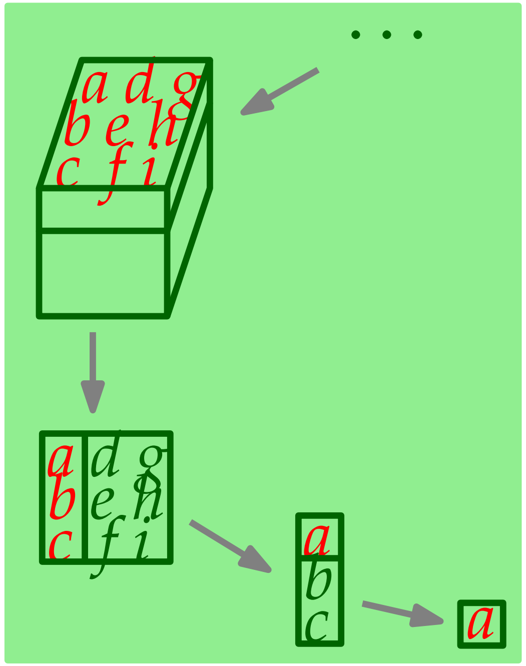
Symmetries and ∞-dim limits of algebro-statistical models

$$\begin{matrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & & & \\ a_{31} & & & \\ \vdots & & & \end{matrix}$$



Jan Draisma

TU Eindhoven

Part I: some infinite-dimensional commutative algebra

What is an infinite-dimensional variety?

inductive limits of finite-dimensional varieties,
projective limits, spectra of infinite-dimensional rings, etc.

What is an infinite-dimensional variety?

inductive limits of finite-dimensional varieties,
projective limits, spectra of infinite-dimensional rings, etc.

This talk

V: countable-dimensional space over $\mathbb{C}$ (or $\mathbb{R}$) of *coordinates*
V^*: dual space, topological space with Zariski topology
Closed subsets $X \subseteq V^*$ are called *infinite-dimensional varieties*.

Example

$V = \langle x_{ij} \mid i, j \in \mathbb{N} \rangle$, $X \subseteq V^*$ defined by equations $x_{ij}x_{kl} - x_{il}x_{kj}$

What is an infinite-dimensional variety?

inductive limits of finite-dimensional varieties,
projective limits, spectra of infinite-dimensional rings, etc.

This talk

V: countable-dimensional space over $\mathbb{C}$ (or $\mathbb{R}$) of *coordinates*
V^*: dual space, topological space with Zariski topology
Closed subsets $X \subseteq V^*$ are called *infinite-dimensional varieties*.

Example

$V = \langle x_{ij} \mid i, j \in \mathbb{N} \rangle$, $X \subseteq V^*$ defined by equations $x_{ij}x_{kl} - x_{il}x_{kj}$

Sequence model

If $V_1 \subseteq V_2 \subseteq \ldots$ finite-dimensional with $V = \bigcup_i V_i$, then
$V^* = \lim_{\leftarrow} V_i^*$ with $V_1^* \twoheadleftarrow V_2^* \twoheadleftarrow \cdots$
(*both as set and as topological space*)

Noetherianity modulo group actions

Assume a group G acts by linear transformations on V
$\rightsquigarrow$ G acts on SV by algebra auto and on V^* by homeo.

Noetherianity modulo group actions

Assume a group G acts by linear transformations on V
$\rightsquigarrow$ G acts on SV by algebra auto and on V^* by homeo.

For $X \subseteq V^*$ a closed set, $\mathbb{C}[X] := SV/I(X)$, where $I(X) \subseteq SV$ is the ideal of polynomials vanishing on X.

Noetherianity modulo group actions

Assume a group G acts by linear transformations on V
$\rightsquigarrow$ G acts on SV by algebra auto and on V^* by homeo.

For $X \subseteq V^*$ a closed set, $\mathbb{C}[X] := SV/I(X)$, where $I(X) \subseteq SV$ is the ideal of polynomials vanishing on X.

Assume $I(X)$ is G-stable, so that G acts on $\mathbb{C}[X]$.

Noetherianity modulo group actions

Assume a group G acts by linear transformations on V
$\rightsquigarrow$ G acts on SV by algebra auto and on V^* by homeo.

For $X \subseteq V^*$ a closed set, $\mathbb{C}[X] := SV/I(X)$, where $I(X) \subseteq SV$ is the ideal of polynomials vanishing on X.

Assume $I(X)$ is G-stable, so that G acts on $\mathbb{C}[X]$.

Definition
$\mathbb{C}[X]$ is *G-Noetherian* if every chain $I_1 \subseteq I_2 \subseteq \ldots$ of G-stable ideals stabilises (⇔ *each G-stable ideal is G-finitely generated.*)

Noetherianity modulo group actions

Assume a group G acts by linear transformations on V
$\rightsquigarrow G$ acts on SV by algebra auto and on V^* by homeo.

For $X \subseteq V^*$ a closed set, $\mathbb{C}[X] := SV/I(X)$, where $I(X) \subseteq SV$ is the ideal of polynomials vanishing on X.

Assume $I(X)$ is G-stable, so that G acts on $\mathbb{C}[X]$.

Definition
$\mathbb{C}[X]$ is *G-Noetherian* if every chain $I_1 \subseteq I_2 \subseteq \ldots$ of G-stable ideals stabilises (⇔ *each G-stable ideal is G-finitely generated.*)

X is *G-Noetherian* if every chain $X_1 \supseteq X_2 \supseteq \ldots$ of G-stable closed subsets stabilises

(⇔ *each G-stable subvariety of X defined by fin many G-orbits of eqs.*)

Assume a group G acts by linear transformations on V
$\rightsquigarrow$ G acts on SV by algebra auto and on V^* by homeo.

For $X \subseteq V^*$ a closed set, $\mathbb{C}[X] := SV/I(X)$, where $I(X) \subseteq SV$ is the ideal of polynomials vanishing on X.

Assume $I(X)$ is G-stable, so that G acts on $\mathbb{C}[X]$.

Definition

$\mathbb{C}[X]$ is *G-Noetherian* if every chain $I_1 \subseteq I_2 \subseteq \ldots$ of G-stable ideals stabilises (⇔ *each G-stable ideal is G-finitely generated.*)

$\Longrightarrow$

X is *G-Noetherian* if every chain $X_1 \supseteq X_2 \supseteq \ldots$ of G-stable closed subsets stabilises

(⇔ *each G-stable subvariety of X defined by fin many G-orbits of eqs.*)

(Non-)Examples of G-Noetherianity

Finite-by-infinite matrices

Fix $k \in \mathbb{N}$;

$\mathrm{Sym}(\mathbb{N})$ acts on $V = \langle x_{ij} \mid i \in [k], j \in \mathbb{N} \rangle$

by $\pi(x_{ij}) = x_{i\pi(j)}$.

$$\begin{matrix} x_{11} & x_{12} & \cdots \\ \vdots & & \\ x_{k1} & x_{k2} & \cdots \end{matrix}$$

Finite-by-infinite matrices

$$\begin{matrix} x_{11} & x_{12} & \cdots \\ \vdots & & \\ x_{k1} & x_{k2} & \cdots \end{matrix}$$

Fix $k \in \mathbb{N}$;
$\mathrm{Sym}(\mathbb{N})$ acts on $V = \langle x_{ij} \mid i \in [k], j \in \mathbb{N} \rangle$
by $\pi(x_{ij}) = x_{i\pi(j)}$.

Theorem [Cohen 87, Hillar-Sullivant 09]
$\mathbb{C}[x_{ij} \mid i \in [k], j \in \mathbb{N}] = \mathbb{C}[V^*]$ is $\mathrm{Sym}(\mathbb{N})$-Noetherian.

Finite-by-infinite matrices

$$\begin{matrix} x_{11} & x_{12} & \cdots \\ \vdots & & \\ x_{k1} & x_{k2} & \cdots \end{matrix}$$

Fix $k \in \mathbb{N}$;
$\mathrm{Sym}(\mathbb{N})$ acts on $V = \langle x_{ij} \mid i \in [k], j \in \mathbb{N} \rangle$
by $\pi(x_{ij}) = x_{i\pi(j)}$.

Theorem [Cohen 87, Hillar-Sullivant 09]
$\mathbb{C}[x_{ij} \mid i \in [k], j \in \mathbb{N}] = \mathbb{C}[V^*]$ is $\mathrm{Sym}(\mathbb{N})$-Noetherian.

Infinite-by-infinite matrices

$$\begin{matrix} x_{11} & x_{12} & \cdots \\ x_{21} & & \\ \vdots & & \end{matrix}$$

$\mathrm{Sym}(\mathbb{N})$ acts by $\pi(x_{ij}) = x_{\pi(i),\pi(j)}$
$\rightsquigarrow \mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}]$ is *not* $\mathrm{Sym}(\mathbb{N})$-Noetherian;
e.g. the $\mathrm{Sym}(\mathbb{N})$-stable ideal generated by
$x_{12}x_{21},\ x_{12}x_{23}x_{31},\ x_{12}x_{23}x_{34}x_{41},\ \ldots$
is not $\mathrm{Sym}(\mathbb{N})$-finitely generated.

(*neither* $\mathrm{Sym}(\mathbb{N}) \times \mathrm{Sym}(\mathbb{N})$-*Noetherian*)

(Non-)Examples, continued

Theorem (*Matrices of bounded rank*)
$\mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}]/(\text{all } (k+1) \times (k+1)\text{-subdeterminants})$
is $\mathrm{Sym}(\mathbb{N})$-Noetherian.
(*uses* $2k \times \mathbb{N}$*-matrices and the FFT, SFT for* GL_k)

Theorem (*Matrices of bounded rank*)
$\mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}]/(\text{all } (k+1) \times (k+1)\text{-subdeterminants})$
is Sym($\mathbb{N}$)-Noetherian.
(*uses* $2k \times \mathbb{N}$*-matrices and the FFT, SFT for* GL_k)

Larger groups
$\mathrm{GL}_{\mathbb{N}} := \{$invertible $\mathbb{N} \times \mathbb{N}$-matrices g with almost all $g_{ii} = 1$ and almost all $g_{ij} = 0 (i \neq j)\}$.

(Non-)Examples, continued

Theorem (*Matrices of bounded rank*)
$\mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}]/(\text{all } (k+1) \times (k+1)\text{-subdeterminants})$
is Sym($\mathbb{N}$)-Noetherian.
(*uses* $2k \times \mathbb{N}$*-matrices and the FFT, SFT for* GL_k)

Larger groups
$\mathrm{GL}_{\mathbb{N}} :=$ {invertible $\mathbb{N} \times \mathbb{N}$-matrices g with almost all $g_{ii} = 1$ and almost all $g_{ij} = 0 (i \neq j)$}.

Theorem (*symmetric matrices*) [Snowden-Sam 2012]
$\mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}, x_{ij} = x_{ji}]$ is $\mathrm{GL}_{\mathbb{N}}$-Noetherian via $g \circ x = gxg^T$.

(Non-)Examples, continued

Theorem (*Matrices of bounded rank*)
$\mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}]/(\text{all } (k+1) \times (k+1)\text{-subdeterminants})$
is Sym($\mathbb{N}$)-Noetherian.
(*uses* $2k \times \mathbb{N}$*-matrices and the FFT, SFT for* GL_k)

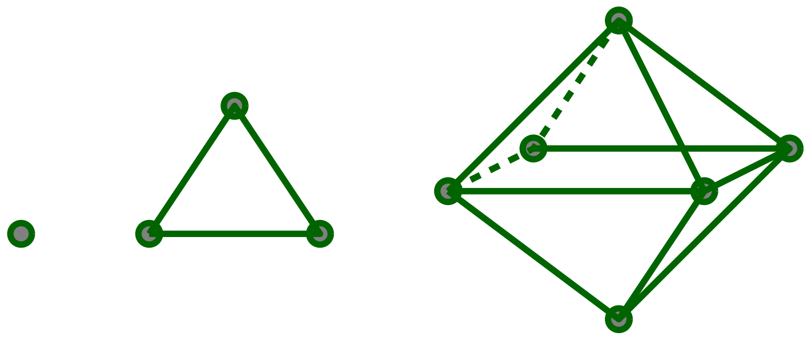
Larger groups
$\mathrm{GL}_{\mathbb{N}} := \{$invertible $\mathbb{N} \times \mathbb{N}$-matrices g with almost all $g_{ii} = 1$ and almost all $g_{ij} = 0 (i \neq j)\}$.

Theorem (*symmetric matrices*) [Snowden-Sam 2012]
$\mathbb{C}[x_{ij} \mid i, j \in \mathbb{N}, x_{ij} = x_{ji}]$ is $\mathrm{GL}_{\mathbb{N}}$-Noetherian via $g \circ x = gxg^T$.

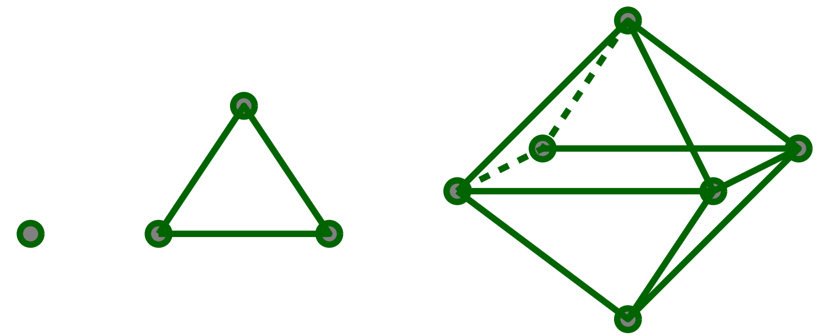
Theorem [D-Eggermont 2014]
$(\mathbb{C}^{\mathbb{N} \times \mathbb{N}})^p$ is $\mathrm{GL}_{\mathbb{N}} \times \mathrm{GL}_{\mathbb{N}}$-Noetherian for each p, via
$(g, h) \circ (x, \ldots, z) := (gxh^{-1}, \ldots, gzh^{-1})$.

Example: second hypersimplex

$P_n := \{v_{ij} = e_i + e_j \mid 1 \leq i \neq j \leq n\}$



Finitely generated symmetric Markov bases



Example: second hypersimplex

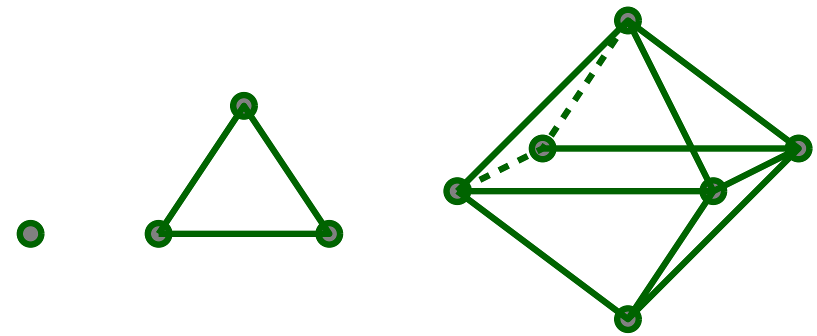
$P_n := \{v_{ij} = e_i + e_j \mid 1 \leq i \neq j \leq n\}$

Markov basis M_n [De Loera-Sturmfels-Thomas 1995]

$v_{ij} = v_{ji}$ and $v_{ij} + v_{kl} = v_{il} + v_{kj}$ for i, j, k, l distinct

$\rightsquigarrow$ if $\sum_{ij} c_{ij} v_{ij} = \sum_{ij} d_{ij} v_{ij}$ with $c_{ij}, d_{ij} \in \mathbb{Z}_{\geq 0}$,
then the expressions are connected by such *moves* without creating negative coefficients

Finitely generated symmetric Markov bases

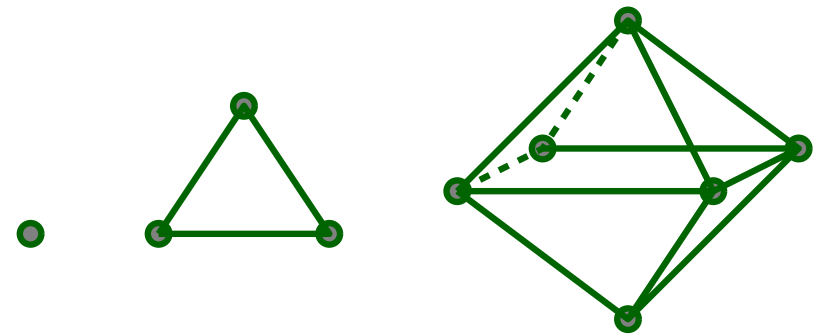


Example: second hypersimplex
$P_n := \{v_{ij} = e_i + e_j \mid 1 \leq i \neq j \leq n\}$

Markov basis M_n [De Loera-Sturmfels-Thomas 1995]
$v_{ij} = v_{ji}$ and $v_{ij} + v_{kl} = v_{il} + v_{kj}$ for i, j, k, l distinct
$\rightsquigarrow$ if $\sum_{ij} c_{ij} v_{ij} = \sum_{ij} d_{ij} v_{ij}$ with $c_{ij}, d_{ij} \in \mathbb{Z}_{\geq 0}$,
then the expressions are connected by such
moves without creating negative coefficients

Theorem [D-Eggermont-Krone-Leykin 2013]
For *any* family $(P_n \subseteq \mathbb{Z}^F \times \mathbb{Z}^{k \times n})$, F finite, if $P_n = \mathrm{Sym}(n) P_{n_0}$ for $n \geq n_0$, then $\exists n_1$: for $n \geq n_1$ has a Markov basis M_n with $M_n = \mathrm{Sym}(n) M_{n_0}$.

Example: second hypersimplex
$P_n := \{v_{ij} = e_i + e_j \mid 1 \leq i \neq j \leq n\}$



Markov basis M_n [De Loera-Sturmfels-Thomas 1995]
$v_{ij} = v_{ji}$ and $v_{ij} + v_{kl} = v_{il} + v_{kj}$ for i, j, k, l distinct
$\rightsquigarrow$ if $\sum_{ij} c_{ij} v_{ij} = \sum_{ij} d_{ij} v_{ij}$ with $c_{ij}, d_{ij} \in \mathbb{Z}_{\geq 0}$,
then the expressions are connected by such *moves* without creating negative coefficients

Theorem [D-Eggermont-Krone-Leykin 2013]
For *any* family $(P_n \subseteq \mathbb{Z}^F \times \mathbb{Z}^{k \times n})$, F finite, if $P_n = \mathrm{Sym}(n) P_{n_0}$ for $n \geq n_0$, then $\exists n_1$: for $n \geq n_1$ has a Markov basis M_n with $M_n = \mathrm{Sym}(n) M_{n_0}$.

Explicit results for width $n_0 = 2$: [Kahle-Krone-Leykin 2014]

Part II: Applications to algebro-statistical models

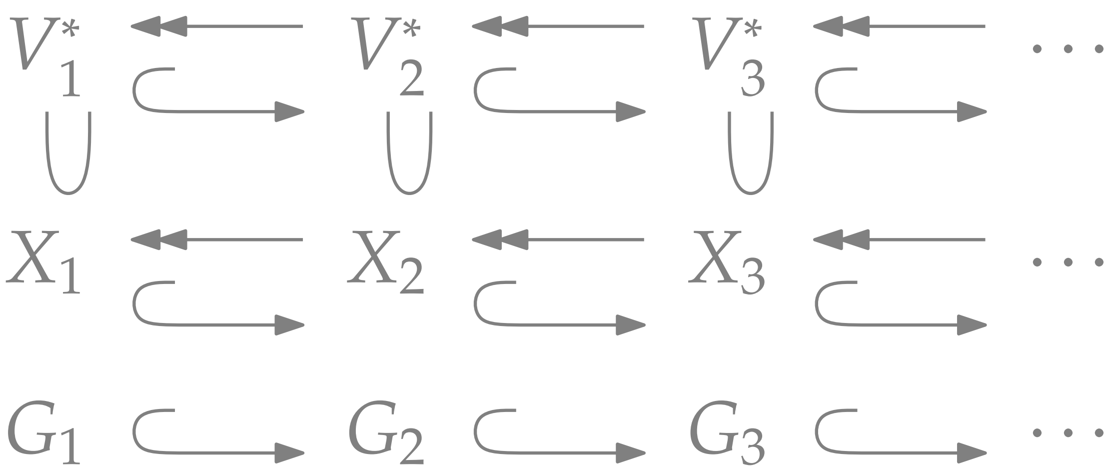
Setting

$V_1^*, V_2^*, \ldots$ fin-dim spaces; $X_i \subseteq V_i^*$ subvariety

G_i group acting linearly on V_i^* preserving X_i

$G_i \subseteq G_{i+1}$ & maps $\pi : V_{i+1}^* \to V_i^*$ and $\iota : V_i^* \to V_{i+1}^*$ both

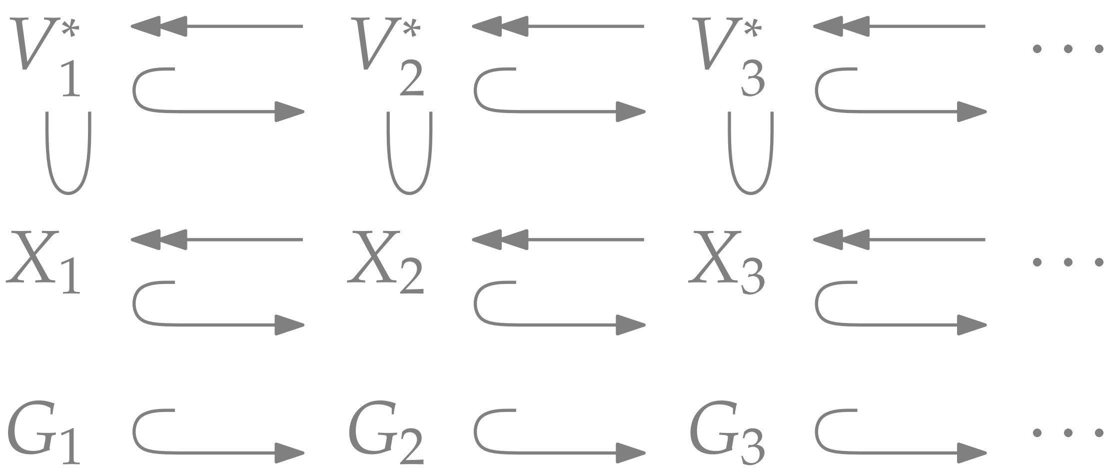
G_i-equivariant, mapping X_{i+1} into X_i and v.v. & $\pi \circ \iota = \mathrm{id}$



Constructing limits

Setting

$V_1^*, V_2^*, \ldots$ fin-dim spaces; $X_i \subseteq V_i^*$ subvariety
G_i group acting linearly on V_i^* preserving X_i
$G_i \subseteq G_{i+1}$ & maps $\pi : V_{i+1}^* \to V_i^*$ and $\iota : V_i^* \to V_{i+1}^*$ both
G_i-equivariant, mapping X_{i+1} into X_i and v.v. & $\pi \circ \iota = \mathrm{id}$



Definition

Sequence $(X_i \subseteq V_i^*)_i$ *stabilises* if for $n \gg 0$:
$p \in V_n^*$ lies in X_n
iff $\forall_{g \in G_n} \pi(gp) \in X_{n-1}$.

Constructing limits

Setting

$V_1^*, V_2^*, \ldots$ fin-dim spaces; $X_i \subseteq V_i^*$ subvariety
G_i group acting linearly on V_i^* preserving X_i
$G_i \subseteq G_{i+1}$ & maps $\pi : V_{i+1}^* \to V_i^*$ and $\iota : V_i^* \to V_{i+1}^*$ both G_i-equivariant, mapping X_{i+1} into X_i and v.v. & $\pi \circ \iota = \mathrm{id}$

$$\begin{array}{ccccccc}
V_1^* & \twoheadleftarrow \atop \hookrightarrow & V_2^* & \twoheadleftarrow \atop \hookrightarrow & V_3^* & \twoheadleftarrow \atop \hookrightarrow & \cdots \\
\cup & & \cup & & \cup & & \\
X_1 & \twoheadleftarrow \atop \hookrightarrow & X_2 & \twoheadleftarrow \atop \hookrightarrow & X_3 & \twoheadleftarrow \atop \hookrightarrow & \cdots \\
 & & & & & & \\
G_1 & \hookrightarrow & G_2 & \hookrightarrow & G_3 & \hookrightarrow & \cdots
\end{array}$$

Definition
Sequence $(X_i \subseteq V_i^*)_i$ *stabilises* if for $n \gg 0$: $p \in V_n^*$ lies in X_n iff $\forall_{g \in G_n} \pi(gp) \in X_{n-1}$.

$\rightsquigarrow V_\infty^* := \lim_\leftarrow V_n$; $X_\infty := \lim_\leftarrow X_n$; $G_\infty := \cup_n G_n$

Lemma Stabilisation is "equivalent" to: $X_\infty \subseteq V_\infty^*$ is defined by finitely many G_∞-orbits of equations.

Fixed row and column sums

$A, B \in \mathbb{Z}_{\geq 0}^{m \times n}$ with $a_{i+} = b_{i+}$ and $a_{+j} = b_{+j}$

$\Rightarrow \exists A = A_0, A_1, \ldots, A_k = B \in \mathbb{Z}_{\geq 0}^{m \times n}$ with

$A_l - A_{l-1} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \rightsquigarrow$ moves "independent" of m, n.

Fixed row and column sums

$A, B \in \mathbb{Z}_{\geq 0}^{m\times n}$ with $a_{i+} = b_{i+}$ and $a_{+j} = b_{+j}$

$\Rightarrow \exists A = A_0, A_1, \ldots, A_k = B \in \mathbb{Z}_{\geq 0}^{m\times n}$ with

$A_l - A_{l-1} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \rightsquigarrow$ moves "independent" of m, n.

Theorem [Diaconis-Sturmfels 1998]

basis of Markov moves = generating set of toric ideal

(e.g. $\ker[y_{ij} \mapsto x_i z_j]$ generated by $\{y_{ij}y_{i'j'} - y_{ij'}y_{i'j}\}$)

Fixed row and column sums

$A, B \in \mathbb{Z}_{\geq 0}^{m\times n}$ with $a_{i+} = b_{i+}$ and $a_{+j} = b_{+j}$

$\Rightarrow \exists A = A_0, A_1, \ldots, A_k = B \in \mathbb{Z}_{\geq 0}^{m\times n}$ with

$A_l - A_{l-1} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \rightsquigarrow$ moves “independent” of m, n.

Theorem [Diaconis-Sturmfels 1998]

basis of Markov moves = generating set of toric ideal

(e.g. $\ker[y_{ij} \mapsto x_i z_j]$ generated by $\{y_{ij}y_{i'j'} - y_{ij'}y_{i'j}\}$)

Conjecture [Hoşten-Sullivant 2007]

Similar stabilisation conjecture for Markov basis for sampling higher-dimensional contingency tables.

The independent set theorem, continued

Hierarchichal models

F family of subsets of $[m]$

$y(i_1, \ldots, i_m)$ and $x(S, (i_s)_{s \in S})$ for $S \in F$ variables

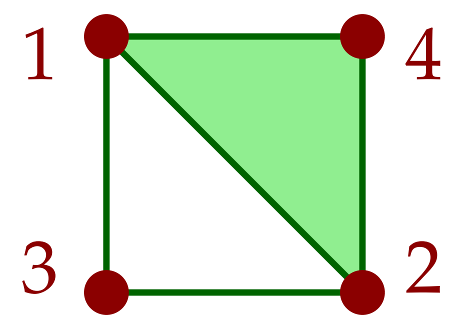
$I := \ker[y(i_1, \ldots, i_m) \mapsto \prod_{A \in S} x(S, (i_s)_{s \in S})]$

Example

$m = 4, F = \{124, 13, 23\}$

variables $y(abcd), x(abd), z(ac), u(bc)$

$I = \ker[y(abcd) \mapsto x(abd)z(ac)u(bc)]$



The independent set theorem, continued

Hierarchichal models

F family of subsets of $[m]$

$y(i_1, \ldots, i_m)$ and $x(S, (i_s)_{s \in S})$ for $S \in F$ variables

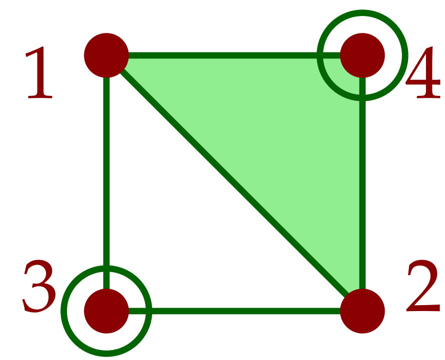
$I := \ker[y(i_1, \ldots, i_m) \mapsto \prod_{A \in S} x(S, (i_s)_{s \in S})]$

Example

$m = 4, F = \{124, 13, 23\}$

variables $y(abcd), x(abd), z(ac), u(bc)$

$I = \ker[y(abcd) \mapsto x(abd)z(ac)u(bc)]$



Theorem [Hillar-Sullivant 2012]

If $T \subseteq [m]$ *independent set* ($|T \cap S| \leq 1$ for $S \in F$);

$i_t, t \in T$ run through $\mathbb{N}$ and $i_t, t \notin T$ through $[r_t]$

$\rightsquigarrow I$ generated by finitely many $\mathrm{Inc}(\mathbb{N})$-orbits

(now this also follows from D-Eggermont-Krone-Leykin)

II: Cloning sinks in a Gaussian Bayesian model

G: directed acyclic graph on [n]
$X_i, i \in [n]$: jointly Gaussian
$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + a_j \epsilon_j$ where the $\epsilon_j \sim N(0,1)$ independent
$\rightsquigarrow \Sigma = (I - \Lambda)^{-T} \mathrm{diag}(a_1^2, \ldots, a_n^2)(I - \Lambda)^{-1}$

II: Cloning sinks in a Gaussian Bayesian model

G: directed acyclic graph on [n]
$X_i, i \in [n]$: jointly Gaussian
$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + a_j \epsilon_j$ where the $\epsilon_j \sim N(0,1)$ independent
$\rightsquigarrow \Sigma = (I - \Lambda)^{-T} \mathrm{diag}(a_1^2, \dots, a_n^2)(I - \Lambda)^{-1}$

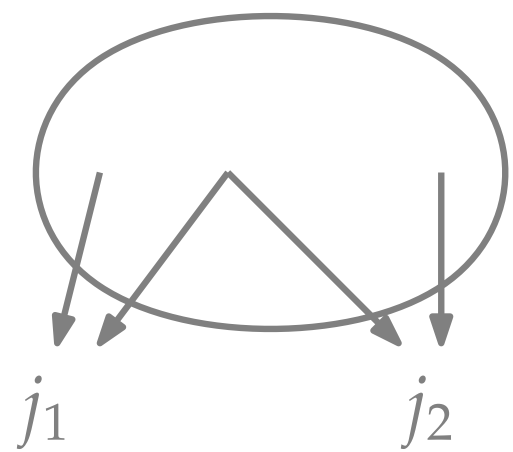
$H \subseteq [m]$ *hidden* $\rightsquigarrow \Sigma_{[n]-H}$ principal submatrix
model: Zariski closure of $\{\Sigma_{[n]-H} | \Lambda, a\}$

II: Cloning sinks in a Gaussian Bayesian model

G: directed acyclic graph on [n]
$X_i, i \in [n]$: jointly Gaussian
$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + a_j \epsilon_j$ where the $\epsilon_j \sim N(0,1)$ independent
$\rightsquigarrow \Sigma = (I - \Lambda)^{-T} \mathrm{diag}(a_1^2, \ldots, a_n^2)(I - \Lambda)^{-1}$

$H \subseteq [m]$ *hidden* $\rightsquigarrow \Sigma_{[n]-H}$ principal submatrix
model: Zariski closure of $\{\Sigma_{[n]-H} | \Lambda, a\}$

Cloning sinks

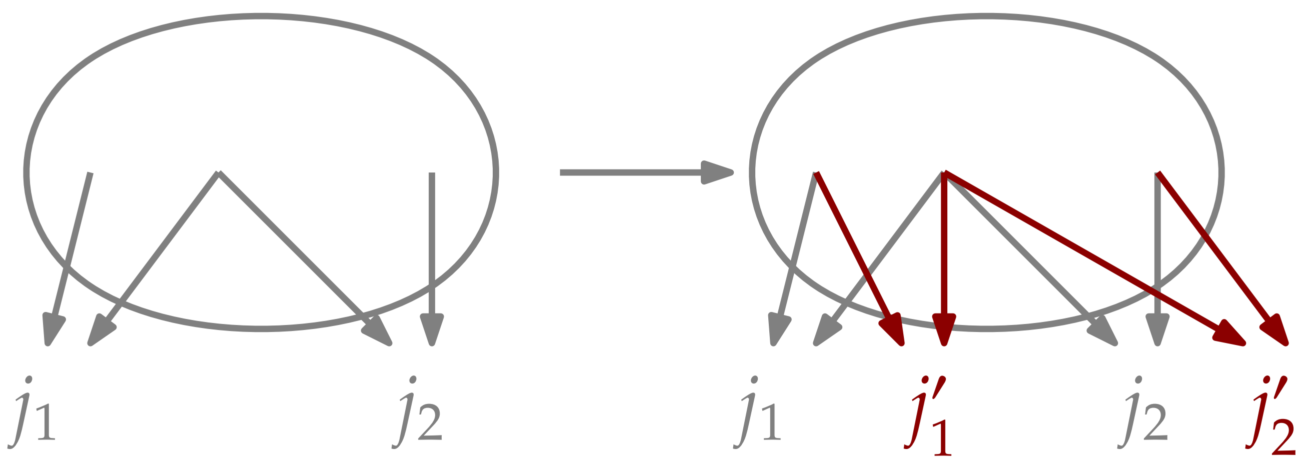


II: Cloning sinks in a Gaussian Bayesian model

G: directed acyclic graph on [n]
$X_i, i \in [n]$: jointly Gaussian
$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + a_j \epsilon_j$ where the $\epsilon_j \sim N(0,1)$ independent
$\rightsquigarrow \Sigma = (I - \Lambda)^{-T} \mathrm{diag}(a_1^2, \ldots, a_n^2)(I - \Lambda)^{-1}$

$H \subseteq [m]$ *hidden* $\rightsquigarrow \Sigma_{[n]-H}$ principal submatrix
model: Zariski closure of $\{\Sigma_{[n]-H} | \Lambda, a\}$

Cloning sinks

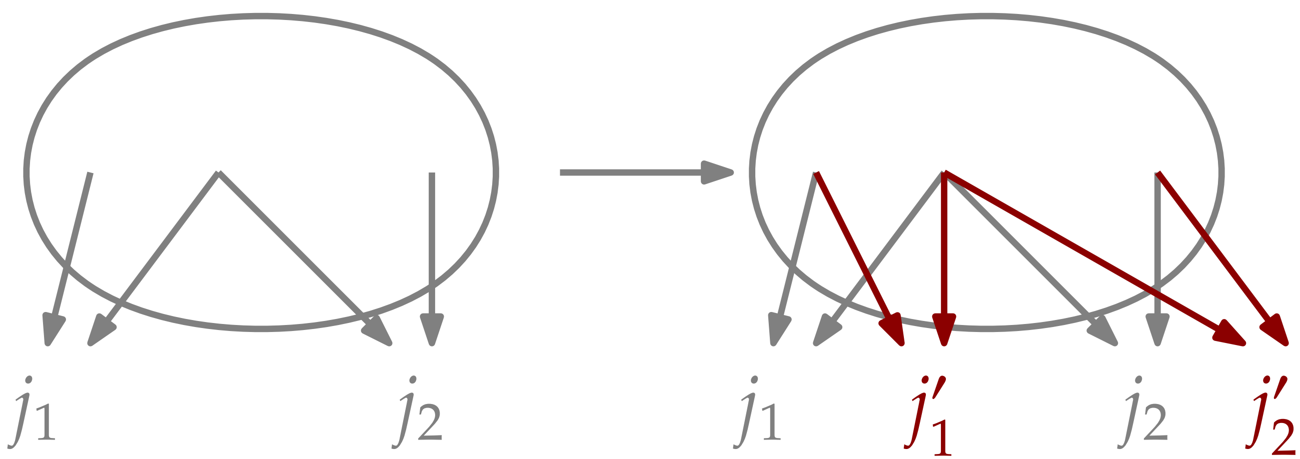


II: Cloning sinks in a Gaussian Bayesian model

G: directed acyclic graph on [n]
$X_i, i \in [n]$: jointly Gaussian
$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + a_j \epsilon_j$ where the $\epsilon_j \sim N(0,1)$ independent
$\rightsquigarrow \Sigma = (I - \Lambda)^{-T} \mathrm{diag}(a_1^2, \ldots, a_n^2)(I - \Lambda)^{-1}$

$H \subseteq [m]$ *hidden* $\rightsquigarrow \Sigma_{[n]-H}$ principal submatrix
model: Zariski closure of $\{\Sigma_{[n]-H} | \Lambda, a\}$

Cloning sinks



Theorem ["D 2010"]
Model stabilises under cloning sinks (via permuting clones).

Stabilisation for parameterised graphical models

undirected	DAG with hidden vars	
$M(G) := \{(p_{i_1,\ldots,i_n} = \prod_C \theta^C_{\mathbf{i}_C})_{i_1,\ldots,i_n}\} \subseteq \mathbb{C}^R$ prod over all cliques C $\theta^C \in \mathbb{C}^{R_C}$ *I: stabilises under increasing r_j for j in independent set*		*discrete* $R = \prod_j [r_j]$
		Gaussian mean 0

Stabilisation for parameterised graphical models

undirected	DAG with hidden vars	
$M(G) := \{(p_{i_1,\ldots,i_n} = \prod_C \theta^C_{\mathbf{i}_C})_{i_1,\ldots,i_n}\} \subseteq \mathbb{C}^R$ prod over all cliques C $\theta^C \in \mathbb{C}^{R_C}$ *I: stabilises under increasing r_j for j in independent set*		*discrete* $R = \prod_j [r_j]$
	$\{\Sigma = (I-\Lambda)^{-T} D (I-\Lambda)^{-1}\}$ $\Lambda_{ij} = 0$ if $i \not\to j$ $M(G) \subseteq \mathbb{C}^{([n]-H)\times([n]-H)}$ *II: stabilises under cloning sinks*	*Gaussian* *mean* 0

Stabilisation for parameterised graphical models

undirected	DAG with hidden vars	
$M(G) := \{(p_{i_1,\ldots,i_n} = \prod_C \theta^C_{\mathbf{i}_C})_{i_1,\ldots,i_n}\} \subseteq \mathbb{C}^R$ prod over all cliques C $\theta^C \in \mathbb{C}^{R_C}$ *I: stabilises under increasing r_j for j in independent set*	$\{(p_{i_1,\ldots,i_n} = \prod_{j\in[n]} \theta_{i_j\|(i_k)_{k\in \mathrm{pa}(j)}})\} \subseteq \mathbb{C}^R$ $\forall \mathbf{i} \in R_{\mathrm{pa}(j)} : \sum_{i_j} \theta_{i_j\|\mathbf{i}} = 1$ hide variables in H *III: stabilises under cloning sinks?? Yes for trees.*	*discrete* $R = \prod_j [r_j]$
	$\{\Sigma = (I-\Lambda)^{-T} D (I-\Lambda)^{-1}\}$ $\Lambda_{ij} = 0$ if $i \not\to j$ $M(G) \subseteq \mathbb{C}^{([n]-H)\times([n]-H)}$ *II: stabilises under cloning sinks*	*Gaussian* *mean* 0

Stabilisation for parameterised graphical models

undirected	DAG with hidden vars	
$M(G) := \{(p_{i_1,\ldots,i_n} = \prod_C \theta^C_{\mathbf{i}_C})_{i_1,\ldots,i_n}\} \subseteq \mathbb{C}^R$ prod over all cliques C $\theta^C \in \mathbb{C}^{R_C}$ *I: stabilises under increasing r_j for j in independent set*	$\{(p_{i_1,\ldots,i_n} = \prod_{j\in[n]} \theta_{i_j\|(i_k)_{k\in \mathrm{pa}(j)}})\} \subseteq \mathbb{C}^R$ $\forall \mathbf{i} \in R_{\mathrm{pa}(j)} : \sum_{i_j} \theta_{i_j\|\mathbf{i}} = 1$ hide variables in H *III: stabilises under cloning sinks?? Yes for trees.*	*discrete* $R = \prod_j [r_j]$
$M(G) := \{\Sigma = K^{-1}\}$ $K_{ij} = 0$ if $ij \notin E(G)$ $M(G) \subseteq \mathbb{C}^{n\times n}$ *stabilisation??*	$\{\Sigma = (I-\Lambda)^{-T} D (I-\Lambda)^{-1}\}$ $\Lambda_{ij} = 0$ if $i \not\to j$ $M(G) \subseteq \mathbb{C}^{([n]-H)\times([n]-H)}$ *II: stabilises under cloning sinks*	*Gaussian mean 0*

Definition

$C \in \mathbb{R}_{\geq 0}^{m \times n} \rightsquigarrow \mathrm{rk}_{\geq 0} C := \min\{r \mid \exists (A, B) \in \mathbb{R}_{\geq 0}^{m \times r} \times \mathbb{R}_{\geq 0}^{r \times n} : C = AB\}$

IV: Nonnegative matrix rank

Definition

$C \in \mathbb{R}_{\geq 0}^{m \times n} \rightsquigarrow \mathrm{rk}_{\geq 0} C := \min\{r \mid \exists (A, B) \in \mathbb{R}_{\geq 0}^{m \times r} \times \mathbb{R}_{\geq 0}^{r \times n} : C = AB\}$

Very ill-behaved [Moitra 2012]

For all $n > 3$ there are $n \times n$-matrices of nonnegative rank > 3 with all proper submatrices of nonnegative rank 3.

Definition
$C \in \mathbb{R}_{\geq 0}^{m \times n} \rightsquigarrow \mathrm{rk}_{\geq 0} C := \min\{r \mid \exists (A, B) \in \mathbb{R}_{\geq 0}^{m \times r} \times \mathbb{R}_{\geq 0}^{r \times n} : C = AB\}$

Very ill-behaved [Moitra 2012]
For all $n > 3$ there are $n \times n$-matrices of nonnegative rank > 3 with all proper submatrices of nonnegative rank 3.

$M_r^{m \times n} := \{A \in \mathbb{R}_{\geq 0}^{m \times n} \mid \mathrm{rk}_{\geq 0} A \leq r\}$
(positive cone over) mixture of r copies of independence
$\partial M_r^{m \times n}$:=topological boundary

IV: Nonnegative matrix rank

Definition

$C \in \mathbb{R}_{\geq 0}^{m\times n} \rightsquigarrow \mathrm{rk}_{\geq 0} C := \min\{r \mid \exists (A,B) \in \mathbb{R}_{\geq 0}^{m\times r} \times \mathbb{R}_{\geq 0}^{r\times n} : C = AB\}$

Very ill-behaved [Moitra 2012]

For all $n > 3$ there are $n \times n$-matrices of nonnegative rank > 3 with all proper submatrices of nonnegative rank 3.

$M_r^{m\times n} := \{A \in \mathbb{R}_{\geq 0}^{m\times n} \mid \mathrm{rk}_{\geq 0} A \leq r\}$

(positive cone over) mixture of r copies of independence

$\partial M_r^{m\times n}$:=topological boundary

Observation/Theorem [Kubjas, Robeva, Sturmfels 2013]

EM-algorithm for $M_r^{m\times n}$ often converges to boundary!

Explicit, quantifier-free expression for $r = 2$.

Algebraic boundary

$\overline{\partial M_r^{m\times n}}$: Zariski closure $\subseteq \mathbb{C}^{m\times n}$
hypersurface in the variety of rank-r matrices

Algebraic boundary

$\overline{\partial M_r^{m\times n}}$: Zariski closure $\subseteq \mathbb{C}^{m\times n}$

hypersurface in the variety of rank-r matrices

Theorem [Kubjas-Robeva-Sturmfels 2013]

Apart from coordinate hyperplanes, for $m, n \geq 4$, $\overline{\partial M_3^{m\times n}}$ has 2 $\mathrm{Sym}(m) \times \mathrm{Sym}(n)$-orbits of irreducible components, parameterised by the following and its transpose:

$$\begin{bmatrix} 0 & * & * \\ * & 0 & * \\ * & * & 0 \\ * & * & * \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & * & * & * \\ * & * & 0 & * & * \\ * & * & * & 0 & * \end{bmatrix}$$

Nonnegative matrix rank, continued

Algebraic boundary

$\overline{\partial M_r^{m\times n}}$: Zariski closure $\subseteq \mathbb{C}^{m\times n}$
hypersurface in the variety of rank-r matrices

Theorem [Kubjas-Robeva-Sturmfels 2013]

Apart from coordinate hyperplanes, for $m, n \geq 4$, $\overline{\partial M_3^{m\times n}}$ has 2 $\mathrm{Sym}(m) \times \mathrm{Sym}(n)$-orbits of irreducible components, parameterised by the following and its transpose:

$$\begin{bmatrix} 0 & * & * \\ * & 0 & * \\ * & * & 0 \\ * & * & * \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & * & * & * \\ * & * & 0 & * & * \\ * & * & * & 0 & * \end{bmatrix}$$

Conjecture

This component has a GB of 4×4 minors plus $\binom{m}{3}$ sextics.

Algebraic boundary

$\overline{\partial M_r^{m\times n}}$: Zariski closure $\subseteq \mathbb{C}^{m\times n}$

hypersurface in the variety of rank-r matrices

Theorem [Kubjas-Robeva-Sturmfels 2013]

Apart from coordinate hyperplanes, for $m, n \geq 4$, $\overline{\partial M_3^{m\times n}}$ has 2 $\mathrm{Sym}(m) \times \mathrm{Sym}(n)$-orbits of irreducible components, parameterised by the following and its transpose:

$$\begin{bmatrix} 0 & * & * \\ * & 0 & * \\ * & * & 0 \\ * & * & * \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & * & * & * \\ * & * & 0 & * & * \\ * & * & * & 0 & * \end{bmatrix}$$

Conjecture

This component has a GB of 4×4 minors plus $\binom{m}{3}$ sextics.

Now a **Theorem** *due to Eggermont-Horobeţ-Kubjas.*
But what about higher nonnegative rank??

Conclusions and questions

- Many algebro-statistical models fit into families with a meaningful limit.
- There is an ever growing body of commutative algebra for dealing with these limits up to symmetry.
- Do discrete Bayesian models stabilise under cloning sinks?
- Do undirected Gaussian graphical models exhibit any kind of stabilisation?
- If you have other families of models where you expect stabilisation, come talk to me!

Conclusions and questions

- Many algebro-statistical models fit into families with a meaningful limit.
- There is an ever growing body of commutative algebra for dealing with these limits up to symmetry.
- Do discrete Bayesian models stabilise under cloning sinks?
- Do undirected Gaussian graphical models exhibit any kind of stabilisation?
- If you have other families of models where you expect stabilisation, come talk to me!

Děkuji!