Intertwining of Markov processes and the contact process on the hierarchical group

Jan M. Swart

April 27, 2010

Outline

- Intertwining of Markov processes

Outline

- Intertwining of Markov processes
- First passage times of birth and death processes

Outline

- Intertwining of Markov processes
- First passage times of birth and death processes
- The contact process on the hierarchical group

A bit of filtering theory

Let $(X,Y)=(X_n,Y_n)_{n\geq 0}$ be a Markov chain with finite state space of the form $S\times T$, which jumps from a point (x,y) to a point (x',y') with transition probability $P(x,y;x',y')$. Set

$$\mathbb{P}(y\,|\,x_0,\ldots,x_n):=\mathbb{P}[Y_n=y\,|\,(X_0,\ldots,X_n)=(x_0,\ldots,x_n)].$$

The so-called *filtering equations* read:

$$\begin{aligned}&\mathbb{P}(y\,|\,x_0,\ldots,x_n)\\&\quad=\frac{\sum_{y'\in S'}\mathbb{P}(y'\,|\,x_0,\ldots,x_{n-1})P(x_{n-1},y';x_n,y)}{\sum_{y',y''\in S'}\mathbb{P}(y'\,|\,x_0,\ldots,x_{n-1})P(x_{n-1},y';x_n,y'')}.\end{aligned}$$

Markov functionals

Proposition Let K be a probability kernel from S to T. Assume that there exists a transition probability Q on S such that

$$Q(x,x')K(x',y') = \sum_y K(x,y)P(x,y;x',y')$$

Then

$$\mathbb{P}[Y_0 = y \,|\, X_0] = K(X_0,y) \quad \text{a.s.}$$

implies that

$$\mathbb{P}[Y_n = y \,|\, (X_0,\ldots,X_n)] = K(X_n,y) \quad \text{a.s.} \quad (n \geq 0)$$

and X, on its own, is a Markov chain with transition kernel Q.

Continuous time

[Rogers & Pitman '81] Let K be a probability kernel from S to T and define $K : \mathbb{R}^{S\times T} \to \mathbb{R}^S$ by

$$Kf(x) := \sum_{y\in S'} K(x,y)f(x,y).$$

Let G and $\hat{G}$ be Markov generators on S resp. $S \times T$. Assume that

$$GK = K\hat{G}.$$

Then the Markov process (X,Y) with generator $\hat{G}$ started in an initial law such that

$$\mathbb{P}[Y_0 = y \,|\, X_0] = K(X_0, y) \quad \text{a.s.}$$

satisfies

$$\mathbb{P}[Y_t = y \,|\, (X_s)_{0\leq s\leq t}] = K(X_t, y) \quad \text{a.s.} \qquad (t \geq 0).$$

and X, on its own, is a Markov process with generator G.

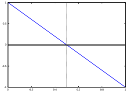
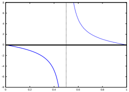
Example: Wright-Fisher diffusion

Let (X, Y) be the Markov process in $[0,1] \times \{0,1\}$ with generator

$$\hat{G}f(x,0) = \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2}f(x,0) + b_0(x)\tfrac{\partial}{\partial x}f(x,0) + \big(f(x,1) - f(x,0)\big),$$
$$\hat{G}f(x,1) = \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2}f(x,1) + b_1(x)\tfrac{\partial}{\partial x}f(x,1),$$

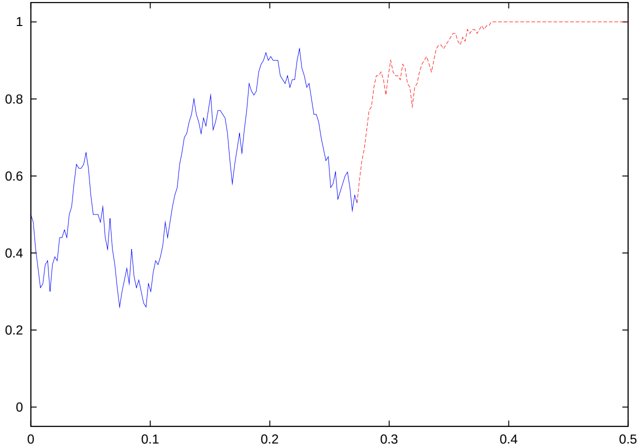
where

$$b_0(x) = 2(\tfrac{1}{2} - x), \quad b_1(x) = \frac{8x(1-x)(x-\frac{1}{2})}{1-4x(1-x)}.$$

b_0  b_1 

Note The process Y is autonomous and jumps $0 \mapsto 1$ with rate one independent of the state of X, but the evolution of X depends on the state of Y.

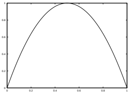
Example: Wright-Fisher diffusion



As long as $Y_t = 0$, X cannot reach the boundary. When $Y_t = 1$, X cannot cross the middle.

Example: Wright-Fisher diffusion

Define a kernel K from $[0,1]$ to $\{0,1\}$ by



$$K(x,0) := 4x(1-x), \quad K(x,1) := 1-4x(1-x)$$

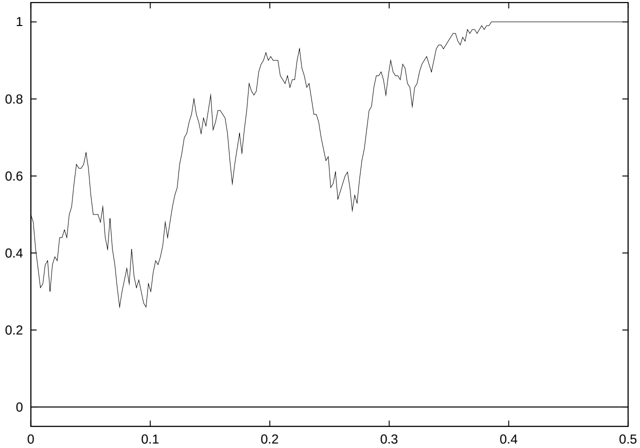
Then $(X_0, Y_0) = (\frac{1}{2}, 0)$ implies

$$\mathbb{P}[Y_t = y \,|\, (X_s)_{0\leq s\leq t}] = K(X_t, y) \quad \text{a.s.} \qquad (t \geq 0).$$

and X, on its own, is a Markov process with generator

$$Gf(x) = \tfrac{1}{2}x(1-x)\tfrac{\partial^2}{\partial x^2}f(x).$$

Example: Wright-Fisher diffusion



Example: Wright-Fisher diffusion

Generalization It is possible to couple a Wright-Fisher diffusion X started in $X_0 = \frac{1}{2}$ to a process Y with state space $\{0, 1, \ldots, \infty\}$, started in $Y_0 = 0$ such that Y jumps $k \mapsto k + 1$ with rate $(k + 1)(2k + 1)$, independent of the state of X, and:

$$\mathbb{P}[Y_t = k \,|\, (X_s)_{0 \leq s \leq t}] = (1 - X_t^2) X_t^{2k},$$

and

$$\inf \big\{ t \geq 0 : X_t \in \{0, 1\} \big\} = \inf\{ t \geq 0 : Y_t = \infty \}.$$

Adding structure to Markov processes

Let X be a Markov processes with state space S and generator G, let K be a probability kernel from S to T and let $(G'_x)_{x\in S}$ be a collection of generators of T-valued Markov processes. Assume that

$$GK = \overline{K}\,\overline{G}$$

where

$$\overline{K}f(x) := \sum_y K(x,y)f(x,y) \quad \text{and} \quad \overline{G}f(x,y) := G'_x f(y).$$

Then X can be coupled to a process Y such that (X,Y) is Markov, Y evolves according to the generator G'_x while X is in the state x, and

$$\mathbb{P}[Y_t = y \,|\, (X_s)_{0\leq s\leq t}] = K(X_t, y) \quad \text{a.s.} \qquad (t \geq 0).$$

We call Y an *added-on process on* X.

Intertwining of semigroups

Proposition Let X and Y be Markov processes with state spaces S and T, semigroups $(P_t)_{t\geq 0}$ and $(P'_t)_{t\geq 0}$, and generators G and G'. Let K be a probability kernel from S to T and define

$$Kf(x) := \sum_y K(x,y)f(y).$$

Then $GK = KG'$ implies the *intertwining relation*

$$P_t K = K P'_t \quad (t \geq 0)$$

and the processes X and Y can be coupled such that

$$\mathbb{P}[Y_t = y \,|\, (X_s)_{0\leq s\leq t}] = K(X_t, y) \quad \text{a.s.} \qquad (t \geq 0).$$

We call Y an *averaged Markov process on* X.

First passage times of birth and death processes

[Karlin & McGregor '59] Let Z be a Markov process with state space $\{0, 1, 2, \ldots\}$, started in $Z_0 = 0$, that jumps $k-1 \mapsto k$ with rate $b_k > 0$ and $k \mapsto k-1$ with rate $d_k > 0$ $(k \geq 1)$. Then

$$\tau_N := \inf\{t \geq 0 : Z_t = N\}$$

is distributed as a sum of independent exponentially distributed random variables whose parameters $\lambda_1 < \cdots < \lambda_N$ are the negatives of the eigenvalues of the generator of the process stopped in N.

A probabilistic proof

[Diaconis & Miclos '09] Let Y be a Markov process with state space $\{0, 1, \ldots, N\}$ that jumps $k-1 \mapsto k$ with rate λ_{N-k+1}. Then Y can be coupled to Z such that

$$\mathbb{P}[Z_{t \wedge \tau_N} = z \,|\, (Y_s)_{0 \leq s \leq t}] = K(Y_t, z) \quad \text{a.s.},$$

where K is a probability kernel on $\{0, \ldots, N\}$ satisfying

$$K(N, N) = 1 \quad \text{and} \quad K(y, \{0, \ldots, y\}) = 1 \quad (0 \leq y < N).$$

Note: here $Z_{t \wedge \tau_N}$ is an averaged Markov process on Y.

A probabilistic proof

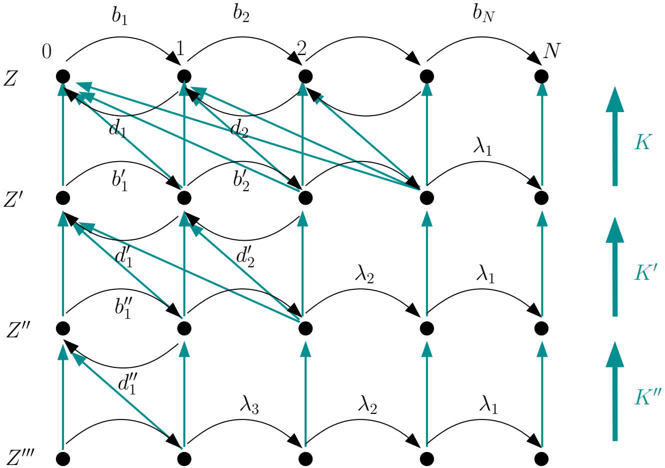
Proof By induction. Assume that Z is a process on $\{0,\ldots,N\}$ such that $b_1,\ldots,b_N>0$, $d_1,\ldots,d_{M-1}>0$, $d_M=\cdots=d_N=0$ for some $1\leq M\leq N$. Then Diaconis and Miclos construct a kernel K such that

$$\begin{array}{ll} K(y,\{0,\ldots,y\})=1 & (0\leq y<M),\\ K(y,y)=1 & (M\leq y\leq N) \end{array}$$

and a proces Z' with $b'_1,\ldots,b'_N>0$, $d'_1,\ldots,d'_{M-2}>0$, $d'_{M-1}=\cdots=d'_N=0$ such that

$$\mathbb{P}[Z_t=z\,|\,(Z'_s)_{0\leq s\leq t}]=K(Z'_t,z)\quad\text{a.s.}$$

A probabilistic proof



A question

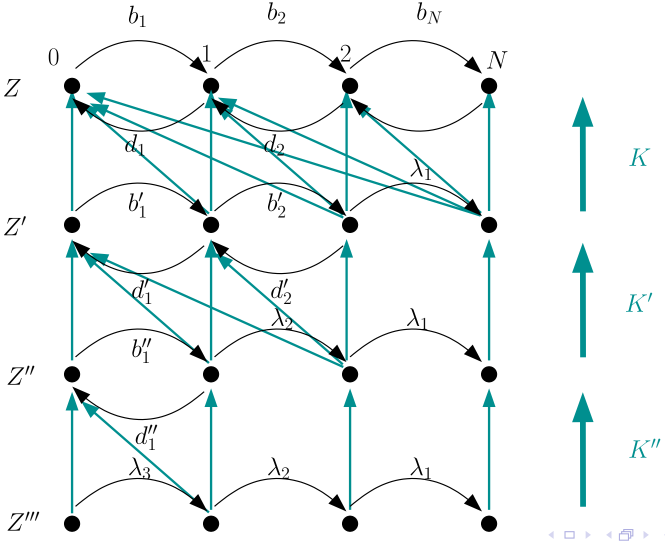
Question Diaconis and Miclos construct $Z_{t\wedge\tau_N}$ as an averaged Markov process on a process Y_t with zero death rates and with birth rates $\lambda_N,\ldots,\lambda_1$, in such a way that $Z_{t\wedge\tau_N} \leq Y_t$. Let V_t be the process with zero death rates and with birth rates $\lambda_1,\ldots,\lambda_N$ (in this order!). Is it possible to contruct V_t as an averaged Markov process on $Z_{t\wedge\tau_N}$ (in this order!) in such a way that $V_t \leq Z_{t\wedge\tau_N}$?

Strong stationary times

Def A *strong stationary time* of a Markov process X is a randomized stopping time τ such that X_τ is independent of τ and X_τ is distributed according to its stationary law.

Claim Let Z be a Markov process with state space $\{0, 1, \ldots, N\}$, started in $Z_0 = 0$, that jumps $k - 1 \mapsto k$ with rate $b_k > 0$ and $k \mapsto k - 1$ with rate $d_k > 0$ $(1 \leq k \leq N)$. Then there exists a strong stationary time for Z and the fastest such time is distributed as a sum of independent exponentially distributed random variables whose parameters $\lambda_1 < \cdots < \lambda_N$ are the negatives of the eigenvalues of the generator of Z.

Strong stationary times

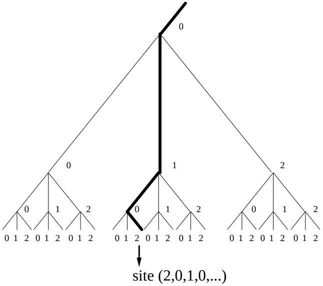


The hierarchical group

By definition, the *hierarchical group with freedom N* is the set

$$\Omega_N := \Big\{ i = (i_0, i_1, \ldots) : \quad \begin{array}{l} i_k \in \{0, \ldots, N-1\}, \\ i_k \neq 0 \text{ for finitely many } k \end{array} \Big\},$$

equipped with componentwise addition modulo N. Think of sites $i \in \Omega_N$ as the leaves of an infinite tree. Then $i_0, i_1, i_2, \ldots$ are the labels of the branches on the unique path from i to the root of the tree.

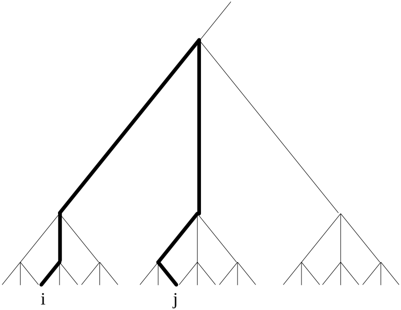


The hierarchical distance

Set

$$|i| := \inf\{k \geq 0 : i_m = 0 \ \forall m \geq k\} \qquad (i \in \Omega_N).$$

Then $|i - j|$ is the *hierarchical distance* between two elements $i, j \in \Omega_N$. In the tree picture, $|i - j|$ measures how high we must go up the tree to find the last common ancestor of i and j.



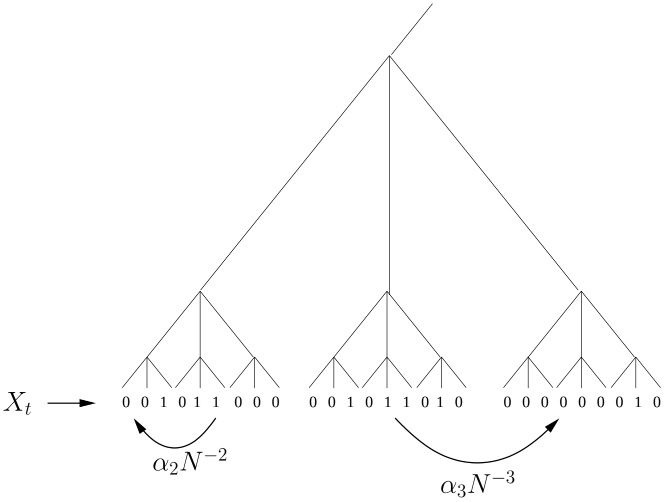
two sites at hierarchical distance 3

Hierarchical contact processes

Fix a *recovery rate* $\delta \geq 0$ and *infection rates* $\alpha_k \geq 0$ such that $\sum_{k=1}^{\infty} \alpha_k < \infty$. The *contact process* on Ω_N with these rates is the $\{0,1\}^{\Omega_N}$-valued Markov process $(X_t)_{t\geq 0}$ with the following description:

If $X_t(i) = 0$ (resp. $X_t(i) = 1$), then we say that the site $i \in \Omega_N$ is *healthy* (resp. *infected*) at time $t \geq 0$. An infected site i infects a healthy site j at hierarchical distance $k := |i - j|$ with rate $\alpha_k N^{-k}$, and infected sites become healthy with rate $\delta \geq 0$.

Hierarchical contact processes



Infection rates on the hierarchical group.

The critical recovery rate

We say that a contact process $(X_t)_{t\geq 0}$ on Ω_N with given recovery and infection rates *survives* if there is a positive probability that the process started with only one infected site never recovers completely, i.e., there are infected sites at any $t \geq 0$. For given infection rates, we let

$$\delta_{\rm c} := \sup \big\{ \delta \geq 0 : \text{the contact process with infection rates } (\alpha_k)_{k\geq 1} \text{ and recovery rate } \delta \text{ survives} \big\}$$

denote the *critical recovery rate*. A simple monotone coupling argument shows that X survives for $\delta < \delta_{\rm c}$ and dies out for $\delta > \delta_{\rm c}$. It is not hard to show that $\delta_{\rm c} < \infty$. The question whether $\delta_{\rm c} > 0$ is more subtle.

(Non)triviality of the critical recovery rate

[Athreya & S. '10] Assume that $\alpha_k = e^{-\theta^k}$ $(k \geq 1)$. Then:

- **(a)** If $N < \theta$, then $\delta_{\rm c} = 0$.
- **(b)** If $1 < \theta < N$, then $\delta_{\rm c} > 0$.

More generally, we show that $\delta_{\rm c} = 0$ if

$$\liminf_{k\to\infty} N^{-k} \log(\beta_k) = -\infty, \quad \text{where} \quad \beta_k := \sum_{n=k}^{\infty} \alpha_n \quad (k \geq 1),$$

while $\delta_{\rm c} > 0$ if

$$\sum_{k=m}^{\infty} (N')^{-k} \log(\alpha_k) > -\infty,$$

for some $m \geq 1$ and $N' < N$.

Proof of survival

We use added-on Markov processes to inductively derive bounds on the finite-time survival probability of finite systems. Let

$$\Omega_2^n := \big\{ i = (i_0, \ldots, i_{n-1}) : i_k \in \{0,1\} \big\}$$

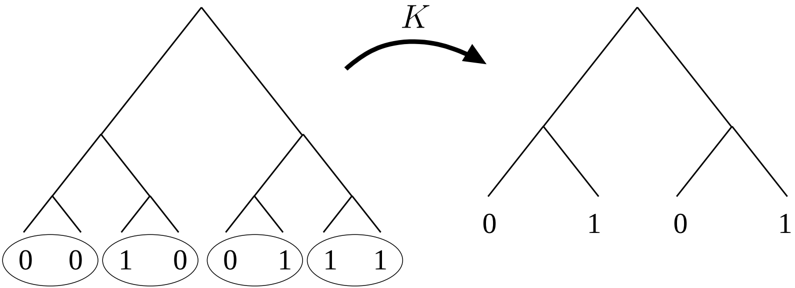
and let $S_n := \{0,1\}^{\Omega_2^n}$. We define a kernel from S_n to S_{n-1} by independently replacing blocks consisting of two spins by a single spin according to the stochastic rules:

$$00 \longrightarrow 0, \qquad 11 \longrightarrow 1,$$

$$\text{and} \quad 01 \text{ or } 10 \longrightarrow \begin{cases} 0 & \text{with probability } \xi, \\ 1 & \text{with probability } 1-\xi, \end{cases}$$

where $\xi \in (0, \frac{1}{2}]$ is a constant, to be determined later.

Renormalization kernel



The probability of this transition is $1 \cdot (1 - \xi) \cdot \xi \cdot 1$.

An added-on process

Let X be a contact process on Ω_2^n with infection rates $\alpha_1, \ldots, \alpha_n$ and recover rate δ. Then X can be coupled to a process Y such that

$$\mathbb{P}[Y_t = y \,|\, (X_s)_{0 \leq s \leq t}] = K(X_t, y) \quad \text{a.s.} \qquad (t \geq 0),$$

where K is the kernel defined before, and

$$\xi := \gamma - \sqrt{\gamma^2 - \tfrac{1}{2}} \quad \text{with} \quad \gamma := \frac{1}{4}\Big(3 + \frac{\alpha_1}{2\delta}\Big).$$

Moreover, the process Y can be coupled to a finite contact process Y' on Ω_2^{n-1} with recovery rate $\delta' := 2\xi\delta$ and infection rates $\alpha'_1, \ldots, \alpha'_{n-1}$ given by $\alpha'_k := \frac{1}{2}\alpha_{k+1}$, in such a way that $Y'_t \leq Y_t$ for all $t \geq 0$.

Renormalization

We may view the map $(\delta, \alpha_1, \ldots, \alpha_n) \mapsto (\delta', \alpha'_1, \ldots, \alpha'_{n-1})$ as an (approximate) renormalization transformation. By iterating this map n times, we get a sequence of recovery rates $\delta, \delta', \delta'', \ldots$, the last of which gives a upper bound on the spectral gap of the finite contact process X on Ω_2^n. Under suitable assumptions on the α_k's, we can show that this spectral gap tends to zero as $n \to \infty$, and in fact, we can derive explicit lower bounds on the probability that finite systems survive till some fixed time t.

A question

Question Can we find an *exact* renormalization map $(\delta, \alpha_1, \ldots, \alpha_n) \mapsto (\delta', \alpha'_1, \ldots, \alpha'_{n-1})$ for hierarchical contact processes, i.e., for each hierarchical contact process X on Ω_2^n, can we find an *averaged Markov process* Y that is itself a hierarchical contact process on Ω_2^{n-1}?