The Second CREST-SBM International Conference

Harmony of Gröbner Bases and the Modern Industrial Society

Algebra of Reversible Markov Chains

Giovanni Pistone

Collegio Carlo Alberto

Maria Piera Rogantin

Università degli Studi di Genova

giovanni.pistone@gmail.com rogantin@dima.unige.it

Osaka, July 1st 2010

Gröbner bases in discrete stochastics

1. Background on Markov chains:
   - detailed balance,
   - reversibility,
   - Kolmogorov's condition.
2. Main result: Binomial ideal of reversibility.
3. Work in progress
   - other algebraic features,
   - Bayes,
   - from invariant probability to reversible transitions, e.g. Metropolis-Hastings.

Detailed balance

- A transition matrix $P_{v \to w}$, $v, w \in V$, satisfies the detailed balance conditions if $\kappa(v) > 0$, $v \in V$, and

$$\kappa(v) P_{v \to w} = \kappa(w) P_{w \to v}, \quad v, w \in V.$$

- It follows that $\pi(v) \propto \kappa(v)$ is an invariant probability and the Markov chain X_n, $n = 0, 1, \ldots$, has reversible two-step joint distribution

$$\mathrm{P}\left(X_n = v, X_{n+1} = w\right) = \mathrm{P}\left(X_n = w, X_{n+1} = v\right), \quad v, w \in V, n \geq 0.$$

- Reversible MCs are important in Statistical Physics, e.g. for entropy production and in the simulation method Monte Carlo Markov Chain MCMC.
- Textbook on simulation: J.S. Liu, Monte Carlo strategies in scientific computing, Springer Series in Statistics (Springer, New York, 2008), ISBN 978-0-387-76369-9; 0-387-95230-6; chapters on-line `http://www.people.fas.harvard.edu/~junliu/`.
- Original papers on MCMC: W.K. Hastings, Biometrika **57**(1), 97 (1970), `http://dx.doi.org/10.1093/biomet/57.1.97` and P.H. Peskun, Biometrika **60**, 607 (1973), ISSN 0006-3444.
- Kolmogorov's contribution: R.L. Dobrushin, Y.M. Sukhov, Ĭ. Fritts, Uspekhi Mat. Nauk **43**(6(264)), 167 (1988), ISSN 0042-1316, `http://dx.doi.org/10.1070/RM1988v043n06ABEH001985`.
- Textbook on MCs: D.W. Strook, An Introduction to Markov Processes, Number 230 in Graduate Texts in Mathematics (Springer-Verlag, Berlin, 2005), Chapter 5 on MCMC.
- In Statistical Physics: J.L. Lebowitz, H. Spohn, J. Statist. Phys. **95**(1-2), 333 (1999), ISSN 0022-4715, `http://dx.doi.org/10.1023/A:1004589714161`

2-reversible processes

- The stochastic process $(X_n)_{n\geq 0}$ with state space V is 2-reversible if

$$\mathrm{P}\left(X_n = v, X_{n+1} = w\right) = \mathrm{P}\left(X_n = w, X_{n+1} = v\right), \quad v, w \in V, n \geq 0.$$

- The process is 1-stationary:

$$\mathrm{P}\left(X_n = v\right) = \mathrm{P}\left(X_{n+1} = v\right) = \pi(v), \quad v \in V, n \geq 0.$$

- Define $V_2 = \{\{v, w\} : v, w \in V, v \neq w\}$, and

$$\theta_{\{v,w\}} = 2P\left(X_n = v, X_{n+1} = w\right), \quad \{v, w\} \in V_2;$$
$$\theta_v = P\left(X_n = v, X_{n+1} = v\right), \quad v \in V.$$

- We have:

$$1 = \sum_{v,w \in V} P\left(X_n = v, X_{n+1} = w\right) = \sum_{v \in V} \theta_v + \sum_{\{v,w\} \in V_2} \theta_{\{v,w\}},$$

so that $\theta = (\theta_V, \theta_{V_2})$ belongs to the simplex $\Delta(V \cup V_2)$.

- This parameterization is used in P. Diaconis, S.W.W. Rolles, Ann. Statist. **34**(3), 1270 (2006), ISSN 0090-5364, http://dx.doi.org/10.1214/009053606000000290

Restriction on a graph

- We assume we are given the (undirected) connected graph $\mathcal{G} = (V, \mathcal{E})$ and $\theta_{\{v,w\}} = 0$ if $\{v, w\} \notin \mathcal{E}$. The the vector of parameters $\theta = (\theta_v \colon v \in V, \theta_e \colon e \in \mathcal{E})$ belong to the simplex $\Delta(V \cup \mathcal{E})$.
- The probability π is a linear function of the θ parameters:

$$\pi(v) = \sum_{w \in V} P\left(X_n = v, X_{n+1} = w\right) = \theta_v + \frac{1}{2} \sum_{y \colon \{x,y\} \in \mathcal{E}} \theta_{\{v,w\}}$$

or, if where Γ is the incidence matrix of the graph $\mathcal{G}$

$$\pi = \theta_V + \frac{1}{2}\Gamma\theta_{\mathcal{E}}.$$

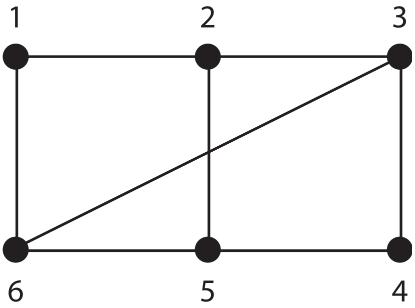
- The map

$$\gamma \colon \Delta(V \cup \mathcal{E}) \ni \theta = \begin{bmatrix} \theta_V \\ \theta_{\mathcal{E}} \end{bmatrix} \longmapsto \pi = \begin{bmatrix} I_V & \frac{1}{2}\Gamma \end{bmatrix} \begin{bmatrix} \theta_V \\ \theta_{\mathcal{E}} \end{bmatrix} \in \Delta(V)$$

is a surjective Markov map.

- The image of $(\theta_V, 0)$, $\theta_V \in \Delta(V)$, is full; the image of $(0, \theta_{\mathcal{E}})$, $\theta_{\mathcal{E}} \in \Delta(\mathcal{E})$, is the convex hull in $\Delta(V)$ of the half points of each edge of the graph $\mathcal{G}$.

Example: 6 vertexes, 8 edges



$\Gamma =$

	{1,2}	{2,3}	{1,6}	{2,5}	{3,4}	{5,6}	{4,5}	{3,6}
1	1	0	1	0	0	0	0	0
2	1	1	0	1	0	0	0	0
3	0	1	0	0	1	0	0	1
4	0	0	0	0	1	0	1	0
5	0	0	0	1	0	1	1	0
6	0	0	1	0	0	1	0	1

Reversible Markov chain

- Assume that the 2-reversible process $(X_n)_{n\in\mathbb{N}}$ is a Markov chain and consider the undirected graph $\mathcal{G} = (V, \mathcal{E})$ such that $\{v, w\} \in \mathcal{E}$ if, and only if, $\theta_{\{v,w\}} > 0$.
- The transition probability are:

$$p_{v\to w} = \frac{\theta_{\{v,w\}}}{\sum_{w:\ \{v,w\}\in\mathcal{E}} \theta_{\{v,w\}}}$$

so that, denoting $\sum_w \theta_{\{x,w\}}$ by $\kappa(v)$, we have the detailed balance conditions

$$\kappa(v)P_{v\to w} = \kappa(w)P_{w\to v}.$$

- Vice-versa, if there exist positive constants $\kappa(v)$, $v \in V$ such that the datailed balance conditions hold, then the process is 2-reversible with $\pi \propto \kappa$.

Reversibility on trajectories

Let $\omega = v_0 \cdots v_n$ be a trajectory (path) in the connected graph $\mathcal{G} = (V, \mathcal{E})$ and let $r\omega = v_n \cdots v_0$ be the reversed trajectory.

Proposition

If the detailed balance holds, the the reversibility condition

$$\mathrm{P}(\omega) = \mathrm{P}(r\omega)$$

holds for each trajectory ω.

Proof.

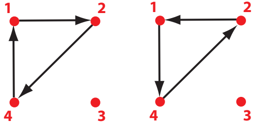
Write the detailed balance along the trajectory,

$$\begin{aligned}
\pi(v_0)P_{v_0 \to v_1} &= \pi(v_1)P_{v_1 \to v_0}, \\
\pi(v_1)P_{v_1 \to v_2} &= \pi(v_2)P_{v_2 \to v_1}, \\
&\vdots \\
\pi(v_{n-1})P_{v_{n-1} \to v_n} &= \pi(v_n)p_{v_n \to v_{n-1}},
\end{aligned}$$

and clear $\pi(v_1) \cdots \pi(v_{n-1})$ in both sides of the product. □

Kolmogorov's condition

We denote by ω a closed trajectory, that is a trajectory on the graph such that the last state coincides with the first one, $\omega = v_0 v_1 \dots v_n v_0$, and by $r\omega$ the reversed trajectory $r\omega = v_0 v_n \dots v_1 v_0$



Theorem (Kolmogorov)

Let the Markov chain $(X_n)_{n\in\mathbb{N}}$ have a transition supported by the connected graph $\mathcal{G}$.

- *If the process is reversible, for all closed trajectory*

$$\boxed{P_{v_0\to v_1}\cdots P_{v_n\to v_0} = P_{v_0\to v_n}\cdots P_{v_1\to v_0}}$$

- *If the equality is true for all closed trajectory, then the process is reversible.*

- The Kolmogorov's condition does not involve the π, whose existence is derived from Doeblin theorem.
- **Detailed balance, reversibility, Kolmogorov's condition are algebraic in nature and define binomial ideals.**

Proof.

- If $\mathrm{P}(\omega) = \mathrm{P}(r\omega)$, then for a closed trajectory we have $\omega = vv_1 \cdots v_{n-1}v$, we have $\mathrm{P}(\omega | X_0 = v) = \mathrm{P}(r\omega | X_n = v)$.
- Vice-versa, assume that all closed trajectory have the displayed property. We denote by x and y the first and the next to last vertices, respectively. By summing on the intermediate vertices on all trajectory with same x and y, we obtain:

$$\sum_{v_2 v_3 \dots v_{n-1}} P_{x \to v_2} P_{v_2 \to v_3} \cdots P_{y \to x} = \sum_{v_2 v_3 \dots v_{n-1}} P_{x \to y} \cdots P_{v_3 \to v_2} P_{v_2 \to x}$$

and

$$P^{(n-2)}_{x \to y} P_{y \to x} = P_{x \to y} P^{(n-2)}_{x \to y}$$

where $P^{(n-2)}_{x \to y}$ denotes the $(n-2)$-step transition probability. If $n \to \infty$, then $P^{(n-2)}_{x \to y} \to \pi(y)$, so that $\pi(y) P_{y \to x} = P_{x \to y} \pi(x)$.

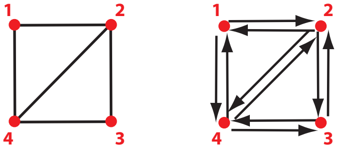
□

Remark

Any algebraic proof?

Transition graph

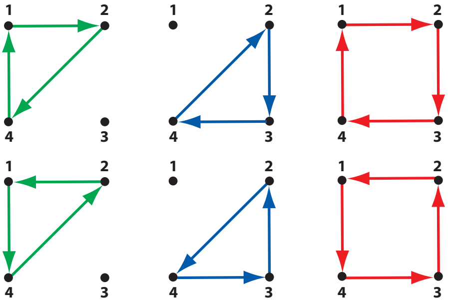
- From $\mathcal{G} = (V, \mathcal{E})$ an (undirected simple) graph, split each edge into two opposite arcs to get a connected directed graph (without loops) $\mathcal{O} = (V, \mathcal{A})$. The arc going from vertex v to vertex w is $(v \to w)$. The reversed arc is $r(v \to w) = (w \to v)$.



- A path or trajectory is a sequence of vertices $\omega = v_0 v_1 \cdots v_n$ with $(v_{k-1} \to v_k) \in \mathcal{A}$, $k = 1, \ldots, n$. The reversed path is $r\omega = v_n v_{n-1} \cdots v_0$. Equivalently, a path is a sequence of inter-connected arcs $\omega = a_1 \ldots a_n$, $a_k = (v_{k-1} \to v_k)$, and $r\omega = r(a_n) \ldots r(a_1)$.

Circuits, cycles

- A closed path $\omega = v_0 v_1 \cdots v_{n-1} v_0$ is any path going from an initial v_0 back to v_0; $r\omega = v_0 v_{n-1} \cdots v_1 v_0$ is the reversed closed path. If we do not distinguish any initial vertex, the equivalence class of closed paths is called a circuit.
- A closed path is elementary if it has no proper closed sub-path, i.e. if does not meet twice the same vertex except the initial one v_0. The circuit of an elementary closed path is a cycle.



- C. Berge, Graphs, Vol. 6 of North-Holland Mathematical Library (North-Holland Publishing Co., Amsterdam, 1985), ISBN 0-444-87603-0, second revised edition of part 1 of the 1973 English version, B. Bollobás, Modern graph theory, Vol. 184 of Graduate Texts in Mathematics (Springer-Verlag, New York, 1998), ISBN 0-387-98488-7.

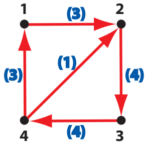
Kolmogorov's ideal

- With indeterminates $P = [P_{v \to w}]$, $(v \to w) \in \mathcal{A}$, form the ring $k[P_{v \to w} : (v \to w) \in \mathcal{A}]$. For a trajectory ω, define the monomial term

$$\omega = a_1 \cdots a_n \mapsto P^\omega = \prod_{k=1}^{n} P_{a_k} = \prod_{a \in \mathcal{A}} P_a^{N_a(\omega)},$$

with $N_a(\omega)$ the number of traversals of the arc a by the trajectory.

2→3 3→4 4→1 1→2
2→3 3→4 4→2
2→3 3→4 4→1 1→2
2→3 3→4 4→1 1→2



$P_{1\to2}^{3} P_{2\to3}^{4} P_{3\to4}^{4} P_{4\to1}^{3} P_{4\to2}$

- $\omega \mapsto P^\omega$ is a representation of the non-commutative path algebra on the commutative product of inderminates. Two closed trajectories associated to the same circuit are mapped to the same monomial term because they have the same traversal counts. The monomial term of a cycle is square-free.

Definition (K-ideal)

The Kolmogorov's ideal or K-ideal of the graph $\mathcal{G}$ is the ideal generated by the binomials $P^\omega - P^{r\omega}$, where ω is any circuit.

Examples

- For a given connected graph $\mathcal{G}$, a transition matrix $P = [P_{v \to w}]$, $u, v \in V$, is compatible with $\mathcal{G}$ if $P_{v \to w} = 0$ whenever $(v \to w) \notin \mathcal{A}$ and $v \neq w$. Let $\text{out}(v)$ be the set of arcs leaving v, and define the simplex

$$\Delta(v) = \left\{ P_{v \to \cdot} \in \mathbb{R}_+^{\text{out}(v)} : \sum_{w \in \text{out}(v)} P_{v \to w}(w) \leq 1 \right\}.$$

- A transition matrix P compatible with $\mathcal{G}$ is a point in the product of simplexes $\Delta(\mathcal{O}) = \times_{u \in V} \Delta(u)$.

Examples of K-ideals

Let P be compatible with $\mathcal{G}$ and reversible.

1. The restriction of a compatible transition matrix $P_{v \to w}$, $(v \to w) \in \mathcal{A}$, is a point of the intersection of the variety of the K-ideal with $\Delta(\mathcal{O})$.
2. Let $(X_n)_{n \geq 0}$ be the stationary Markov chain with reversible transition P. Then the joint probabilities $p(v, w) = \mathrm{P}\left(X_n = u, X_{n+1} = v\right)$, $(v \to w) \in \mathcal{A}$, are points in the intersection of the K-variety and the simplex $\Delta(\mathcal{A}) = \left\{p \in \mathbb{R}_+^{\mathcal{A}} : \sum_{a \in \mathcal{A}} P(a) \leq 1\right\}$.

Basis of the K-ideal

Finite basis of the K-ideal

The K-ideal is generated by the set of binomials $P^{\omega} - P^{r\omega}$, where ω is cycle.

Proof.

Let $\omega = v_0 v_1 \cdots v_0$ be a closed path which is not elementary and consider the least $k \geq 1$ such that $v_k = v_{k'}$ for some $k' < k$. Then the sub-path ω_1 between the k'-th vertex and the k-th vertex is an elementary closed path and the residual path $\omega_2 = v_0 \cdots v_{k'} v_{k+1} \cdots v_0$ is closed and shorter than the original one. The arcs of ω are in 1-to-1 correspondence with the arcs of ω_1 and ω_2. The procedure can be iterated and stops in a finite number of steps. Hence, given any closed path ω, there exists a finite sequence of cycles $\omega_1, \ldots, \omega_l$, such that the list of arcs in ω is partitioned into the lists of arcs of the ω_i's. From $P^{\omega_i} - P^{r\omega_i} = 0$, $i = 1, \ldots, l$, it follows

$$P^{\omega} = \prod_{i=1}^{l} P^{\omega_i} = \prod_{i=1}^{l} P^{r\omega_i} = P^{r\omega}.$$

□

Gröbner basis: recap

- The K-ideal is generated by a finite set of binomials. A Gröbner basis is a special class of generating set of an ideal. We refer to D. Cox, J. Little, D. O'Shea, Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra, Undergraduate Texts in Mathematics, 2nd edn. (Springer-Verlag, New York, 1997), ISBN 0-387-94680-2 and M. Kreuzer, L. Robbiano, Computational commutative algebra. 1 (Springer-Verlag, Berlin, 2000), ISBN 3-540-67733-X for the relevant necessary and sufficient conditions.
- The theory is based on the existence of a monomial order, i.e. a total order on monomial term which is compatible with the product. Given such an order, the leading term $\text{LT}(f)$ of the polynomial f is defined. A generating set is a Gröbner basis if the set of leading terms of the ideal is generated by the leading terms of monomials in the generating set. A Gröbner basis is reduced if the coefficient of the leading term of each element of the basis is 1 and no monomial in any element of the basis is in the ideal generated by the leading terms of the other element of the basis. The Gröbner basis property depend on the monomial order. However, a generating set is a universal Gröbner basis if it is a Gröbner basis for all monomial orders.
- The finite algorithm for computing a Gröbner basis depends on the definition of sygyzy. Given two polynomial f and g in the polynomial ring K, their sygyzy is the polynomial

$$S(f,g) = \frac{\text{LT}(g)}{\gcd(\text{LT}(f),\text{LT}(g))} f - \frac{\text{LT}(f)}{\gcd(\text{LT}(f),\text{LT}(g))} g.$$

  A generating set of an ideal is a Gröbner basis if, and only if, it contains the sygyzy $S(f,g)$ whenever it contains f and g, see Chapter 6 in D. Cox, J. Little, D. O'Shea, Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra, Undergraduate Texts in Mathematics, 2nd edn. (Springer-Verlag, New York, 1997), ISBN 0-387-94680-2 or Theorem 2.4.1 p. 111 of M. Kreuzer, L. Robbiano, Computational commutative algebra. 1 (Springer-Verlag, Berlin, 2000), ISBN 3-540-67733-X.

Universal G-basis of the K-ideal

Universal G-basis

The binomials $P^{\omega} - P^{r\omega}$, where ω is any cycle, form a reduced universal Gröbner basis of the K-ideal.

Proof.

Let ω_1 and ω_2 be two cycles with $\omega_i \succ r\omega_i$, $i = 1, 2$. Assume first they do not have any arc in common. Then $\gcd(P^{\omega_1}, P^{\omega_2}) = 1$ and the sygyzy is

$$S(P^{\omega_1} - P^{r\omega_1}, P^{\omega_2} - P^{r\omega_2}) = P^{\omega_2}(P^{\omega_1} - P^{r\omega_1}) - P^{\omega_1}(P^{\omega_2} - P^{r\omega_2}) = P^{\omega_1}P^{r\omega_2} - P^{r\omega_1}P^{\omega_2},$$

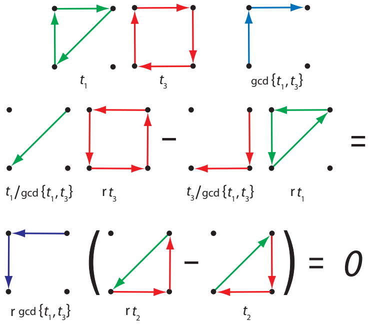
which belongs to the K-ideal.
Let now α be the common part. The sygyzy of $P^{\omega_1} - P^{r\omega_1}$ and $P^{\omega_2} - P^{r\omega_2}$ is

$$P^{\omega_1-\alpha}P^{r\omega_2} - P^{\omega_2-\alpha}P^{r\omega_1} = P^{r\alpha}(P^{\omega_1-\alpha}P^{r\omega_2-r\alpha} - P^{\omega_2-\alpha}P^{r\omega_1-r\alpha}) = 0,$$

which belongs to the K-ideal because $\omega_1 - \alpha + r(\omega_2 - \alpha)$ is a union of cycles. In fact $\omega_1 - \alpha$ and $\omega_2 - \alpha$ have in common the extreme vertices, corresponding to tre extreme vertices of α. Notice that α is the common part of ω_1 and ω_2 only if it is traversed in the same direction by both the cycles. □

Example: square with 1 diagonal

Six cycles: $\omega_1 = 1 \to 2\ 2 \to 4\ 4 \to 1$ (green), $\omega_2 = 2 \to 3\ 3 \to 4\ 4 \to 2$, $\omega_3 = 1 \to 2\ 2 \to 3\ 3 \to 2\ 4 \to 1$ (red), $\omega_4 = r\omega_1$, $\omega_5 = r\omega_2$,$\omega_6 = r\omega_3$.



- ω_1 In blue we have represented the common part of ω_1 and ω_3. $t_i = P^{\omega_i}$, $rt_i = P^{r\omega_i}$, $i = 1, \ldots, 6$.
- A monomial order is obtained by first introducing a total order on arcs. For example, one could give a total order on vertexes, then order lexicographically the arc. We do not see any special order with particular meaning in this problem. The issue is related with the monomial basis which is linear basis of the quotient ring.

Cycle space of $\mathcal{O}$

- For each cycle ω define cycle vector

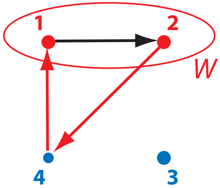
$$z_a(\omega) = \begin{cases} +1 & \text{if } a \text{ is an arc of } \omega, \\ -1 & \text{if } r(a) \text{ is an arc of } \omega, \\ 0 & \text{otherwise.} \end{cases} \qquad a \in \mathcal{A}.$$

- The binomial $P^\omega - P^{r\omega}$ is written as $P^{z^+(\omega)} - P^{z^-(\omega)}$.
- The definition of z can be is extended to any circuit $\bar{\omega}$ by $z_a(\bar{\omega}) = N_a(\omega) - N_a(r\omega)$.
- There exists a sequence of cycles such that $z(\bar{\omega}) = z(\omega_1) + \cdots + z(\omega_l)$.
- We can find nonnegative integers $\lambda(\omega)$ such that $z(\bar{\omega}) = \sum_{\omega \in \mathcal{C}} \lambda(\omega) z(\omega)$, i.e. it belongs to the integer lattice generated by the cycle vectors.
- $Z(\mathcal{O})$ is the cycle space, i.e. the vector space generated in $k^{\mathcal{A}}$ by the cycle vectors.

Cocycle space of $\mathcal{O}$

- For each subset W of V, define cocycle vector

$$u_a(W) = \begin{cases} +1 & \text{if } a \text{ exits from } W, \\ -1 & \text{if } a \text{ enters into } W, \\ 0 & \text{otherwise.} \end{cases} \qquad a \in \mathcal{A}.$$



- The generated subspace of $k^{\mathcal{A}}$ is the cocycle space $U(\mathcal{O})$
- The cycle space and the cocycle space orthogonally split the vector space $\{y \in k^{\mathcal{A}} : y_a = -y_{r(a)}, a \in \mathcal{A}\}$.
- Note that for each cycle vector $z(\omega)$, cocycle vector $u(W)$, $z_a(\omega)u_a(W) = z_{r(a)}(\omega)u_{r(a)}(W)$, $a \in \mathcal{A}$, hence

$$z(\omega) \cdot u(W) = 2\sum_{a\in\omega} u_a(W) = 2\left[\sum_{a\in\omega, u_a(W)=+1} 1 - \sum_{a\in\omega, u_a(W)=-1} 1\right] = 0.$$

- Chapter 2 of C. Berge, Graphs, Vol. 6 of North-Holland Mathematical Library (North-Holland Publishing Co., Amsterdam, 1985), ISBN 0-444-87603-0, second revised edition of part 1 of the 1973 English version; Section II.3 of B. Bollobás, Modern graph theory, Vol. 184 of Graduate Texts in Mathematics (Springer-Verlag, New York, 1998), ISBN 0-387-98488-7.

Toric ideals

- Let U be the matrix whose rows are the cocycle vectors $u(W)$, $W \subset V$. We call the matrix $U = [u_a(W)]_{W \subset V, a \in \mathcal{A}}$ the cocycle matrix.
- Consider the ring $k[P_a \colon a \in \mathcal{A}]$ and the Laurent ring $k(t_W \colon W \subset V)$, together with their homomorphism h defined by

$$h \colon P_a \longmapsto \prod_{W \subset V} t_W^{u_a(W)} = t^{u_a}.$$

- The kernel $I(U)$ of h is the toric ideal of U. It is a prime ideal and the binomials $P^{z^+} - P^{z^-}$, $z \in \mathbb{Z}^{\mathcal{A}}$, $Uz = 0$ are a generating set of $I(U)$ as a k-vector space.
- As for each cycle ω we have $Uz(\omega) = 0$, the cycle vector $z(\omega)$ belongs to $\ker_{\mathbb{Z}} U = \left\{ z \in \mathbb{Z}^{\mathcal{A}} : Uz = 0 \right\}$. Moreover, $P^{z^+(\omega)} = P^{\omega}$, $P^{z^-(\omega)} = P^{r\omega}$, therefore the K-ideal is contained in the toric ideal $I(U)$.
- Chapter 4 B. Sturmfels, Gröbner bases and convex polytopes (American Mathematical Society, Providence, RI, 1996), ISBN 0-8218-0487-1, A. Bigatti, L. Robbiano, Matemática Contemporânea **21**, 1 (2001).

The K-ideal is toric

The K-ideal is the toric ideal of the cocycle matrix.

- Let $\mathcal{C}$ denote the set of cycles and let $z = \sum_{\omega \in \mathcal{C}} \lambda(\omega) z(\omega)$ be a nonzero element of $\ker_{\mathbb{Z}}(U)$.
- For all $\omega \in \mathcal{C}$ we have $-u(\omega) = u(r\omega)$, so that we can assume all the $\lambda(\omega)$'s to be non-negative.
- Notice also that we can arrange things in such a way that at most one of the two direction of each cycle has a positive $\lambda(\omega)$. We define

$$\mathcal{A}_+(z) = \{a \in \mathcal{A} : z_a > 0\}, \quad \mathcal{A}_-(z) = \{a \in \mathcal{A} : z_a < 0\},$$

  and consider two subgraph of $\mathcal{O}$ with a restricted set of arcs, $\mathcal{O}_+(z) = (V, \mathcal{A}_+(z))$, $\mathcal{O}_-(z) = (V, \mathcal{A}_-(z))$. We drop from now on the dependence on z for ease of notation. We note that $r\mathcal{A}_+ = \mathcal{A}_-$ and $r\mathcal{A}_- = \mathcal{A}_+$.

Proof

1. We show first that $\mathcal{A}_+$ must contain a cycle. If $\mathcal{O}_+$ where acyclic, it would exists a vertex v such that $\mathrm{out}(v) \cap \mathcal{A}_+ = \emptyset$ and $\mathrm{in}(v) \cap \mathcal{A}_+ \neq \emptyset$. Let $u(v)$ be the cocycle vector of $\{v\}$; we derive a contradiction to the assumption $z \cdot u(v) = 0$. In fact,

$$z \cdot u(v) = \sum_{a \in \mathcal{A}_+} z_a u_a(v) + \sum_{a \in \mathcal{A}_-} z_a u_a(v)$$
$$= 2 \sum_{a \in \mathcal{A}_+} z_a u_a(v) = 2 \sum_{a \in \mathcal{A}_+ \cap \mathrm{in}(v)} z_a u_a(v) \leq -1.$$

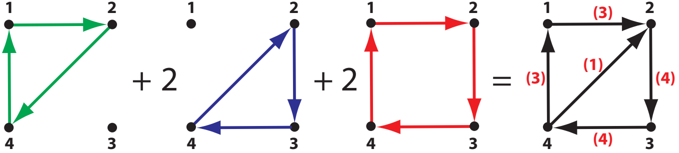
2. Let ω be a cycle in $\mathcal{A}_+$ and define an integer $\alpha(\omega) \geq 1$ such that $z^+ - \alpha(\omega) z^+(\omega) \geq 0$ and it is zero for at least one $a\omega$. The vector $z^1 = z - \alpha(\omega) z(\omega)$ is a cycle vector. i.e. belongs to $\ker_{\mathbb{Z}} U$, and $\mathcal{A}_+(z^1) \subset \mathcal{A}_+(z)$.
3. By repeating the same step a finite number of times we obtain a new representation of z in the form $z = \sum_{\omega \in \mathcal{C}} \alpha(\omega) z(\omega)$ where the support of each $\alpha(\omega) z^+(\omega)$ is contained in $\mathcal{A}_+$. It follows $z^+ = \sum_{\omega \in \mathcal{C}} \alpha(\omega) z^+(\omega)$ and $z^- = \sum_{\omega \in \mathcal{C}} \alpha(\omega) z^-(\omega)$.
4. It follows that

$$P^{z^+} - P^{z^-} = \prod_{\omega \in \mathcal{C}} (P^{z^+(\omega)})^{\alpha(\omega)} - \prod_{\omega \in \mathcal{C}} (P^{z^-(\omega)})^{\alpha(\omega)}$$

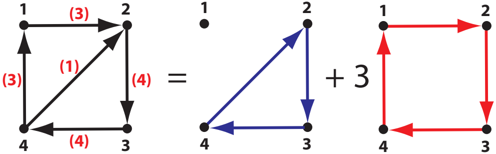
belongs to the K-ideal.

Example of proof

	1 → 2	2 → 1	2 → 3	3 → 2	3 → 4	4 → 3	4 → 1	1 → 4	2 → 4	4 → 2
$z(\omega_A) =$	1	−1	0	0	0	0	1	−1	1	−1
$z(\omega_B) =$	0	0	1	−1	1	−1	0	0	−1	1
$z(\omega_C) =$	1	−1	1	−1	1	−1	1	−1	0	0



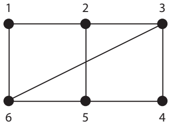
$$z(\omega) = z(\omega_A) + 2z(\omega_B) + 2z(\omega_C) = (3\,,-3\,,4\,,-4\,,4\,,-4\,,0\,,0\,,-1\,,1)$$
$$z^+(\omega) = z^+(\omega_B) + 3z^+(\omega_C) = (3\,,\;0\;,4\,,\;0\;,4\,,\;0\;,0\,,0\,,\;0\;,1)$$



Work in progress

- In the ring $\mathbb{R}[t, P_{v\to w} : (v \to w) \in \mathcal{A}]$, the positive Kolmogorov's ideal is the binomial ideal sum of the Kolmogorov's ideal and $J = \text{Ideal}\left(t \prod_{(v\to w)\in\mathcal{A}} P_{v\to w} - 1\right)$.
- Let $\omega_1, \dots, \omega_m$, be the elementary path obtained from a spanning tree. The sequence $z(\omega_i)$, $i = 1, \dots, m$, is a basis of the elementary closed paths. The positive K-ideal of $\mathcal{O}$ is the sum of J with the ideal generated by $P^{\omega_i} - P^{r\omega_i}$, $i = 1, \dots, m$, i.e. binomials on a basis of the closed paths. In fact, we can generate any elementary closed path and clear the common factors by J.
- The monomial parameterization of the positive K-ideal leads to an alternative presentation of the statistical model.
- The detailed balance ideal is the ideal of $\mathbb{Q}[k(v) : v \in V, P_{v\to w} : (v \to w) \in \mathcal{E}]$ generated by $\prod_{v\in V} k(v) - 1$, $\sum_v P_{v\to w} - 1$, and $k(u)P_{v\to w} - k(v)P_{v\to u}$, $(v \to w) \in \mathcal{E}$.
- If the graph is connected, then the Kolmogorov ideal is the k-elimination ideal of the detailed balance ideal.

CoCoA elimination




```
Use S::=Q[t,k[1..6],p[1..6,1..6]];
Set Indentation;
NI:=6; M:=[];
Define Lista(L,NI);
   For I:=1 To NI Do
       For J:=1 To I-1 Do
           Append(L,k[I]p[I,J]-k[J]p[J,I]); EndFor;
   EndFor; Return L; EndDefine;
N:=Lista(M,NI);
LL:=t*Product([k[I]|I In 1..NI])-1; Append(N,LL);
P0:=[p[1,3],p[1,4],p[1,5],p[2,4],p[2,6], p[3,1],p[3,5],
p[4,1],p[4,2],p[4,6],p[5,1],p[5,3],p[6,2],p[6,4]];
N:=Concat(N,P0);
E:=Elim(k,Ideal(N)); GB:=ReducedGBasis(E);  GB;
```

CoCoA output

```
GB;
[
 p[1,3],  p[1,4],  p[1,5],  p[2,4],  p[2,6],  p[3,1],  p[3,5],
 p[4,1],  p[4,2],  p[4,6],  p[5,1],  p[5,3],  p[6,2],  p[6,4],

  p[2,3]p[3,4]p[4,5]p[5,2] - p[2,5]p[3,2]p[4,3]p[5,4],
  p[1,2]p[2,3]p[3,6]p[6,1] - p[1,6]p[2,1]p[3,2]p[6,3],
  p[1,2]p[2,5]p[5,6]p[6,1] - p[1,6]p[2,1]p[5,2]p[6,5],
  p[2,5]p[3,2]p[5,6]p[6,3] - p[2,3]p[3,6]p[5,2]p[6,5],
  p[3,4]p[4,5]p[5,6]p[6,3] - p[3,6]p[4,3]p[5,4]p[6,5],
  p[1,2]p[2,5]p[3,6]p[4,3]p[5,4]p[6,1] -
        p[1,6]p[2,1]p[3,4]p[4,5]p[5,2]p[6,3],
  p[1,2]p[2,3]p[3,4]p[4,5]p[5,6]p[6,1] -
        p[1,6]p[2,1]p[3,2]p[4,3]p[5,4]p[6,5]]
```

Joint 2-distributions with a given stationary π

- Given π, the fiber $\gamma^{-1}(\pi)$ is contained in an affine space parallel to the subspace $\theta_v + (1/2)\sum_{y\colon \{x,y\}\in\mathcal{E}} \theta_{\{x,y\}} = 0$.
- Each fiber contains special solutions.
  - One is the zero transition case $(\pi, 0_{\mathcal{E}})$.
  - If the graph has full connections, $\mathcal{G} = (V, V_2)$, there is the independence solution $\theta_v = \pi(v)^2$, $\theta_{\{v,w\}} = 2\pi(v)\pi(w)$.
  - If $\pi(v) > 0$, $v \in V$, a strictly positive solution is obtained as follows. Let $d(v) = \#\{w\colon \{v,w\} \in \mathcal{E}\}$ be the degree of the vertex v and define a transition probability by $A_{v\to w} = 1/2d(w)$ if $\{v,w\} \in \mathcal{E}$, $A_{v\to v} = 1/2$, and $A_{v\to w} = 0$ otherwise. A is the transition matrix of a random walk on the graph $\mathcal{G}$, stopped with probability $1/2$. Define a probability on $V \times V$ with $Q(v,w) = \pi(v)A_{v\to w}$. If $Q(v,w) = Q(w,v)$, we have a 2-reversible probability with marginal π. Otherwise, take $Q(v,w) \wedge Q(w,v)$, $\{v,w\} \in \mathcal{E}$.

Metropolis–Hastings algorithm

Proposition

Let Q be a probability on $V \times V$, strictly positive on $\mathcal{E}$, and let $\pi(x) = \sum_y Q(x,y)$. If $f :]0,1[\times]0,1[\to]0,1[$ is a simmetric function such that $f(u,v) \leq u \wedge v$ then

$$P(x,y) = \begin{cases} f(Q(x,y), Q(y,x)) & \{x,y\} \in \mathcal{E} \\ \pi(x) - \sum_{y:\ y \neq x} P(x,y) & x = y \\ 0 & \text{otherwise,} \end{cases}$$

is a 2-reversible probability on $\mathcal{E}$ such that $\pi(x) = \sum_y P(x,y)$, positive if Q is positive.

The proposition applies to

- $f(u,v) = u \wedge v$. This is the Hastings case: $u \wedge v = u(1 \wedge (v/u))$
- $f(u,v) = uv/(u+v)$. This is the Barker case: $uv/(u+v) = u(1 + u/v)^{-1}$
- $f(u,v) = uv$. This is one of the Hastings general form.

Proof.

For $\{x, y\} \in \mathcal{E}$ we have $P(x, y) = P(y, x) > 0$. As $P(x, y) \leq Q(x, y)$, $x \neq y$, it follows

$$\begin{aligned} P(x, x) &= \pi(x) - \sum_{y:\, y \neq x} P(x, y) \\ &\geq \sum_{y} Q(x, y) - \sum_{y:\, y \neq x} Q(x, y) \\ &= Q(x, x) > 0. \end{aligned}$$

We have $\sum_y P(x, y) = \pi(x)$ by construction and, in particular, P is a probability on $V \times V$. □

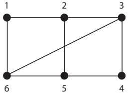
- Given a positive Q, the corresponding parameters for P

$$\theta_{\{x,y\}} = 2P(x, y), \quad \theta_{\{x\}} = P(x, x)$$

  are strictly positive. We have shown the existence of a mapping from the interior of $\Delta(V)$ to the interior of $\Delta(V_1 \cup \mathcal{E})$.

- The mapping $\theta \mapsto (\pi, P_{xy} = \frac{P(x,y)}{\pi(x)})$ is a rational mapping from $\Delta(V_1 \cup V_2)$ into $\Delta(V) \otimes \Delta(V)^{\otimes V}$.

Example 1B



$$Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & \frac{1}{2}\pi(1) & 0 & 0 & 0 & \frac{1}{2}\pi(1) \\ \frac{1}{3}\pi(2) & 0 & \frac{1}{3}\pi(2) & 0 & \frac{1}{3}\pi(2) & 0 \\ 0 & \frac{1}{3}\pi(3) & 0 & \frac{1}{3}\pi(3) & 0 & \frac{1}{3}\pi(3) \\ 0 & 0 & \frac{1}{2}\pi(4) & 0 & \frac{1}{2}\pi(4) & 0 \\ 0 & \frac{1}{3}\pi(5) & 0 & \frac{1}{3}\pi(5) & 0 & \frac{1}{3}\pi(5) \\ \frac{1}{3}\pi(6) & 0 & \frac{1}{3}\pi(6) & 0 & \frac{1}{3}\pi(6) & 0 \end{array} \right]$$

$$P = \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array}$$

$$P = \left[\begin{array}{cccccc} P(11) & \frac{1}{6}\pi(1)\pi(2) & 0 & 0 & 0 & \frac{1}{6}\pi(1)\pi(6) \\ \frac{1}{6}\pi(1)\pi(2) & P(22) & \frac{1}{9}\pi(2)\pi(3) & 0 & \frac{1}{9}\pi(2)\pi(5) & 0 \\ 0 & \frac{1}{9}\pi(2)\pi(3) & P(33) & \frac{1}{6}\pi(3)\pi(4) & 0 & \frac{1}{9}\pi(3)\pi(6) \\ 0 & 0 & \frac{1}{6}\pi(3)\pi(4) & P(44) & \frac{1}{6}\pi(4)\pi(5) & 0 \\ 0 & \frac{1}{9}\pi(2)\pi(5) & 0 & \frac{1}{6}\pi(4)\pi(5) & P(55) & \frac{1}{9}\pi(5)\pi(6) \\ \frac{1}{6}\pi(1)\pi(6) & 0 & \frac{1}{9}\pi(3)\pi(6) & 0 & \frac{1}{9}\pi(5)\pi(6) & P(66) \end{array} \right]$$

Example 1C

$$9\theta_{\mathcal{E}} = \begin{array}{c} \text{edges} \\ \{1,2\} \\ \{2,3\} \\ \{1,6\} \\ \{2,5\} \\ \{3,4\} \\ \{5,6\} \\ \{4,5\} \\ \{3,6\} \end{array} \begin{bmatrix} 3\pi(1)\pi(2) \\ 2\pi(2)\pi(3) \\ 3\pi(1)\pi(6) \\ 2\pi(2)\pi(5) \\ 3\pi(3)\pi(4) \\ 2\pi(5)\pi(6) \\ 3\pi(4)\pi(5) \\ 2\pi(3)\pi(6) \end{bmatrix} \qquad \text{and} \qquad \theta_V = \pi - \frac{1}{2}\Gamma\theta_{\mathcal{E}}$$

$$\boxed{\begin{array}{c} \log \bar{\theta}_{\mathcal{E}} = \text{const} + \Gamma^t\pi \\ \bar{\theta}_V = \pi - \frac{1}{2}\Gamma\bar{\theta}_{\mathcal{E}} \\ \delta_V + \frac{1}{2}\Gamma\delta_{\mathcal{E}} = 0 \end{array}} \Longleftrightarrow \theta = \bar{\theta} + \delta \in \gamma^{-1}(\pi)$$

Example 1D

$$\pi(x) = \text{Binomial}(5,p)(x-1) \implies$$

$$9\theta_{\mathcal{E}} = \begin{array}{c} \text{edges} \\ \{1,2\} \\ \{2,3\} \\ \{1,6\} \\ \{2,5\} \\ \{3,4\} \\ \{5,6\} \\ \{4,5\} \\ \{3,6\} \end{array} \begin{bmatrix} 3\binom{5}{0}p^0(1-p)^5\binom{5}{1}p^1(1-p)^4 \\ 2\binom{5}{1}p^1(1-p)^4\binom{5}{2}p^2(1-p)^3 \\ 3\binom{5}{0}p^0(1-p)^5\binom{5}{5}p^5(1-p)^0 \\ 2\binom{5}{1}p^1(1-p)^4\binom{5}{4}p^4(1-p)^1 \\ 3\binom{5}{2}p^2(1-p)^3\binom{5}{3}p^3(1-p)^2 \\ 2\binom{5}{4}p^4(1-p)^1\binom{5}{5}p^5(1-p)^0 \\ 3\binom{5}{3}p^3(1-p)^2\binom{5}{4}p^4(1-p)^1 \\ 2\binom{5}{2}p^2(1-p)^5\binom{5}{5}p^5(1-p)^0 \end{bmatrix} = \begin{bmatrix} 3\binom{5}{0}\binom{5}{1}p^1(1-p)^9 \\ 2\binom{5}{1}\binom{5}{2}p^3(1-p)^4 \\ 3\binom{5}{0}\binom{5}{5}p^5(1-p)^5 \\ 2\binom{5}{1}\binom{5}{4}p^5(1-p)^5 \\ 3\binom{5}{2}\binom{5}{3}p^5(1-p)^5 \\ 2\binom{5}{4}\binom{5}{5}p^9(1-p)^1 \\ 3\binom{5}{3}\binom{5}{4}p^7(1-p)^3 \\ 2\binom{5}{2}\binom{5}{5}p^7(1-p)^3 \end{bmatrix}$$