Applied algebraic geometry: algebraic statistics

Jan Draisma, Universität Bern

In more complicated models, do such connected graphs exist?

Definition

Let $A \in \mathbb{Z}_{\geq 0}^{m\times n}$. A *Markov basis* for A is a subset S of $\ker A : \mathbb{Z}^n \to \mathbb{Z}^m$ with the property that if $u, v \in \mathbb{Z}_{\geq 0}^n$ satisfy $Au = Av$, then there exists a sequence $u_0 = u, u_1, \ldots, u_k = v$ in $\mathbb{Z}_{\geq 0}^n$ such that $u_i - u_{i+1} \in \pm S$ for all i.

In our 2×2-table example, $m = n = 4$ and A looks like this:

	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
row 1	1	1		
row 2			1	1
column 1	1		1	
column 2		1		1

Observation
If S is a Markov basis for A, then for every $b \in A\mathbb{Z}^n_{\geq 0}$ the *fibre* $(A^{-1}b) \cap \mathbb{Z}^n_{\geq 0}$ is connected via moves of the form $v \mapsto v + u$ with $u \in \pm S$. If S is finite and $(\ker A) \cap \mathbb{Z}_{\geq 0} = \{0\}^n$, then this fibre is finite for every b, and we have a finite graph as required by M-H.

Theorem (Diaconis-Sturmfels)
For any $A \in \mathbb{Z}^{m\times n}_{\geq 0}$ there exists a finite Markov basis.

Proof
Consider the $\mathbb{Z}$-algebra homomorphism $\varphi : \mathbb{Z}[y_1, \ldots, y_n] \to \mathbb{Z}[x_1, \ldots, x_m]$ that sends y_j to $x^{Ae_j} := \prod_{i=1}^m x_i^{a_{ij}}$.

Note that $\varphi y^u = x^{Au}$ for all $u \in \mathbb{Z}^n_{\geq 0}$.

The existence of finite Markov bases

Suppose that $f = \sum_u c_u y^u \in \ker \varphi$. Then $0 = \sum_u c_u x^{Au}$, and hence for each fixed $b \in A\mathbb{Z}^n_{\geq 0}$ we have $\sum_{u \in (A^{-1}b) \cap \mathbb{Z}^n_{\geq 0}} c_u = 0$. This implies that f lies in the $\mathbb{Z}$-span of all *binomials* $x^u - x^v$ where $u, v \in \mathbb{Z}^n_{\geq 0}$ satisfy $Au = Av$. Conversely, these binomials lie in $\ker \varphi$.

So, by Hilbert's basis theorem, $\ker \varphi$ generated by finitely many binomials $x^{u_i} - x^{v_i}, i = 1, \ldots, k$ where $u_i, v_i \in \mathbb{Z}^n_{\geq 0}$ satisfy $Au_i = Av_i$ and moreover $\operatorname{supp}(u_i) \cap \operatorname{supp}(v_i) = \emptyset$ for all i.

Set $S := \{u_i - v_i \mid i = 1, \ldots, k\}$. We claim that this is a Markov basis. Indeed, suppose that $Au = Av$ where $u, v \in \mathbb{Z}_{\geq 0}$. Then $y^u - y^v \in \ker \varphi$ and hence $y^u - y^v = \sum_{i=1}^k f_i(y^{u_i} - y^{v_i})$ for suitable polynomials f_i.

Rewrite this as $y^u - y^v = \sum_{j=1}^l \pm y^{w_j}(y^{u_{i_j}} - y^{v_{i_j}})$.

Then there is a p such that either $w_p + u_{i_p}$ or $w_p + v_{i_p}$ equals u and has sign $+1$ on the left. Wlog consider the first case.

Then $u' := u - u_{i_p} + v_{i_p} \in u - S \cap \mathbb{Z}^n_{\geq 0}$ is another element in the fibre through u, and $y^{u'} - y^v = \sum_{j \neq p} \pm y^{w_j}(y^{u_{i_j}} - y^{v_{i_j}})$.

Continuing in this fashion we obtain a path from u to v by means of steps from $\pm S$, while keeping all entries nonnegative. □

Exercise
Show that, conversely, if S is a Markov basis of A, then $\ker \varphi$ is generated by all binomials of the form $y^{u_+} - y^{u_-}$ where u_+ is the componentwise maximum of u and 0 and u_- is the componentwise maximum of $-u$ and 0 (so that $u = u_+ - u_-$).

Consequence
The ideal of {rank-one matrices} is generated by 2×2-minors.

Example: no three-way interaction

This is a model for three random variables X_1, X_2, X_3 taking values in $[r_1], [r_2], [r_3]$, respectively.

$p_{ijk} = \text{Prob}(X_1 = i, X_2 = j, X_3 = k)$ equals $a_{jk} b_{ik} c_{ij}$ (normalised such that the probabilities add up to 1)

To use the MH-algorithms for rejecting/accepting that an $r_1 \times r_2 \times r_3$-table of observations comes from this distribution, one needs to sample tables M with prescribed marginals labelled $(j,k), (i,k), (i,j)$ (in total, $r_2 r_3 + r_1 r_3 + r_1 r_2$ marginals). For instance, m_{+jk}.

Take $r_3 = 2$ and $r_1 = r_2 = 5$.

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

This pair represents a possible element of the Markov basis. As $r_3 = 2$ and $r_1, r_2 = n \to \infty$, the maximal degree necessarily grows.

Setting

d random variables taking values in $[r_j]$, $j = 1, \ldots, d$

$\mathcal{F}$ a collection of subsets of $[d]$; for each $A \in \mathcal{F}$ and $\alpha \in \prod_{j \in A}[r_j]$ have a parameter $c_{A,\alpha}$

For $\alpha \in \prod_{j \in [d]}[r_j]$ have $\text{Prob}(\alpha) = \prod_{A \in \mathcal{F}} c_{A,\alpha|_A}$ (forget normalisation).

Example

Independence: $\mathcal{F} = \{\{1\}, \ldots, \{d\}\}$.
No 3-way interaction: $\mathcal{F} = \{\{2,3\},\{1,3\},\{1,2\}\}$

Theorem (Independent set theorem, Hillar-Sullivant)

Fix $\mathcal{F}$ and a subset $T \subseteq [d]$. If $|T \cap A| \leq 1$ for all $A \in \mathcal{F}$, then the Markov degree of the model is bounded as we fix r_j with $j \in [d] \setminus T$ and we let the r_j with $j \in T$ arbitrary elements of $\mathbb{Z}_{\geq 0}$.

Definition
A partial order $\leq$ on S is called a *well-partial order* if for all $s_1, s_2, \ldots \in S$ there exist $i < j$ with $s_i \leq s_j$.

Exercise
If $(S, \leq)$ is wpo then each sequence $s_1, s_2, \ldots$ has an infinite ascending subsequence $s_{i_1} \leq s_{i_2} \leq \cdots$ where $i_1 < i_2 < \ldots$

Lemma
If S, T are wpo, then so is $S \times T$ ordered by $(s, t) \leq (s', t')$ if and only if $s \leq s'$ and $t \leq t'$.

Corollary (Dickson's Lemma)
$\mathbb{Z}^n_{\geq 0}$ with $\alpha \leq \beta$ iff $\beta - \alpha \in \mathbb{Z}^n_{\geq 0}$ is wpo.

Higman's Lemma

If $(S, \leq)$ is wpo, then so is $S^* := \bigcup_n S^n$ with the partial order $(s_1, \ldots, s_m) \leq (t_1, \ldots, t_n)$ if and only if $\exists \pi : [m] \to [n]$ strictly increasing such that $s_i \leq t_{\pi(i)}$ for each $i \in [m]$.

Proof

If not, then there is a counterexample $\mathbf{s}^1, \mathbf{s}^2, \ldots$ where the length of $\mathbf{s}^i$ is minimal among all counterexamples starting with $\mathbf{s}^1, \ldots, \mathbf{s}^{i-1}$.

None of the $\mathbf{s}^i$ is the empty string (); so write $\mathbf{s}^i = (a_i, \mathbf{t}^i)$.

There exists a subsequence $i_1 < i_2 < \ldots$ such that $a_{i_1} \leq a_{i_2} \leq \ldots$ in the wpo S.

Check: then $\mathbf{s}^1, \ldots, \mathbf{s}^{i_1-1}, \mathbf{t}^{i_1}, \mathbf{t}^{i_2}, \ldots$ is a smaller counterexample. □