WOGAS2 - Workshop on Geometric and Algebraic Statistics 2
CRiSM - Centre for Research in Statistical Methodology
University of Warwick Apr 6-7 2010

Algebra of Reversible Markov Chains

Giovanni Pistone giovanni.pistone@gmail.com
Maria Piera Rogantin rogantin@dima.unige.it

POLITECNICO DI TORINO
Collegio Carlo Alberto
UNIVERSITÀ DEGLI STUDI DI GENOVA

April 6th, 2010

Abstract and a few references

A Markov Chain with stationary un-normalized probability π and transitions P_{xy}, $x, y \in V$, is reversible if the detailed balance conditions

$$\pi(x)P_{xy} = \pi(y)P_{yx}, \quad x \neq y,$$

are satisfied.
In turn, the detailed balance conditions generate a toric ideal of the ring $\mathbb{Q}[\pi(x), P_{xy}]$, so that reversible Markov Chains are part of a toric variety.
Elimination of the indeterminate π produces again a toric ideal, whose binomial generators are the Kolmogorov's equations on the closed paths of the undirected graph of the transition matrix P.
We discuss the parameterization of this model and the effect of imposing a toric model on the π's.

- 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`
- Chapter 5 in D.W. Strook, *An Introduction to Markov Processes*, Number 230 in Graduate Texts in Mathematics (Springer-Verlag, Berlin, 2005)
- 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

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

In particular, the process is 1-stationary:

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

Let $V_1 = \binom{N}{1}$, $V_2 = \binom{V}{2}$. In

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

the following parameterization is considered:

$$\theta_{\{x,y\}} = 2P\left(X_n = x, X_{n+1} = y\right), \quad \{x, y\} \subset V_2;$$
$$\theta_{\{x\}} = P\left(X_n = x, X_{n+1} = x\right), \quad x \in V_1.$$

We have:

$$1 = \sum_{x,y\in V} P\left(X_n = x, X_{n+1} = y\right) = \sum_{\{x\}\in V_1} \theta_{\{x\}} + \sum_{\{x,y\}\in V_2} \theta_{\{x,y\}},$$

so that $\theta = (\theta_{\{x\}} \colon \{x\} \in V_1, \theta_{\{x,y\}} \colon \{x,y\} \in V_2)$ belongs to the simplex $\Delta(V_1 \cup V_2)$.

Restriction on a graph

We consider here a special case. We assume we are given the undirected connected graph $\mathcal{G} = (V, \mathcal{E})$ such that $\theta_{\{x,y\}} = 0$ if $\{x, y\} \notin \mathcal{E}$. The the vector of parameters $\theta = (\theta_{\{x\}} \colon \{x\} \in V_1, \theta_{\{x,y\}} \colon \{x, y\} \in \mathcal{E})$ belong to the simplex $\Delta(V_1 \cup \mathcal{E})$.
The probability $\pi(x)$ can is a linear function of the θ parameters:

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

or, in matrix form,

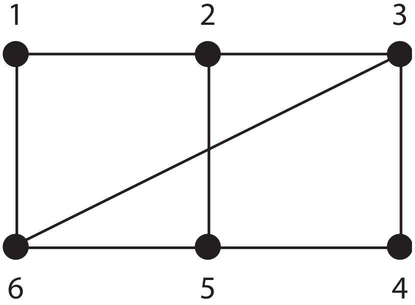
$$\pi = \theta_V + \frac{1}{2}\Gamma\theta_{\mathcal{E}},$$

where Γ is the incidence matrix of the graph $\mathcal{G}$. The map

$$\gamma \colon \Delta(V_1 \cup \mathcal{E}) \ni \theta = \begin{bmatrix} \theta_V \\ \theta_{\mathcal{E}} \end{bmatrix} \longmapsto \pi = \begin{bmatrix} I_{V_1} & \frac{1}{2}\Gamma \end{bmatrix} \begin{bmatrix} \theta_V \\ \theta_{\mathcal{E}} \end{bmatrix} \in \Delta(V_1)$$

is a surjective Markov map. In fact, the image of the convex set $\Delta(V_1 \cup \mathcal{E})$ is the convex hull of the extreme points e_i, $i \in V$, and $e_{\{x,y\}}$, $\{x, y\} \in \mathcal{E}$, hence $\gamma\left(\Delta(V_1 \cup \mathcal{E})\right)$ is the convex hull of the columns of the matrix $\begin{bmatrix} I_{V_1} & \frac{1}{2}\Gamma \end{bmatrix}$. The image of $(\theta_V, 0)$, $\theta_V \in \Delta(V_1)$, 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 1A



$\Gamma =$

	$\{1,2\}$	$\{2,3\}$	$\{1,6\}$	$\{2,5\}$	$\{3,4\}$	$\{5,6\}$	$\{4,5\}$	$\{3,6\}$
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

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 $\theta_{\{x\}} + (1/2) \sum_{y\colon \{x,y\} \in \mathcal{E}} \theta_{\{x,y\}} = 0$.
- 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 $\theta_{\{x\}} = \pi(x)^2$, $\theta_{\{x,y\}} = 2\pi(x)\pi(y)$.
- If $\pi(x) > 0$, $x \in V$, a strictly positive solution is obtained as follows. Let $d(x) = \#\{y\colon \{x,y\} \in \mathcal{E}\}$ be the degree of the vertex x and define a transition probability by $A(x,y) = 1/2d(x)$ if $\{x,y\} \in \mathcal{E}$, $A(x,x) = 1/2$, and $A(x,y) = 0$ 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 $Q(x,y) = \pi(x)A(x,y)$. If $Q(x,y) = Q(y,x)$, we have a 2-reversible probability with marginal π. Otherwise, there is a general construction as follows.

Metropolis

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 standard Metropolis case: $u \wedge v = u(1 \wedge \frac{v}{u})$
- $f(u,v) = uv$. In fact, as $v \leq 1$, we have $uv \leq u$. For our purposes, this case is intersting, because it is an algebraic function.

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

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

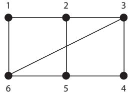
We have $\sum_y P(x, y) = \pi(x)$ by construction and, in particular, P is a probability on $V \times V$. □

- 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

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

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

Example 1D

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

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

Reversible Markov chain

Assume 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 $\{x, y\} \in \mathcal{E}$ if, and only if, $\theta_{\{x,y\}} > 0$. The transition probability are:

$$p_{xy} = \frac{P(X_n = x, X_{n+1} = y)}{P(X_n = x)} = \frac{\theta_{\{x,y\}}}{\sum_{y:\ \{x,y\}\in\mathcal{E}} \theta_{\{x,y\}}}$$
$$p_{yx} = \frac{P(X_n = y, X_{n+1} = x)}{P(X_n = y)} = \frac{\theta_{\{x,y\}}}{\sum_{y:\ \{x,y\}\in\mathcal{E}} \theta_{\{x,y\}}}$$

so that, denoting $\sum_y \theta_{\{x,y\}}$ by $k(x)$, we have the detailed balance conditions

$$k(x)P_{xy} = k(y)P_{yx}, \quad x \neq y.$$

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

Reversibility

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

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

Proposition

If the detailed balance holds with $\prod_v \pi(v) \neq 0$, the reversibility condition

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

holds for all path ω.

Kolmogorov's condition

- We denote now by ω a circuit, that is a path on the graph such that the last vertex coincides with the first one, $\omega = v_0 v_1 \dots v_n v_1$, and by $-\omega$ the reverse circuit $-\omega = v_0 v_n \dots v_1 v_0$.
- A 'basis' of circuits is obtained by considering a spanning tree and adding the remaining vertices one by one.

Theorem

Let then the Markov chain $(X_n)_{n \in \mathbb{N}}$ have support on the connected graph $\mathcal{G}$. If the process is reversible, for all circuit ω

$$P(\omega | X_0 = v_0) = P(-\omega | X_0 = v_0).$$

If the equality is true on a basis of circuits, then the process is reversible.

- A 'basis' of circuits is a basis of the polinomial (binomial) ideal.

Proof.

- If $\mathrm{P}(\omega) = \mathrm{P}(-\omega)$, then for a circuit $\omega = vv_1 \cdots v_{n-1}v$, we have $\mathrm{P}(\omega | X_0 = v) = \mathrm{P}(-\omega | X_n = v)$.
- Vice-versa, assume that all the circuit 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 circuits with same x and y, we obtain:

$$\sum_{v_2 v_3 \dots v_{n-1}} P_{xv_2} P_{v_2 v_3} \cdots P_{yx} = \sum_{v_2 v_3 \dots v_{n-1}} P_{xy} \cdots P_{v_3 v_2} P_{v_2 x}$$

and

$$P_{xy}^{(n-2)} P_{yx} = P_{xy} P_{xy}^{(n-2)}$$

where $P_{xy}^{(n-2)}$ denotes the $(n-2)$-step transition probability. If $n \to \infty$, then $\pi(j) P_{yx} = P_{xy} \pi(x)$, and the chain is reversible.

□

CoCoA: Square

```
Use S::=Q[k[1..4],p[1..4,1..4]];S;
Set Indentation;
NI:=4;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:=Product([k[I]|I In 1..NI])-1;-- non-zero k's
J:=Ideal(p[1,3]^1,p[3,1]^1,p[2,4]^1,p[4,2])^1;
Append(N,LL);N;
I:=Ideal(N);
K:=I+J;
E:=Elim(k,K);E;
```

CoCoA: Square out

```
  Q[k[1..4],p[1..4,1..4]]
-------------------------------
[ -k[1]p[1,2] + k[2]p[2,1],
  -k[1]p[1,3] + k[3]p[3,1],
  -k[2]p[2,3] + k[3]p[3,2],
  -k[1]p[1,4] + k[4]p[4,1],
  -k[2]p[2,4] + k[4]p[4,2],
  -k[3]p[3,4] + k[4]p[4,3],
   k[1]k[2]k[3]k[4] - 1]
-------------------------------
Ideal(
  p[4,2],
  p[2,4],
  p[3,1],
  p[1,3],
  p[1,2]p[2,3]p[3,4]p[4,1] - p[1,4]p[2,1]p[3,2]p[4,3])
-------------------------------
```

CoCoA: Square with diag 13

```
Use S::=Q[k[1..4],p[1..4,1..4]];
Set Indentation;
NI:=4;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:=Product([k[I]|I In 1..NI])-1;-- non zero k's
J:=Ideal(p[2,4]^1,p[4,2]^1); -- diagonal 13
Append(N,LL);
I:=Ideal(N);
K:=I+J;K;
E:=Elim(k,K);E;
```

CoCoA: Square with diag 13 out

```
Q[k[1..4],p[1..4,1..4]]
-------------------------------
Ideal(
  -k[1]p[1,2] + k[2]p[2,1],
  -k[1]p[1,3] + k[3]p[3,1],
  -k[2]p[2,3] + k[3]p[3,2],
  -k[1]p[1,4] + k[4]p[4,1],
  -k[2]p[2,4] + k[4]p[4,2],
  -k[3]p[3,4] + k[4]p[4,3],
  k[1]k[2]k[3]k[4] - 1,
  p[2,4],
  p[4,2])
-------------------------------
Ideal(
  p[4,2],
  p[2,4],
  p[1,3]p[3,4]p[4,1] - p[1,4]p[3,1]p[4,3],
  p[1,2]p[2,3]p[3,1] - p[1,3]p[2,1]p[3,2],
  p[1,2]p[2,3]p[3,4]p[4,1] - p[1,4]p[2,1]p[3,2]p[4,3])
-------------------------------
```