Normality of the three-state toric homogeneous Markov chain model

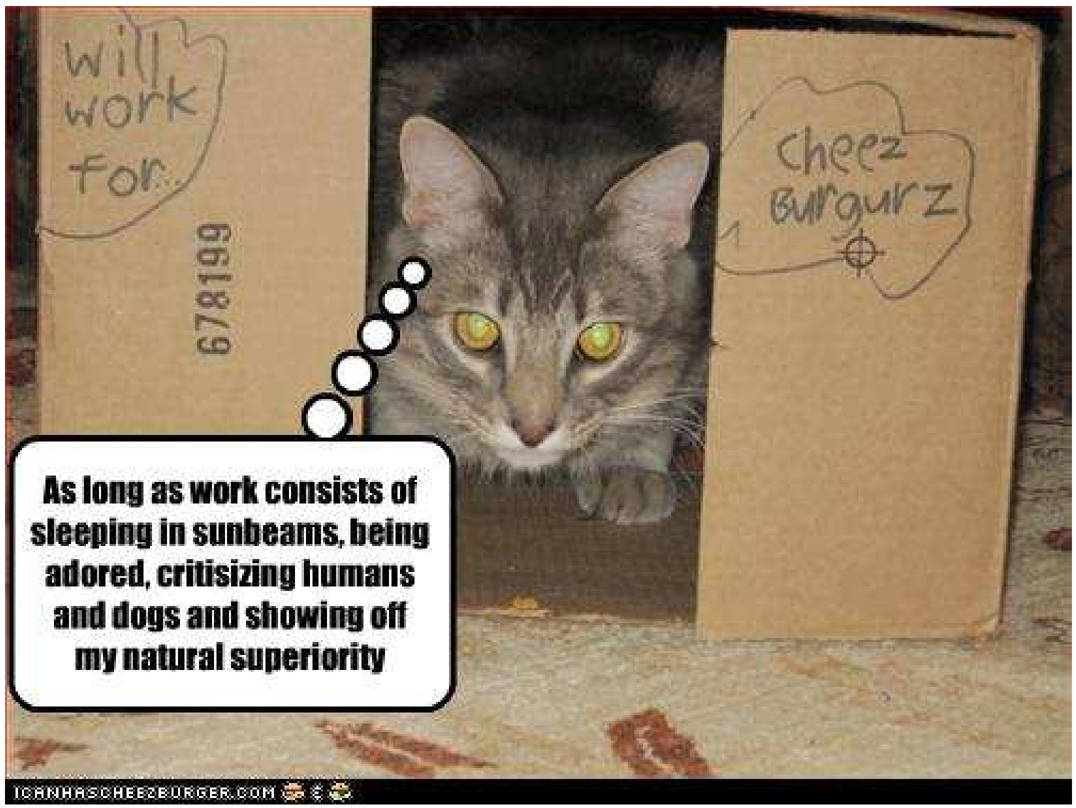
Ruriko Yoshida
Dept. of Statistics University of Kentucky

Joint work with A. Tamekura, D. Haws, and A. Martín del Campo

`polytopes.net`

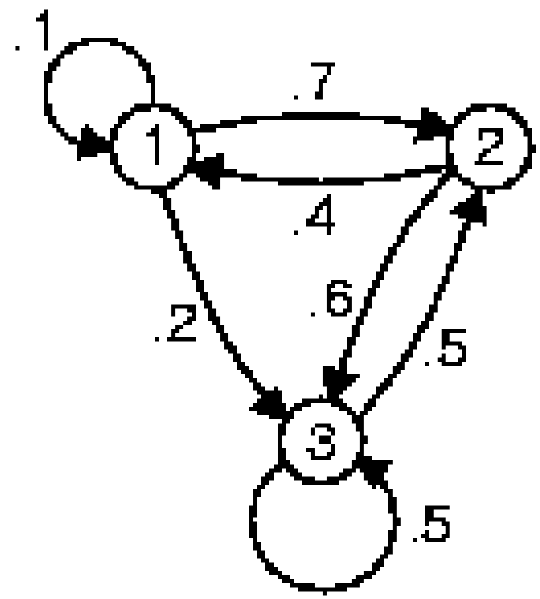
Thank you....

Dave Haws and Abraham Martín del Campo found jobs!



Discrete time Markov chain

We consider a discrete time Markov chain X_t, with $t = 1, \ldots, T$ $(T \geq 3)$, over a finite space of states $[S] = \{1, \ldots, S\}$.



Toric homogeneous Markov chain

Let $\mathbf{w} = (s_1, \ldots, s_T)$ be a path of length T on states $[S]$, which is sometimes written as $\omega = (s_1 \cdots s_T)$ or simply $\omega = s_1 \cdots s_T$. We are interested in Markov bases of toric ideals arising from the following statistical models

$$p(\omega) = c\gamma_{s_1}\beta_{s_1,s_2} \cdots \beta_{s_{T-1},s_T}. \tag{1}$$

where c is a normalizing constant, γ_{s_i} indicates the probability of the initial state, and β_{s_i,s_j} are the transition probabilities from state s_i to s_j. The model (1) is called a toric homogeneous Markov chain (THMC) model.

Problem We want to understand a **Markov basis** under THMC model as $T \to \infty$.

Four models

We refer to them as Model (a), Model (b), Model (c), and Model (d), according to the following:

(a) THMC model (1)

(b) THMC model without initial parameters.

(c) THMC model without self-loops: $\beta_{s_i,s_j} = 0$ whenever $s_i = s_j$.

(d) THMC model without initial parameters and without self-loops, i.e., both (b) and (c) are satisfied

Design matrix for Model (a)

Ordering $[S] \cup [S]^2$ and $[S]^T$ lexicographically, the matrix $A^{(a)}$ is:

	1111	1112	1121	1122	1211	1212	1221	1222	2111	2112	2121	2122	2211	2212	2221	2222
1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
11	3	2	1	1	1	0	0	0	2	1	0	0	1	0	0	0
12	0	1	1	1	1	2	1	1	0	1	1	1	0	1	0	0
21	0	0	1	0	1	1	1	0	1	1	2	1	1	1	1	0
22	0	0	0	1	0	0	1	2	0	0	0	1	1	1	2	3

Design matrix for Model (b)

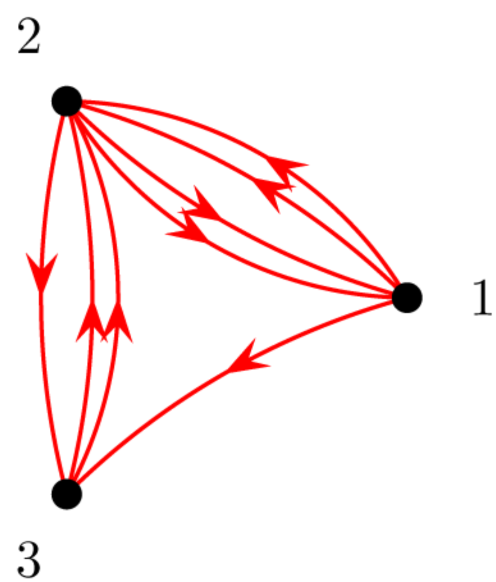
Ordering $[S]^2$ and $[S]^T$ lexicographically with $S = 2$ and $T = 4$ the matrix $A^{(b)}$ is:

	1111	1112	1121	1122	1211	1212	1221	1222	2111	2112	2121	2122	2211	2212	2221	2222
11	3	2	1	1	1	0	0	0	2	1	0	0	1	0	0	0
12	0	1	1	1	1	2	1	1	0	1	1	1	0	1	0	0
21	0	0	1	0	1	1	1	0	1	1	2	1	1	1	1	0
22	0	0	0	1	0	0	1	2	0	0	0	1	1	1	2	3

Example for Model (d)

The state graph $G(W)$ of $W = \{(12132), (12321)\}$. Also the state graph $G(\overline{W})$ where $\overline{W} = \{(13212), (21232)\}$.

Test statistics for both sets of paths is $[2, 1, 2, 1, 0, 2]$.



Two-state THMC

Hara and Takemura (2010) provided a full description of the *Markov bases* for the THMC model (on Model (a) and Model (b)) in two states (i.e. when $S = 2$) that does not depend on T.

Inspired by their work, we study the algebraic and polyhedral properties of the Markov bases of the three-state THMC model for any time $T > 3$.

We hoped we could have the same result for the three-state THMC model without initial parameters and without self-loops (however not yet!).

Recall Markov basis

Suppose $P = \{x \in \mathbb{R}^d | Ax = b,\ x \geq 0\} \neq \emptyset$ and let M be a finite set such that $M \subset \{x \in \mathbb{Z}^d | Ax = 0\}$.

We define the graph G_b such that:

- Nodes of G_b are all the lattice points inside of P.
- We draw an undirected edge between a node u and a node v iff $u - v \in M$.

Definition :

M is called a **Markov basis** if G_b is a connected graph for all b.

Example

				Total
	? ? ?	? ? ?	? ? ?	6
	? ? ?	? ? ?	? ? ?	6
Total	4	4	4	

Table 1: 2×3 tables with 1-marginals.

There are 19 tables with these marginals.

$$\pm \begin{array}{|c|c|c|} \hline 1 & -1 & 0 \\ \hline -1 & 1 & 0 \\ \hline \end{array} \qquad \pm \begin{array}{|c|c|c|} \hline 0 & 1 & -1 \\ \hline 0 & -1 & 1 \\ \hline \end{array}$$

$$\pm \begin{array}{|c|c|c|} \hline 1 & 0 & -1 \\ \hline -1 & 0 & 1 \\ \hline \end{array}$$

There are 3 elements in a Markov basis modulo signs.

4	0	2
0	4	2

+

-1	0	1
1	0	-1

=

3	0	3
1	4	1

A table with the marginals plus an element of a Markov basis is also a table with the given marginals.

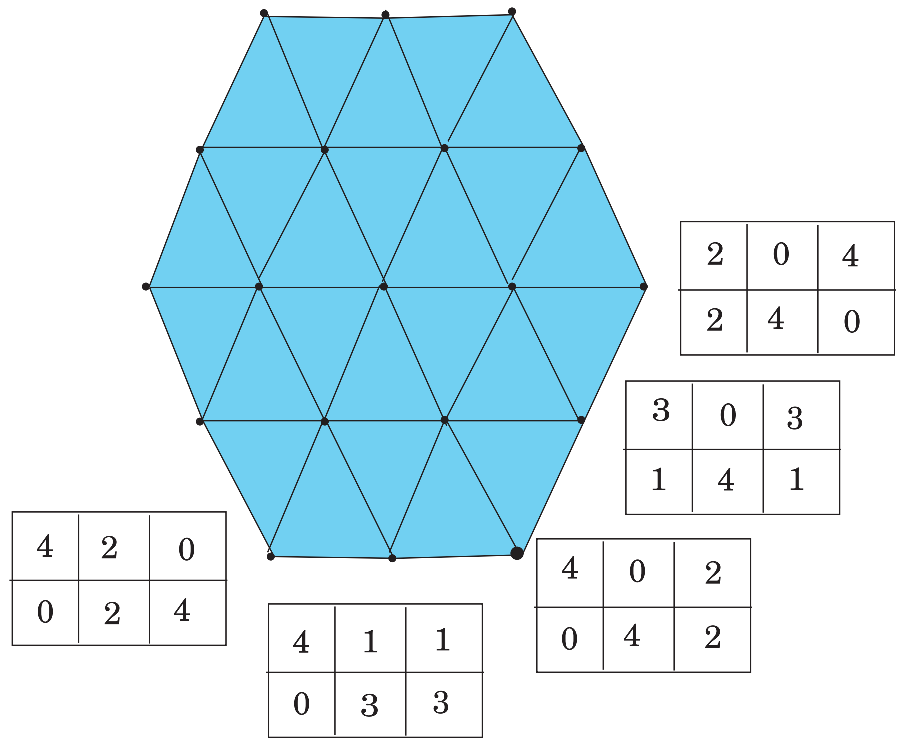


Figure 1: A Markov basis for 2×3 tables. An element of the Markov basis is a undirected edge between integral points in the polytope.

Good news

Theorem: For any $T \geq 4$, a minimum Markov basis for the toric ideal $I_{A^{(d)}}$, where $A^{(d)}$ is the design matrix under Model (d), consists of binomials of degree less than or equal to $d = 6$.

We used polyhedral geometry to prove this theorem.

Polyhedral geometry

Here we focus on Model (d) and $S = 3$.

Look closely at $P^{(d)}$, the convex hull generated by the columns of the design matrix for Model (d).

Recall: The *convex hull* of $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \subset \mathbb{R}^n$ is defined as

$$\text{conv}(\mathbf{a}_1, \ldots, \mathbf{a}_m) := \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{a}_i,\ \sum_{i=1}^m \lambda_i = 1,\ \lambda_i \geq 0 \right\}.$$

A *polytope* P is the convex hull of finitely many points.

For $k \in \mathbb{N}$, we define the k-th dilation of P as $kP := \{\, k\mathbf{x} \mid \mathbf{x} \in P, \}$. A point $\mathbf{x} \in P$ is a *vertex* if and only if it can not be written as a convex combination of points from $P \backslash \{\mathbf{x}\}$.

The *cone* of $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \subset \mathbb{R}^n$ is defined as

$$\text{cone}(\mathbf{a}_1, \ldots, \mathbf{a}_m) := \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{a}_i,\ \lambda_i \geq 0 \right\}.$$

Integer lattice $L := \mathbb{Z}A = \{n_1 a_1 + \cdots + n_m a_m \mid n_i \in \mathbb{Z}\}$.

The semigroup $S := \mathbb{N}A := \{n_1 a_1 + \cdots + n_m a_m \mid n_i \in \mathbb{N}\}$.

Let $P^{(d)}$ be the convex hull generated by the columns of the design matrix for Model (d), let $C^{(d)}$ be the cone generated by the columns of the design matrix for Model (d), let $L^{(d)}$ be the lattice generated by the columns of the design matrix for Model (d), and let $S^{(d)}$ be the semigroup generated by the columns of the design matrix for Model (d).

Prop: $kP^{(d)} = C^{(d)} \cap \{\sum_{i=1}^{n} x_i = k\}$ and $kP^{(d)} \cap \mathbb{Z}^n = C^{(d)} \cap L^{(d)}$.

Note: A semigroup is normal if and only if the semigroup is equal to the intersection between the cone and the lattice.

Theorem: We consider Model (d) and $S = 3$. The semigroup generated by the columns of the design matrix $A^{(d)}$ is normal for $T \geq 5$.

One notices that the set of columns of $A^{(d)}$ is a graded set.

Theorem 13.14 in [Sturmfels 1996] Let $A \subset \mathbb{Z}^d$ be a graded set such that the semigroup generated by the elements in A is normal. Then the toric ideal I_A associate with the set A is generated by homogeneous binomials of degree at most d.

Theorem: For any $T \geq 4$, a minimum Markov basis for the toric ideal $I_{A^{(d)}}$, where $A^{(d)}$ is the design matrix under Model (d), consists of binomials of degree less than or equal to $d = 6$.

Polyhedral geometry

Theorem Let $S = 3$. The number of vertices of $P^{(d)}$ is bounded by some constant C which does not depend on T.

Also we found their hyperplane representations.

Theorem For $T \geq 5$, the number of facets is 24 and we described explicitely the these 24 facet description of $P^{(d)}$ depend on T mod 6.

The number of Hilbert basis elements (`normaliz`) and f-vectors (`Polymake`) for Model (d) where $S = 3$. The running time of `normaliz` was under two seconds for all data sets.

T	#HB	f_0	f_1	f_2	f_3	f_4
4	20	20	69	90	51	12
5	30	27	114	167	102	24
6	48	24	111	176	111	24
7	66	41	144	189	108	24
8	96	42	171	230	123	24
9	123	45	186	245	126	24
10	166	56	201	252	129	24
11	207	63	216	257	126	24
12	264	54	189	236	123	24
13	320	77	246	279	132	24
14	396	54	189	236	123	24
15	468	63	216	257	126	24

Here we summarize all the inequalities in their original form and in their inhomogeneous form, that is using the equality $n(T-1) = x_{12} + x_{13} + x_{21} + x_{23} + x_{31} + x_{32}$ where $n \geq 1$. Index are ordered by lexicographically. Permute on $[S]$.

For any $T \geq 5$, a row vector equivalent to

$$\mathbf{c} = [1, 0, 0, 0, 0, 0] \cdot x \geq 0$$

For any $T \geq 5$, a row vector equivalent to

$$\mathbf{c} = [T, T, -(T-2), 1, -(T-2), 1)] \cdot x \geq 0$$

inhomogeneous

$$\mathbf{c} = [1, 1, -1, 0, -1, 0] \cdot x \geq -n.$$

For any T odd, $T \geq 5$, a row vector equivalent to

$$\mathbf{c} = [1, 1, -1, -1, 1, 1] \cdot x \geq 0.$$

For any $T \geq 4$ of the form $T = 3k + 1$, $k \geq 1$, a row vector equivalent to

$$\mathbf{c} = [2, -1, -1, -1, 2, 2] \cdot x \geq 0.$$

For any $T \geq 5$ of the form $T = 3k + 2$, $k \geq 1$, a row vector equivalent to

$$\mathbf{c} = [2k + 1, -k, -k, -k, 2k + 1, 2k + 1] \cdot x \geq 0$$

inhomogeneous

$$(3 - n)(x_{12} + x_{31} + x_{32}) - n(x_{13} + x_{21} + x_{23}) \geq -n.$$

For any $T \geq 6$, T:even, a row vector equivalent to

$$[\frac{3}{2}T - 1, \frac{T}{2}, -\frac{T}{2} + 1, -\frac{T}{2} + 1, -\frac{T}{2} + 1, \frac{T}{2}] \cdot x \geq 0$$

inhomogeneous

$$3x_{12} + x_{13} - x_{21} - x_{23} - x_{31} + x_{32} \geq -n.$$

For $T = 6k + 3$, a row vector equivalent to

$$[5k + 2, 2k + 1, -4k - 1, -k, -k, 2k + 1] \cdot x \geq 0$$

inhomogeneous

$$(6 - n)x_{12} + (3 - n)x_{13} - (3 + n)x_{21} + (3 - n)x_{32} - nx_{23} - nx_{31} \geq -2n.$$

For $T = 6k$, a row vector equivalent to

$$[10k - 1, 4k, -8k + 2, -2k + 1, -2k + 1, 4k] \cdot x \geq 0$$

inhomogeneous

$$(6 - n)x_{12} + (3 - n)x_{13} - (3 + n)x_{21} + (3 - n)x_{32} - nx_{23} - nx_{31} \geq -2n.$$

Bad news

For Model (a),

Theorem The **semigroup** generated by the columns of the design matrix $A^{(a)}$ is not normal for $S \geq 3$ and $T \geq 4$.

For Model (b),

Theorem The semigroup generated by the columns of the design matrix $A^{(b)}$ is not normal for $S \geq 2$ and $T \geq 3$.

So it is very hard to understand a Markov basis for $T \to \infty$.

For Model (d),

The semigroup generated by the columns of the design matrix $A^{(d)}$ is not normal for $S = 4$ and $T \geq 5$.

Big conjecture

On the experimentations we ran, we found evidence that more should be true.

Conjecture Fix $S \geq 3$; then, for every $T \geq 4$, there is a Markov basis for the toric ideal $I_{A^{(d)}}$ consisting of binomials of degree at most $S-1$, and there is a Gröbner basis with respect to some term ordering consisting of binomials of degree at most S.

Despite the computational limitations (the number of generators grows exponentially when T grows,) we were able to test this conjecture using the software `4ti2` for $T = 4, 5, 6$ with $S = 3$ and $T = 4, 5$ with $S = 4$.

Problem Provide a full description of the *Markov bases* for the THMC model (on Model (d)) in three states (i.e. when $S = 3$) that does not depend on T.

Question??

D Haws, A Martn del Campo, A Takemura, RY

Normality of the three-state toric homogeneous Markov chain model

http://arxiv.org/abs/1204.3070

Thank you!