Finiteness Theorems in Algebraic Statistics

Seth Sullivant

North Carolina State University

November 12, 2008

Notation

- For sets S, T, let $X_{S,T} = (x_{s,t})_{s \in S, t \in T}$ be the matrix of indeterminates $x_{s,t}$.
- Let $\mathbb{K}[X_{S,T}] := \mathbb{K}[x_{s,t} : s \in S, t \in T]$ the polynomial ring with coefficients in $\mathbb{K}$.
- Let $\mathfrak{S}_T$ be the symmetric group on T; the group of all permutations of T.

Example

Main interest is finite by infinite matrices

$$X_{[k],\mathbb{N}} = \begin{pmatrix} x_{10} & x_{11} & x_{12} & x_{13} & \cdots \\ x_{20} & x_{21} & x_{22} & x_{23} & \cdots \\ x_{30} & x_{31} & x_{32} & x_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ x_{k0} & x_{k1} & x_{k2} & x_{k3} & \cdots \end{pmatrix}, \quad \mathbb{K}[X_{[k],\mathbb{N}}], \quad \mathfrak{S}_{\mathbb{N}}$$

Notation

- Let $\mathbb{K}[X_{S,T}][\mathfrak{S}_T]$ be the group ring of $\mathfrak{S}_T$ with coefficients in $\mathbb{K}[X_{S,T}]$.

$$(f\sigma)\cdot(g\tau) = (fg)(\sigma\tau).$$

- Let $\mathfrak{S}_T$ act on $\mathbb{K}[X_{S,T}]$ by permuting columns of $X_{S,T}$:

$$\sigma \cdot x_{s,t} = x_{s,\sigma(t)}.$$

- $\mathbb{K}[X_{S,T}]$ becomes a (left) $\mathbb{K}[X_{S,T}][\mathfrak{S}_T]$-module:

$$(f\sigma)\cdot x_{s,t} = f x_{s,\sigma(t)}$$

Proposition

$\mathbb{K}[X_{S,T}][\mathfrak{S}_T]$-submodules of $\mathbb{K}[X_{S,T}]$ are the $\mathfrak{S}_T$-stable ideals; that is, ideals $I \subseteq \mathbb{K}[X_{S,T}]$ such that:

$$\sigma \cdot f \in I \quad \text{for all } f \in I, \sigma \in \mathfrak{S}_T.$$

The Main Theorem

Theorem (Aschenbrenner-Hillar, Hillar-S)

- $\mathbb{K}[X_{[k],\mathbb{N}}]$ *is a Noetherian* $\mathbb{K}[X_{[k],\mathbb{N}}][\mathfrak{S}_{\mathbb{N}}]$*-module.*
- *Every* $\mathfrak{S}_{\mathbb{N}}$ *stable ideal of* $\mathbb{K}[X_{[k],\mathbb{N}}]$ *has a finite symmetric generating set.*

Corollary (Finite to Infinite Ascending Chain Condition)

For $m \in \mathbb{N}$, *let* $I_m \in \mathbb{K}[X_{[k],[m]}]$ *be* $\mathfrak{S}_{[m]}$*-stable ideals such that for all* $m_1 \leq m_2$

$$\mathbb{K}[X_{[k],[m_2]}][\mathfrak{S}_{[m_2]}] \cdot I_{m_1} \subseteq I_{m_2}.$$

Then there exists an $M \in \mathbb{N}$ *such that* $\mathbb{K}[X_{[k],[m]}][\mathfrak{S}_{[m]}] \cdot I_M = I_m$ *for all* $m > M$. *That is, there is a finite set of polynomials that generate all the* I_m *up to symmetry.*

Outline

Markov Bases

Definition

Let $A : \mathbb{Z}^n \to \mathbb{Z}^d$ be a linear transformation. A Markov Basis for A is a finite subset $\mathcal{B} \subset \ker_{\mathbb{Z}}(A)$ such that for all $u, v \in \mathbb{N}^n$ with $A(u) = A(v)$ there is a sequence $b_1, \ldots, b_L \in \mathcal{B}$ such that

1. $u = v + \sum_{i=1}^{L} b_i$, and
2. $v + \sum_{i=1}^{l} b_i \geq 0$ for $l = 1, \ldots, L$.

Markov bases allow us to take random walks over the set of nonnegative integral points inside of polyhedra.

Example: 2-way tables

Let $A : \mathbb{Z}^{k \times m} \to \mathbb{Z}^{k+m}$ such that

$$\begin{aligned} A(u) &= \left(\sum_{j=1}^{m} u_{1j}, \ldots, \sum_{j=1}^{m} u_{kj}; \sum_{i=1}^{k} u_{i1}, \ldots, \sum_{i=1}^{k} u_{im} \right) \\ &= \text{vector of row and column sums of } u \end{aligned}$$

$\ker_{\mathbb{Z}}(A) = \{u \in \mathbb{Z}^{k \times m} : \text{ row and columns sums of } u \text{ are } 0\}$
Markov basis consists of the $2\binom{k}{2}\binom{m}{2}$ moves like:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

Note: there is only one move in Markov basis up to symmetry of $\mathfrak{S}_k \times \mathfrak{S}_m$.

3-way tables

Let $A : \mathbb{Z}^{k_1 \times k_2 \times m} \longrightarrow \mathbb{Z}^{k_1 \times k_2 + k_1 \times m + k_2 \times m}$ be the linear transformation such that

$$\begin{aligned} A(u) &= \left((\sum_{i_3} u_{i_1 i_2 i_3})_{i_1, i_2}; (\sum_{i_2} u_{i_1 i_2 i_3})_{i_1 i_3}; (\sum_{i_1} u_{i_1 i_2 i_3})_{i_2, i_3} \right) \\ &= \text{all 2-way margins of 3-way table } u \\ &= \text{all “line sums” of } u. \end{aligned}$$

Markov basis depends on k_1, k_2, m, contains moves like:

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

but also non-obvious moves like:

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Aoki and Takemura's Surprising Result

Theorem (Aoki-Takemura)

There is a finite set of moves $\mathcal{B}$ *that, up to symmetry, form a Markov basis for* $k_1 = 3, k_2 = 3$ *and all values of* m *arbitrary.*

Most complicated move has format $3 \times 3 \times 5$:

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Question

- Is there something special about $3 \times 3 \times m$ 3-way tables?
- What about other, higher dimensional, Markov basis problems?

Higher Lawrence Liftings

Theorem (Santos-Sturmfels, Hoşten-S)

Fix $k_1, k_2, \ldots, k_{N-1}$. There exists a finite set of moves that form a Markov basis up to symmetry for $k_1 \times k_2 \times \cdots k_{N-1} \times m$ contingency tables for every value of m.

Proof idea: study Markov bases of higher Lawrence liftings:

$$\Lambda_m(A, B) = \begin{pmatrix} A & 0 & 0 & \cdots & 0 \\ 0 & A & 0 & \cdots & 0 \\ 0 & 0 & A & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A \\ B & B & B & \cdots & B \end{pmatrix}$$

as m varies.

Fundamental Theorem of Markov Bases

Definition

Let $A : \mathbb{Z}^n \to \mathbb{Z}^d$. The toric ideal I_A is the ideal

$$\langle x^u - x^v : u, v \in \mathbb{N}^n, Au = Av \rangle \subset \mathbb{K}[x_1, \dots, x_n].$$

Theorem (Diaconis-Sturmfels)

The set of moves $\mathcal{B} \subseteq \ker_{\mathbb{Z}} A$ is a Markov basis for A if and only if the set of binomials $\{x^{b^+} - x^{b^-} : b \in \mathcal{B}\}$ generates I_A.

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix} \longrightarrow x_{21}x_{33} - x_{23}x_{31}$$

Using the Main Theorem for Markov Bases

Corollary

For $m \in \mathbb{N}$, let $I_m \in \mathbb{K}[X_{[k],[m]}]$ be $\mathfrak{S}_{[m]}$-stable ideals such that for all $m_1 \leq m_2$

$$\mathbb{K}[X_{[k],[m_2]}][\mathfrak{S}_{[m_2]}] \cdot I_{m_1} \subseteq I_{m_2}.$$

Then there exists an $M \in \mathbb{N}$ such that $\mathbb{K}[X_{[k],[m]}][\mathfrak{S}_{[m]}] \cdot I_M = I_m$ for all $m > M$. That is, there is a finite set of polynomials that generate all the I_m up to symmetry.

- Let $k = \prod_{i=1}^{N-1} k_i$.
- Each of the toric ideals I_{A_m} for $k_1 \times \cdots k_{N-1} \times m$ Markov basis problem is an ideal $I_{A_m} \in \mathbb{K}[X_{[k],[m]}]$.
- I_{A_m} is $\mathfrak{S}_m$ invariant (permuting levels)
- $\mathbb{K}[X_{[k],[m_2]}][\mathfrak{S}_{[m_2]}] \cdot I_{A_{m_1}} \subseteq I_{A_{m_2}}$.

Beyond Markov Bases

Theorem

Consider a statistical model $\mathcal{M}_{k,m}$ *for two discrete random variables* X, Y*, with discrete state spaces* $[k]$ *and* $[m]$*. Suppose that the sequences* $\mathcal{M}_{k,m}$*,* $m \in \mathbb{N}$ *satisfies:*

- $\mathcal{M}_{k,m}$ *is* $\mathfrak{S}_{[m]}$*-invariant, and*
- *If* $p \in \mathcal{M}_{k,m_2}$*, then* $P(X, Y | Y \leq m_1) \in \mathcal{M}_{k,m_1}$*.*

Then the vanishing ideals $\mathcal{I}(\mathcal{M}_{k,m})$*,* $m \in \mathbb{N}$ *have a finite symmetric generating set.*

This applies in very general circumstances:

- Directed or undirected graphical models with and without hidden variables.
- Any “reasonable” model for categorical data with unordered categories.

The Action of Two Symmetric Groups

Now let $\mathfrak{S}_S \times \mathfrak{S}_T$ act on $\mathbb{K}[X_{S,T}]$ by permuting rows and columns.

Proposition

$\mathbb{K}[X_{\mathbb{N},\mathbb{N}}]$ *is* not *a Noetherian* $\mathbb{K}[X_{\mathbb{N},\mathbb{N}}][\mathfrak{S}_{\mathbb{N}} \times \mathfrak{S}_{\mathbb{N}}]$*-module.*

Proof.

Consider

$$\begin{aligned} I \quad = \quad & \langle x_{11}x_{12}x_{22}x_{21}, x_{11}x_{12}x_{22}x_{23}x_{33}x_{31}, \\ & \ldots \\ & x_{11}x_{12}x_{22}x_{23}\cdots x_{mm}x_{m1}, \ldots \rangle . \end{aligned}$$

There are infinite sequences of bipartite graphs such that none is isomorphic to an induced subgraph of another. ☐

Proof of the Main Theorem

Theorem (Hillar-S)

- $\mathbb{K}[X_{[k],\mathbb{N}}]$ *is a Noetherian* $\mathbb{K}[X_{[k],\mathbb{N}}][\mathfrak{S}_\mathbb{N}]$*-module.*
- *Every* $\mathfrak{S}_\mathbb{N}$ *stable ideal of* $\mathbb{K}[X_{[k],\mathbb{N}}]$ *has a finite symmetric generating set.*

Proof.

- Pass to the initial ideal.
- Show that symmetric monomial ideals are finitely generated.
- Deduce that there are finite symmetric Gröbner bases.
- Unfortunately...

□

Unfortunately...

Observation

Regardless of the term order $\prec$, the symmetric group action does not preserve leading terms. That is, for $f \in \mathbb{K}[X_{[k],\mathbb{N}}][\mathfrak{S}_{\mathbb{N}}]$ and $g \in \mathbb{K}[X_{[k],\mathbb{N}}]$ it often happens that:

$$\mathrm{in}_{\prec}(f \cdot g) \neq \mathrm{in}_{\prec}(f \cdot \mathrm{in}_{\prec}(g)).$$

Example

Let $f = (12)$, $g = x_{11} + x_{12}$. If $\mathrm{in}_{\prec}(g) = x_{12}$, then

$$\mathrm{in}_{\prec}(f \cdot g) = x_{12} \neq x_{11} = \mathrm{in}_{\prec}(f \cdot \mathrm{in}_{\prec}(g)).$$

Proposition

There is no term order on $\mathbb{K}[X_{[k],\mathbb{N}}]$ such that the $\mathbb{K}[X_{[k],\mathbb{N}}]$ initial ideal of every $\mathfrak{S}_{\mathbb{N}}$ stable ideal is $\mathfrak{S}_{\mathbb{N}}$ stable.

Increasing Functions

Notation

- Let Π be the monoid of increasing functions from $\mathbb{N}$ to $\mathbb{N}$:

$$\Pi = \{\pi : \mathbb{N} \to \mathbb{N} : \pi(i) < \pi(i+1) \text{ for all } i \in \mathbb{N}\}.$$

- Let Π act on $\mathbb{K}[X_{[k],\mathbb{N}}]$ by acting on the second coordinate: $\pi \cdot x_{s,t} = x_{s,\pi(t)}$.

Proposition

- *Every $\mathfrak{S}_{\mathbb{N}}$-stable ideal of $\mathbb{K}[X_{[k],\mathbb{N}}]$ is a Π-stable ideal.*
- *If $\prec_{\mathrm{lex}}$ is the column-wise lexicographic term order with $x_{s_1,t_1} \prec_{\mathrm{lex}} x_{s_2,t_2}$ if $t_1 < t_2$ or $t_1 = t_2$ and $s_1 < s_2$, then if $I \subset \mathbb{K}[X_{[k],\mathbb{N}}]$ is a Π-stable ideal, so is $\mathrm{in}_{\prec_{\mathrm{lex}}}(I)$.*

The Main Lemma

Lemma

- *Every* Π*-stable monomial ideal in* $\mathbb{K}[X_{[k],\mathbb{N}}]$ *is finitely generated as a* $\mathbb{K}[X_{[k],\mathbb{N}}][\Pi]$*-module.*
- *The* Π*-divisibility partial order on the set of monomials in* $\mathbb{K}[X_{[k],\mathbb{N}}]$ *(where* $x^u|_\Pi x^v$ *if there is a* $\pi \in \Pi$ *such that* $\pi \cdot x^u | x^v$*) has no infinite antichains.*

Proof.

- The Π-divisibility partial order is the Higman partial order obtained from the usual divisibility order $\mathbb{K}[x_1, \dots, x_k]$.
- The Higman partial order of a partial order without infinite antichains has no infinite antichains (Higman 1952).

$\square$

Computational Framework: Increasing Functions Only

Observation

Computing “symmetric” Gröbner bases is hard, because a symmetry group does not respect the term order.

Problem

Design an algorithm to compute Gröbner basis for ideals stable with respect to a monoid (like Π) that preserves the term order.

- Checking Π-divisibility is easy to implement.
- Can use this notion even in ideals $I \subseteq \mathbb{K}[X_{[k],[m]}]$ that are column homogeneous, because

$$\mathrm{in}_{\prec}(\pi \cdot f) = \pi \cdot \mathrm{in}_{\prec}(f).$$

A Conjecture for Affine Semigroup Rings

Definition

- An (affine) semigroup ring $\mathbb{K}[Q]$ is a subring of a polynomial ring generated by monomials.
- A semigroup ring $\mathbb{K}[Q] \subseteq \mathbb{K}[X_{[k],\mathbb{N}}]$ is Π-stable if $x^u \in \mathbb{K}[Q] \implies \pi \cdot x^u \in \mathbb{K}[Q]$ for all $\pi \in \Pi$.

Conjecture

If $\mathbb{K}[Q] \subseteq \mathbb{K}[X_{[k],\mathbb{N}}]$ is a Π-stable semigroup ring, then $\mathbb{K}[Q]$ is a Noetherian $\mathbb{K}[Q][\Pi]$-module.

Proposition

Let $L \subseteq M \subseteq N$ by R-modules such that L is finitely generated and N/L is a Noetherian R-module. Then M is a finitely generated R-module.

Divisible Semigroup Rings

Definition

A semigroup ring $\mathbb{K}[Q]$ is divisible if $x^u | x^v$ in $\mathbb{K}[X_{[k],\mathbb{N}}]$ implies that $x^u | x^v$ in $\mathbb{K}[Q]$.

Corollary

If $\mathbb{K}[Q] \subseteq \mathbb{K}[X_{[k],\mathbb{N}}]$ *is a divisible* Π*-invariant semigroup ring, then* $\mathbb{K}[Q]$ *is a Noetherian* $\mathbb{K}[Q][\Pi]$ *module.*

Example (Infinite Dimensional Segre Products)

$\mathbb{K}[Q] = \mathbb{K}[x_{1i}x_{2j} : i, j \in \mathbb{N}]$ is a divisible semigroup ring.
If $J \subseteq \mathbb{K}[Y_{\mathbb{N},\mathbb{N}}]$ is Π-stable (acting on both indices) and

$$J \supseteq \left\langle \det \begin{pmatrix} y_{i_1,j_1} & y_{i_1,j_2} \\ y_{i_2,j_1} & y_{i_2,j_2} \end{pmatrix} : i_1, i_2, j_1, j_2 \in \mathbb{N} \right\rangle$$

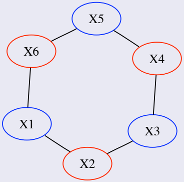
then J is finitely generated as a $\mathbb{K}[Y_{\mathbb{N},\mathbb{N}}][\Pi]$-module.

Independent Set Conjecture For Markov Bases

Theorem (Hillar-S)

Let $\Gamma \subseteq 2^{[N]}$ *be a simplicial complex, and suppose that* $\{\ell+1,\ldots,N\}$ *is an independent set in* Γ*. Fix the number of states of random variables* $X_1, X_2, \ldots, X_\ell$ *as* $k_1, k_2, \ldots, k_\ell$*. Then, up to symmetry, there is a finite set of moves that form a Markov basis for* Γ*, regardless of the number of states of* $X_{\ell+1}, X_{\ell+2}, \ldots, X_N$*.*

Example (6-cycle)



- Fix the number of states of X_1, X_3, X_5, at k_1, k_3, k_5.
- There is a finite symmetric Markov basis of Γ as number of states of $X_2, X_4, X_6 \to \infty$.

Generally: Models Contained in a CI Model

Again: there is nothing special about toric ideals!!!

Theorem (Hillar-S)

Let $\mathcal{M}_{k,m}$ be a family of statistical models for random variables $Y, X_1, \ldots, X_\ell$ that is symmetric with respect to states of the X_i and conditioning invariant. Suppose further that

$$\mathcal{M}_{k,m} \subseteq \mathcal{M}_{X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_\ell | Y}.$$

Then, if the number of states k of Y is fixed, there is a finite, up to symmetry, set of generators of $\mathcal{I}(\mathcal{M}_{k,m})$ for all m.

Open Problems and Future Directions

- Prove the Noetherian conjecture for semigroup rings (and possibly other rings).
- Find effective versions of these results: How large does M need to be in terms of input data?
- Computing monoid invariant Gröbner bases in finite polynomial rings.
- Are there stopping criteria in implicitization computations?