{z} \end{array}
\begin{array}{c}
\begin{array}{ccc|ccc|ccc}
\mathrm{p}_1 & \cdots & \mathrm{p}_n & \mathrm{l}_1 & \cdots & \mathrm{l}_m & \mathrm{x} & \mathrm{y}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

The rest is $19^{\text{th}}$ century projective geometry.

# Von Staudt constructions

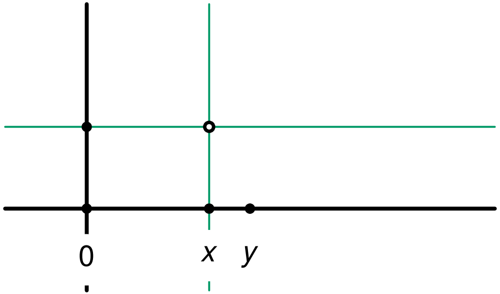



Addition

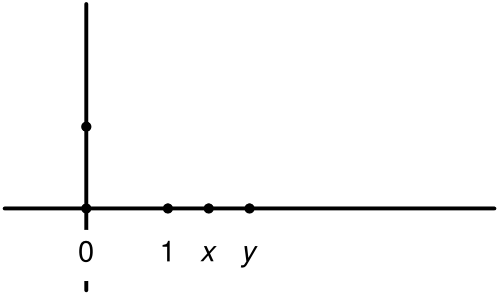



Multiplication

# Von Staudt constructions

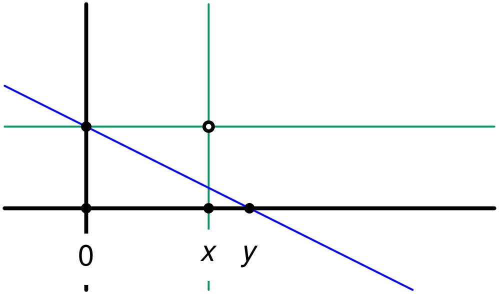



Addition

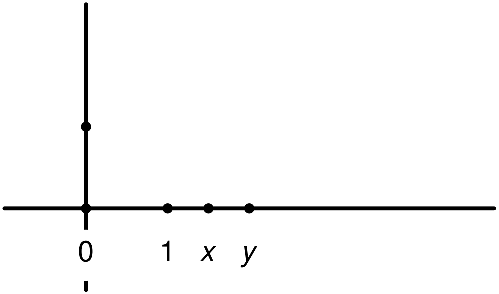



Multiplication

# Von Staudt constructions

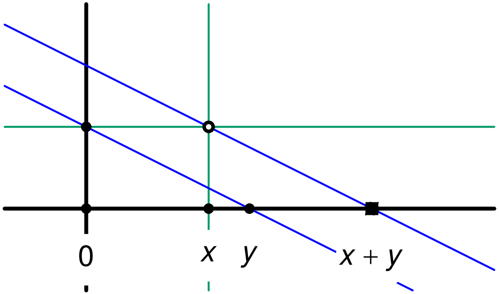



Addition

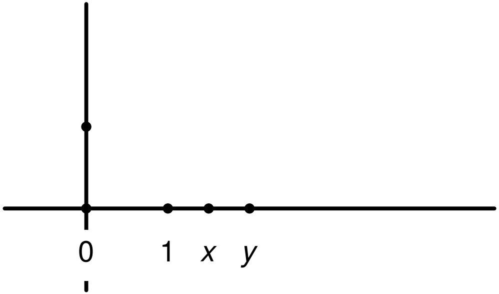



Multiplication

# Von Staudt constructions

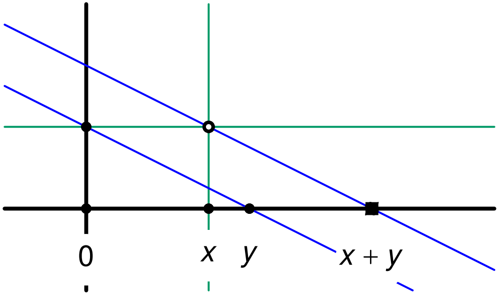



Addition

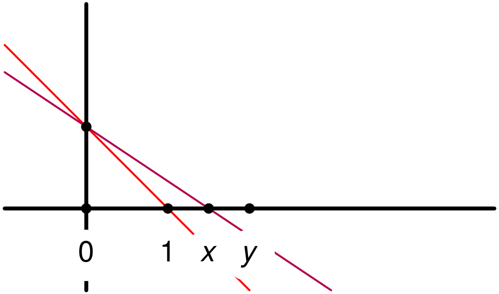



Multiplication

# Von Staudt constructions

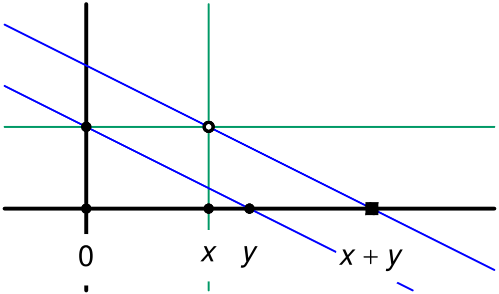



Addition

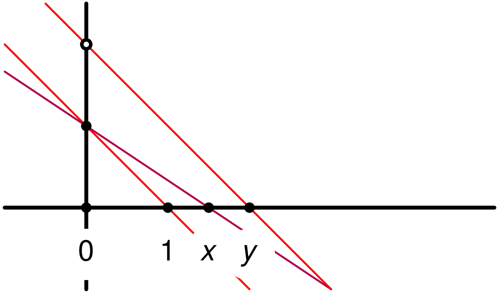



Multiplication

# Von Staudt constructions

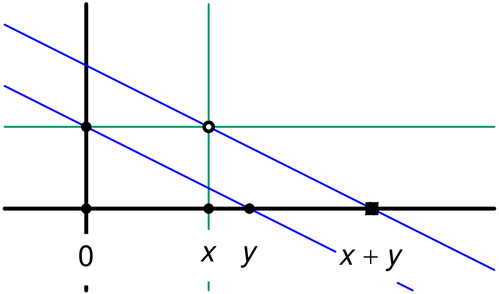



Addition

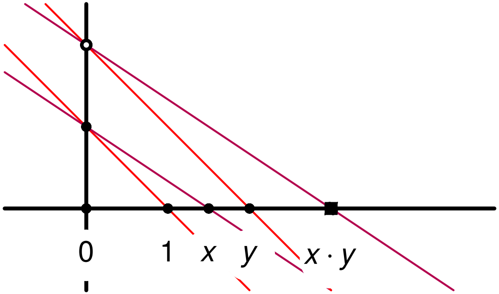



Multiplication

# Von Staudt constructions

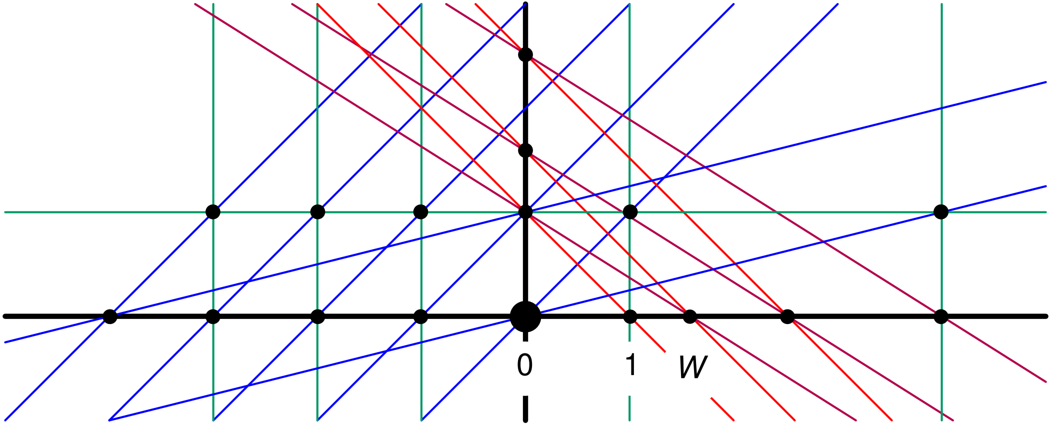



Addition

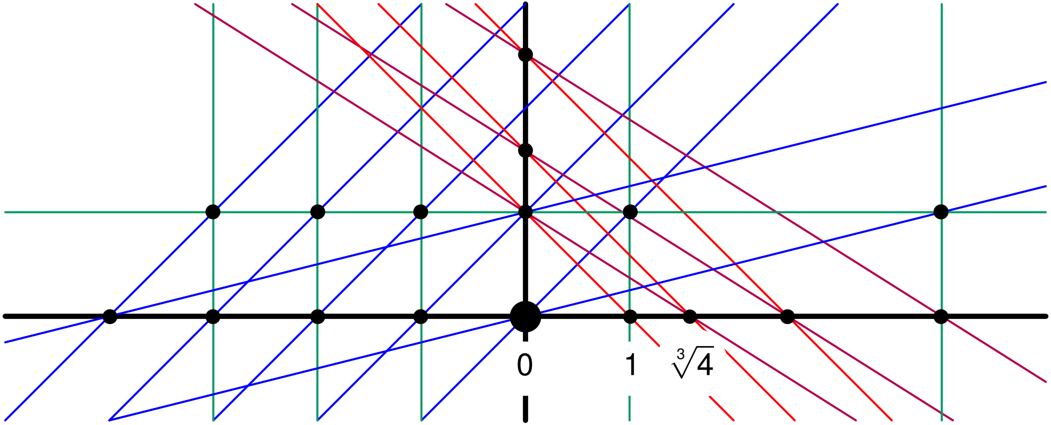



Multiplication

# Where is Waldo?



# Where is Waldo? On the cube root of 4!



# Bonus question

**Matúš's Question (1999); also Sturmfels (2007)**

*Can every invalid inference for discrete CI structures be refuted over* $\mathbb{Q}$*?*

# Bonus question

**Matúš's Question (1999); also Sturmfels (2007)**

*Can every invalid inference for discrete CI structures be refuted over* $\mathbb{Q}$*?*

```
load "simecek.m2";
f = binaryMomentMap 4; -- ... more setup

-- (Linear slice of) CI equations of binary RVs in moment coordinates
Eqns {{1},{2},{}}     --> e_12
Eqns {{1},{2},{3}}    --> e_12 - e_13 e_23
Eqns {{1},{2},{3,4}}  --> e_12 - e_13 e_23 - e_14 e_24 + e_34 e_1234 and
                      --> e_1234 - e_14 e_23 - e_13 e_24 + e_12 e_34
```

# 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.
- 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.