Complements and additional material for “Markov Chains” (2022/23)

Giambattista Giacomin

5045 Bât. Sophie Germain (Campus PRG)
Email address: giacomin@lpsm.paris
URL: https://www.lpsm.paris/users/giacomin/index

1. Markov chains: definitions, basic properties and examples

Like *always* in probability, we work on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that is, at the same time, crucial and useless. At a certain point we will need to make $(\Omega, \mathcal{F}, \mathbb{P})$ explicit, but for most of the time it is just abstract nonsense.

To define a Markov Chain (MC) we need to provide a *state space* E and a probability kernel p:

- the state space E is just a set, but it comes with its own σ-algebra $\mathcal{E}$ that tells us which subsets of E are measurable: so $(E, \mathcal{E})$ is a measurable space;
- p is an application from $E \times \mathcal{E}$ such that
  1. $p(x, \cdot)$ is a probability on $(E, \mathcal{E})$ for every $x \in E$;
  2. $p(\cdot, A) : E \to \mathbb{R}$ is a measurable function for every $A \in \mathbb{E}$ (the measurable subsets of $\mathbb{R}$ are the Borel subsets of $\mathbb{R}$).

A Markov Chain with state space E and probability kernel p (in short: (E, p)-MC, when E is obvious we just write p-MC, sometimes we omit also p) is a sequence $(X_n)_{n=0,1,\ldots}$ of random variables taking values in E with the property that for every n and every $A \in \mathcal{E}$

$$\mathbb{P}\left(X_{n+1} \in A \,\middle|\, \mathcal{F}_n\right) \,=\, p\left(X_n, A\right) \,, \tag{1}$$

where $\mathcal{F}_n = \sigma\left(X_0, X_1, \ldots, X_n\right)$ is the σ-algebra generated by $X_0, X_1, \ldots, X_n$. Of course $\mathcal{F}_n \prec \mathcal{F}$, where $\prec$ means simply that $\mathcal{F}_n \subset \mathcal{F}$, but it reminds you that both $\mathcal{F}_n$ and $\mathcal{F}$ are σ-algebras.

We stress also that the equality in (1) is meant only *almost surely* because a priori the left-hand side is defined only almost surely.

More importantly, note that if we se $\mu(A) := \mathbb{P}(X_0 \in A)$, then μ is a probability on $(E, \mathcal{E})$ and by the tower property of conditional expectation

$$\begin{aligned}\mathbb{P}(X_0 \in A_0, X_1 \in A_1) \,&=\, \mathbb{E}\left[\mathbb{P}\left(X_1 \in A_1 \middle| \mathcal{F}_0\right) \mathbf{1}_{\{X_0 \in A_0\}}\right] \\ &=\, \int_{A_0} \left(\int_{A_1} p(x_0, \mathrm{d}x_1)\right) \mu(\mathrm{d}x_0) \,=\, \int_{A_0}\int_{A_1} \mu(\mathrm{d}x_0) p(x_0, \mathrm{d}x_1),\end{aligned} \tag{2}$$

for every A_0 and $A_1 \in \mathcal{E}$. Of course this generalizes to

$$\mathbb{P}(X_0 \in A_0, X_1 \in A_1, \ldots, X_n \in A_n) \,=\, \int_{A_0}\int_{A_1} \ldots \int_{A_n} \mu(\mathrm{d}x_0) p(x_0, \mathrm{d}x_1) p(x_1, \mathrm{d}x_2) \ldots p(x_{n-1}, \mathrm{d}x_n), \tag{3}$$

for every $A_0, \ldots, A_n \in \mathcal{E}$.

PROPOSITION 1.1. *X est une p-MC avec $X_0 \sim \mu$ (i.e., la loi de X_0 est μ) si et seulement si* (3) *holds for every $n = 0, 1, \ldots$ and for every $A_0, \ldots, A_n \in \mathcal{E}$.*

Before giving examples of MC let us give the following result, which is central for us.

Proposition 1.2. *$(\xi_j)_{j=1,2,\ldots}$ is an IID sequence of random variables that take values in a measurable space $(E', \mathcal{E}')$ and if $h: E \times E' \to E$ is measurable. If X_0 is independent of (ξ_j) and if we set recursively $X_{n+1} = h(X_n, \xi_{n+1})$, $n = 0, 1, \ldots$, we have that (X_n) is a (E, p)-MC with*

$$p(x, A) \,=\, \mathbb{P}\left(h(x, \xi_1) \in A\right)\,, \tag{4}$$

for every $x \in E$ and every $A \in \mathcal{E}$.

Essentially without loss of generality we can choose $E' = \mathbb{R}$ or $E' = (0,1)$, but sometimes if is practical to deal with more general spaces (and we will see it with the first examples). Moreover if we introduce the notation $h_\xi(x) = h(x, \xi)$ we have the convenient notation

$$X_n \,=\, h_{\xi_n} \circ h_{\xi_{n-1}} \circ \ldots \circ h_{\xi_1}\left(X_0\right)\,, \tag{5}$$

so that X_n is just the result of applying n random functions to the initial condition X_0. And (5) is one of the most efficient ways to simulate a Markov chain.

Proof. From (5) we see that X_n is measurable with respect to $\sigma\left(X_0, \xi_1, \ldots, \xi_n\right)$. This implies both that $\mathcal{F}_n \prec \sigma\left(X_0, \xi_1, \ldots, \xi_n\right)$ and that X_n and ξ_{n+1} are independent. And of course ξ_{n+1} is independent of $\sigma\left(X_0, \xi_1, \ldots, \xi_n\right)$. Therefore

$$\mathbb{E}\left[\mathbf{1}_{\{h(X_n, \xi_{n+1}) \in A\}} \,\middle|\, X_0, \xi_1, \ldots, \xi_n\right] \,=\, p\left(X_n, A\right)\,. \tag{6}$$

and

$$\begin{aligned}
\mathbb{P}\left(X_n \in A \,\middle|\, \mathcal{F}_n\right) \,&=\, \mathbb{E}\left[\mathbf{1}_{\{h(X_n, \xi_{n+1}) \in A\}} \,\middle|\, \mathcal{F}_n\right] \\
&=\, \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{h(X_n, \xi_{n+1}) \in A\}} \,\middle|\, X_0, \xi_1, \ldots, \xi_n\right] \,\middle|\, \mathcal{F}_n\right] \\
&=\, p\left(X_n, A\right)\,.
\end{aligned} \tag{7}$$

□

In order to give the first examples of Markov chains let us consider for the moment just the class of random walks: this is a very limited context, but it contains already a lot of examples.

(1) If $E = E' = \mathbb{R}^d$ and $h(x, y) = x + y$ then the arising MC is just a random walk on $\mathbb{R}^d$: for $n = 1, 2, \ldots$

$$X_n \,=\, X_0 + \sum_{j=1}^{n} \xi_j\,, \tag{8}$$

and $p(x, \cdot)$ coincides with the law of ξ_1 for every x.

(2) If $E = E' = \mathbb{Z}^d$ and $h(x, y) = x + y$ then the arising MC is just a random walk on $\mathbb{Z}^d$: note that (8) still holds

(3) If If $E = E' = \mathbb{Z}$, $d = 1$ and $\mathbb{P}(\xi_1 = +1) = 1 - \mathbb{P}(\xi_1 = -1) = p$ then X_n is just a one dimensional simple random walk (simple refers that it jumps just to nearest neighbors). If $p = 1/2$ we speak of *simple symmetric random walk*.

(4) Random walks are naturally defined for example on a graph (N, L) where N is (finite or countably infinite) ensemble and L is a subset of $\{(x,y) \in \mathrm{N}^2 : x, y \in (x,y)\}$. N is the set of *nodes* (or *sites*) and L is the sent of *links*. We say that the graph is symmetric is $x, y) \in \mathrm{L}$ implies $(y,x) \in \mathrm{L}$. If $n_x := |\{y : (x,y) \in \mathrm{L}\ \}| < \infty$ for every $x \in \mathrm{N}$, so we can write $\{y : (x,y) \in \mathrm{L}\ \} = \{y_{x,1}, y_{x,2}, \ldots, y_{x,n_x}\}$, we define for $u \in (0,1)$

$$h(x,u) \,=\, \sum_{j=1}^{n_x} y_{x,j} \mathbf{1}_{((j-1)/n_x, j/n_x]}(u)\,. \tag{9}$$

This way if (ξ_j) is an IID sequence of variables that are uniformly distributed over $(0,1)$ (notation: $U(0,1)$), then, given X_0, $X_{n+1} = h(X_n, U_{j+1})$ defines the *simple random walk* on the graph (N, L). In particular, if $\mathrm{N} = \mathbb{Z}$ and $\mathrm{L} = \{(x,y) \in \mathbb{Z}^2 : |x-y| = 1\}$, then (N, L) is a symmetric graph and the MC we have just defined is the simple symmetric random on $\mathbb{Z}$.

2. Branching process (Bienaymé-Galton-Watson process)

The BGW process (Z_n) is a MC on $E := \mathbb{N} \cup \{0\}$ defined starting from the IID family $\xi := (\xi_{n,j})_{(n,j)\in\mathbb{N}^2}$, with $\mathbb{P}(\xi_{1,1} \in E) = 1$. We use the notation $p_j := \mathbb{P}(\xi_{1,1} = j)$. The chain can be introduced by iteration once Z_0 independent of ξ is given (unless otherwise said, we choose $Z_0 = 1$) via

$$Z_{n+1} \,=\, \begin{cases} \xi_{n+1,1} + \xi_{n+1,2} + \ldots + \xi_{n+1,Z_n} & \text{if } Z_n > 0, \\ 0 & \text{if } Z_n = 0. \end{cases} \tag{10}$$

We assume hat $\mu = \mathbb{E}[\xi_{1,1}] = \sum_j j p_i = \mu \in (0,\infty)$ and that $p_1 < 1$ (to avoid trivialities).

It is useful to establish that (Z_n/μ^n) is a (non-negative) martingale with respect to the natural filtration of the MC (*Exercise*). Hence $\lim_n Z_n/\mu^n$ exists a.s. and we denote the limit (non-negative) random variable by H. Check that if $\mathbb{E}[\xi_{1,1}^2] < \infty$ then the martingale is UI (Uniformly Integrable), hence in this case $H \not\equiv 0$. For $s \in (0,1]$ we introduce also $\varphi(s) = \mathbb{E}[s^{\xi_{1,1}}]$. Note that (*Exercise*) $\varphi(\cdot)$ is convex, increasing and smooth. Since $\varphi(0) = \lim_{s\searrow 0} \varphi(s) = p_0$ and $\varphi(1) = 1$, there exists only one solution in $[0,1)$ to the the fixed point equation $s = \varphi(s)$. Call this solution ϱ

PROPOSITION 2.1. *0 is a recurrent state for the chain (Z_n): note that 0 is accessible from any other state (i.e., $\rho_{n,0} > 0$) if and only if $p_0 > 0$. All other states n are transient. Moreover*

1. *if $\mu \le 1$ then $\sum_n \mathbf{1}_{Z_n > 0} < \infty$ a.s. (hence $H \equiv 0$);*
2. *if $\mu > 1$ then $\mathbb{P}(H = 0) = \rho$, hence $\mathbf{1}_{H=0}\sum_n \mathbf{1}_{Z_n>0} < \infty$ a.s., and if $\mathbb{E}[\xi_{1,1}^2] < \infty$ on the event $\{H > 0\}$ we have $Z_n \sim H\mu^n$ a.s. (here $\sim$ is aymptotic equivalence.*

The only (σ-finite) invariant measure can be normalized and it is δ_0.

Proof: *Exercise.*

3. Birth and death chain

$E = \mathbb{N} \cup \{0\}$ and Q is defined by

$$Q(j, j+1) \,=\, p_j\,, \quad Q(j, j-1) \,=\, q_j\,, \quad Q(j,j) \,=\, r_j, \tag{11}$$

with $p_j + q_j + r_j = 1$ for every j. We assume that $p_j > 0$ for every $j \in E$, $q_j > 0$ for every $j \in E \setminus \{0\}$ and $q_0 = 0$.

This MC is irreducible, moreover it is aperiodic if and only if there exists j such that $r_j > 0$ (*Exercise*).

We introduce the function $\varphi : E \mapsto [0, \infty)$ defined by $\varphi(0) = 0$, $\varphi(1) = 1$ and by imposing that

$$(Q\varphi)(k) = \varphi(k)\,, \tag{12}$$

for $k \in \mathbb{N}$. This yields $(\varphi(k+1) - \varphi(k)) = (q_k/p_k)(\varphi(k) - \varphi(k-1))$ for $k \geq 1$ and therefore for $n \geq 2$

$$\varphi(n) \;=\; 1 + \sum_{m=1}^{n-1} \prod_{j=1}^{m} \frac{q_j}{p_j}\,. \tag{13}$$

Note that $\lim_{n\to\infty} \varphi(n) =: \varphi(\infty)$ exists and takes value in $(1, \infty]$. We set $T_a := \inf\{n = 0, 1, \ldots :\, X_n = a\}$. By the Optional Stopping Theorem we have that, for $a < x < b$, $\varphi(x) = \mathbb{E}_x[\varphi(X_{T_a \wedge T_b}]$ from which we readily extract

$$\mathbb{P}_x(T_b > T_a) \;=\; \frac{\varphi(x) - \varphi(a)}{\varphi(b) - \varphi(a)} \;\Longrightarrow\; \mathbb{P}_x(T_b > T_0) \;=\; \frac{\varphi(x)}{\varphi(b)}\,. \tag{14}$$

Note finally that $\nu(x) := \prod_{k=1}^{x} (p_{k-1}/q_k)$ is reversible for Q.

We can therefore conclude that (*Exercise*):

> Proposition 3.1. *The birth and death MC is recurrent if and only if $\varphi(\infty) = \infty$. In this case ν is unique up to a multiplicative factor. Moreover it is finite recurrent if and only if $\sum_x \prod_{k=1}^{x} (p_{k-1}/q_k) < \infty$.*

Note that this result implies in particular that the symmetric simple random walk is null recurrent, as well as the well known fact that an asymmetric simple random walk is transient.

4. Discrete renewal processes

The basic object is $\tau = \{\tau_j\}_{j=0,1,\ldots}$ with $\tau_0 = 0$ and $\tau_j - \tau_{j-1} =: \eta_j$ and $(\eta_j)j = 1, 2, \ldots$ are IID variables taking values in $\{1, 2, \ldots, \infty\} = \mathbb{N} \cup \{\infty\}$. We set $K(n) := \mathbb{P}(\eta_1 = n)$ so in general $\sum_{n\in\mathbb{N}} K(n) \leq 1$, while $\sum_{n\in\mathbb{N}} K(n) + K(\infty) = 1$. We can view τ as a random subset of $\mathbb{N}\cup\{\infty\}$. In fact either $K(\infty) = 0$ and $|\tau| = \infty$ a.s. or $K(\infty) > 0$ and τ is a.s. a finite set (containing ∞). We set

$$u(n) \;:=\; \mathbb{P}(n \in \tau) \;=\; \mathbb{P}\left(\exists j \text{ such that } \tau_j = n\right)\,, \tag{15}$$

and $u(\cdot)$ is called *renewal function*. We say that τ is aperiodic if there exists no integer $p > 1$ such that $\{n \in \mathbb{N} : K(n) > 0\} \subset p\mathbb{N}$. We write η for η_1

THEOREM 4.1 (Renewal Theorem). *For an aperiodic renewal*

$$\lim_{n \to \infty} u(n) = \frac{1}{\mathbb{E}[\eta]} . \tag{16}$$

Of course the right-hand side of (16) is zero if $\mathbb{E}[\eta_1] = \infty$ and the statement holds without aperiodicity condition in this case. The proof is a direct application (*Exercise*) of the Ergodic Theorem for MC to one of the two MC's that we are going to build (and study) now.

4.1. The backward recurrence time. We set $A_n := n - \sup(\tau \cap [0, n]) \in E := \mathbb{N} \cup \{0\}$. Looking from time n, this is the time elapsed since the last renewal.

PROPOSITION 4.2. *A is a Q-MC with*

$$Q(j, j+1) = \mathbb{P}(\eta > j+1 | \eta > j) \quad \text{and} \quad Q(j, 0) = 1 - \mathbb{P}(\eta > j+1 | \eta > j) ,$$

for $j = 0, 1, \ldots$ with the convention that $\mathbb{P}(\eta > j+1 | \eta > j) = 0$ if $\mathbb{P}(\eta > j) = 0$. Various facts:

- *A is transient if and only if $\mathbb{P}(\eta = \infty) > 0$;*
- *A is positive recurrent if and only if $\mathbb{E}[\eta] < \infty$;*
- *A is aperiodic if and only if τ is;*
- *the unique (up to a factor) invariant measure μ satisfies*

$$\mu(n) = \mu(0)\mathbb{P}(\eta > n) ;$$

- *A has no invariant measure if (and only if) it is transient;*
- *A is positive recurrent if and only if $\mathbb{E}[\eta] < \infty$ and we remark, once μ is normalized, that $\mu(0) = 1/\mathbb{E}[\eta]$.*

Proof: *Exercise.*

4.2. The forward recurrence time. We set $B_n = \inf(\tau \cap [n, \infty]) - n \in E = \mathbb{N} \cup \{0, \infty\}$. Looking from time n, this is the time that is missing till the next renewal (unless at time n there is a renewal, so the next renewal is considered to be the one at n and $B_n = 0$).

PROPOSITION 4.3. *B is a Q-MC on E with*

$$Q(j, j-1) = 1 \text{ for } j \in \mathbb{N} \text{ and } Q(0, j) = \mathbb{P}(\eta = j+1) \text{ for } j \in \mathbb{N} \cup \{0, \infty\} ,$$

and $Q(\infty, \infty) = 1$. B is not irreducible: it is impossible to go from ∞ to $n = 0, 1, \ldots$. Various facts:

- *$E \setminus \{\infty\}$ is transient (in the obvious sense) if and only if $\mathbb{P}(\eta = \infty) > 0$;*
- *$E \setminus \{\infty\}$ is positive recurrent if and only if $\mathbb{E}[\eta] < \infty$;*
- *B is aperiodic if and only if τ is;*

- *the unique (up to a factor) invariant measure μ supported on $E \setminus \{\infty\}$ satisfies*
$$\mu(n) = \mu(0)\mathbb{P}(\eta > n) \quad \textit{for every } n \in E \setminus \{\infty\}\,;$$
- *B has δ_∞ as unique invariant measure if (and only if) $E \setminus \{\infty\}$ is transient;*
- *Assume $B_0 \neq \infty$ a.s.. Then B is positive recurrent if and only if $\mathbb{E}[\eta] < \infty$ and we remark, once μ is normalized, that $\mu(0) = 1/\mathbb{E}[\eta]$.*

Proof: *Exercise.*

5. Lindley process

The Lindley process is the MC on $E = [0, \infty)$ ($\mathcal{E} := \mathcal{B}(E)$) defined by the iteration map $f_\xi(x) := (x + \xi)_+$, $x \geq 0$ and $\xi \in \mathbb{R}$. That means that, given $W_0 \in E$ and an IID sequence $(\xi_j)_{j=1,2,\ldots}$, with W_0 and $(\xi_j)_{j=1,2,\ldots}$ independent, we define for $n \in \mathbb{N}$
$$W_n \,:=\, f_{\xi_n}(W_{n-1})\,. \tag{17}$$
Hence (W_n) is a $p-MC$ with $p(x, [0, y]) = \mathbb{P}(\xi_1 \leq y - x)$. In particular, $p(x, \{0\}) = \mathbb{P}(\xi_1 \leq -x)$. This process has immediate interpretations in terms of a storage processes, of basic queueing systems and of random walks with one barrier [**1**, Ch. III, Sec. 6]. In order to avoid trivialities, we assume that $\mathbb{P}(\xi_1 < 0) > 0$ as well as $\mathbb{P}(\xi_1 > 0) > 0$.

There is an interesting direct link between the Lindley process with $W_0 = 0$ and the random walk (S_n) with $S_0 = 0$ and increments $(\xi_j)_{j=1,2,\ldots}$. In fact, it is straightforward to see that the hitting times $(T_0^{(j)})_{j=0,1,\ldots}$ of 0 by (W_n) are the descending ladder times of (S_n): the descending ladder times are defined by setting $\tau_0 = 0$ and, for $k \geq 0$, $\tau_{k+1} := \inf\{n > \tau_k \,:\, S_n \leq S_{\tau_k}\}$. So the ladder times are the times in which the walk hits a new minimum. The identity we just claimed is actually *pathwise*: that is $T_0^{(j)}(\omega) = \tau_j(\omega)$ for every ω.

Since it is immediate to check that 0 is accessible for every $x \in E$, i.e. $\mathbb{P}_x(T_0 < \infty) > 0$, showing recurrence or transience of 0 is crucial. One can actually show that (S_n) has only three possible behaviors [**1**, p. 224, Th. 2.4]:

(1) either $S_n \to \infty$ a.s.;
(2) or $S_n \to -\infty$ a.s.;
(3) or $\limsup_n S_n = +\infty$ and $\liminf_n S_n = -\infty$ a.s.;

and $\mathbb{P}(\tau_1 < \infty) = 1$ if and only if we are in cases (2) or (3). Therefore ((2) or (3)) is a necessary and sufficient condition for recurrence of the Lindley process. One can actually show that (W_n) is null recurrent if and only if (S_n) is in case (3), and this happens if and only if $\xi \in \mathbb{L}^1$ and $\mathbb{E}\xi_1 = 0$ or $\xi \sim -\xi$.

These results are rather advanced, but it is for example immediate to see that $W_n \geq S_n$, hence $W_n \to \infty$ if $S_n \to \infty$ (in fact: $W_n \sim S_n$ because S_n differs from W_n only because of what happens up to an a.s. finite time). Therefore, this happens in particular if $\xi \in \mathbb{L}^1$ and $\mathbb{E}[\xi_1] > 0$ and in this case one obtains without difficulty also that $W_n/n \to \mathbb{E}[\xi_1]$ a.s.. Therefore (W_n) is transient for $\mathbb{E}[\xi_1] > 0$ (in the sense that 0 is visited only a finite number of times, but of course this result says that also $[0, x]$ is visited only finitely many times).

Another case that we can treat in detail with relatively elementary methods is the case in which $S_n \to -\infty$, hence in particular if $\xi \in \mathbb{L}^1$ and $\mathbb{E}[\xi_1] < 0$. In analyzing this case the following magic identity (that involves no probability!) will the of help:

LEMMA 5.1. *If $\xi_1, \xi_2, \ldots$ are real numbers, then for every $n = 1, 2, \ldots$ and $x \geq 0$ we have*

$$f_{\xi_1} \circ f_{\xi_2} \circ \ldots \circ f_{\xi_2}(x) = \max\left(0, \xi_1, \xi_1 + \xi_2, \ldots, \xi_1 + \ldots + \xi_{n-1}, \xi_1 + \ldots + \xi_{n-1} + \xi_n + x\right) . \quad (18)$$

PROOF. Of course it suffices to consider the case $x = 0$. We can proceed by induction: the case $n = 1$ follows from the definition of $f_\xi(\cdot)$. To prove the induction step it suffices to exploit that $(x+y_+)_+ = \max(0, x, x+y)$ (this identity is established by considering separately the cases $y \geq 0$ and $y < 0$). □

The consequences of Lemma 5.1 are immediate: if $\lim_n \xi_1 + \xi_2 + \ldots \xi_n = -\infty$ we have that for every x

$$\lim_n f_{\xi_1} \circ f_{\xi_2} \circ \ldots \circ f_{\xi_2}(x) = \max\left(0, \xi_1, \xi_1 + \xi_2, \xi_1 + \xi_2 + \xi_3, \ldots\right) \in [0, \infty). \quad (19)$$

This is still just a deterministic result, but of course if (ξ_j) are IID and $S_n \to -\infty$ a.s. we have that, with $Y_n := f_{\xi_1} \circ f_{\xi_2} \circ \ldots \circ f_{\xi_2}(x)$, $\lim_n Y_n$ converges almost surely to a limit that we call Y_∞ (see the right-hand side of (19) for an explicit expression).

Note however that we are not interested in (Y_n), but in (W_n)! This is not a minor change:

- (Y_n) is not (at all) a MC and it the trajectories of (Y_n) have very little to do with those of (W_n). In fact, Lemma 5.1 is telling us that it is a non decreasing process (while (W_n) oscillates a lot!).
- nevertheless $Y_n \sim W_n$ for every n, just because the ξ_j variables are exchangeable. We stress that this holds only for the 1-marginal: we do not claim (at all) for example that $(Y_1, Y_2) \sim (W_1, W_2)$.

From these considerations we can extract

PROPOSITION 5.2. *If $S_n \to -\infty$ a.s., in particular if $\mathbb{E}[(\xi_1)_+] < \infty$ and $\mathbb{E}[\xi_1] \in [-\infty, 0)$, we have that (W_n) converges in law, for every choice of $W_0 = x$, to W_∞. The law of W_∞ is the unique invariant probability of the Markov chain and $\mathbb{P}(W_\infty = 0) \in (0, 1)$.*

PROOF. *Exercise* □

Proposition 5.2 collects the results that one obtains directly from Lemma 5.1 and the trick of inversing the order of the ξ variables. But once we know that there is an invariant probability for which 0 has positive probability, we can extract that 0 is recurrent and exploit this to obtain (*Exercise*)

PROPOSITION 5.3. *If $S_n \to -\infty$ a.s. we have that the law of W_n converges in total variation, for every choice of $W_0 = x$, to the law of W_∞.*

Note that if there exists $a > 0$ such that $\mathbb{P}(S_n \in a\mathbb{Z}$ for every $\mathbb{Z})$ (for this we need that $x \in a\mathbb{Z}$, but even if x is not in this lattice, the process will reach it the first time it hits 0), the the Lindley MC lives on $a\mathbb{N} \cup \{0\}$ (and, of course, our analysis applies also to this case).

the Lindley process lives on a lattice if this latti

6. Random Coefficient Autoregressive MC

The basic Random Coefficient Autoregressive (RCA) MC is defined iteratively, once the initial condition $X_0 = x \in \mathbb{R}$ is given, by

$$X_{n+1} \,=\, A_{n+1}X_n + B_{n+1} \,=:\, f_{A_{n+1},B_{n+1}}(X_n)\,, \tag{20}$$

with $((A_n, B_n))_{n\in\mathbb{N}}$ and IID sequence of random variables taking values in $\mathbb{R}^2$. Note that in this case

$$|f_{A,B}(x) - f_{A,B}(y)| \,\le\, |A|\,|x-y|\,, \tag{21}$$

so this process has good contractive properties if $\log|A| \in \mathbb{L}^1$ and $\mathbb{E}\log|A| < 0$. In order to make this clear and explicit let us consider the particular case in which $B_n = A_n$ for every n, so

$$X_{n+1} \,=\, A_{n+1}\,(1+X_n) \,=:\, f_{A_{n+1}}(X_n)\,, \tag{22}$$

and let us simplify things a bit by choosing the A variables non negative and non trivial (recall that we assume $\log A \in \mathbb{L}^1$ and $\mathbb{E}\log A < 0$). In this case it is not difficult to see that if $x < 0$, the stopping time $\inf\{n : X_n \ge 0\}$ is a.s. finite, so the negative semi-axis is transient, but we will work in any case with $E = \mathbb{R}$ and what we just claimed will come out of the analysis we will develop.

By direct inspection it is not difficult to see that

$$\begin{aligned} X_n \,&=\, f_{A_n}\circ\ldots\circ f_{A_1}(x) \,= \\ &A_n + A_nA_{n-1} + A_nA_{n-1}A_{n-2} + \ldots + A_nA_{n-1}\cdots A_2 + A_nA_{n-1}\cdots A_1(1+x)\,. \end{aligned} \tag{23}$$

If we *reverse the engine* of this MC we obtain

$$\begin{aligned} Y_n \,&:=\, f_{A_1}\circ\ldots\circ f_{A_n}(x) \,= \\ &\qquad A_1 + A_1A_2 + A_1A_2A_3 + \ldots + A_1A_2\cdots A_{n-1} + A_1A_2\cdots A_n(1+x)\,, \end{aligned} \tag{24}$$

which is an increasing process if $x \ge 0$. But even if $x < 0$, it is not difficult to see that $\lim_n Y_n := Y_\infty$

$$Y_\infty \,:=\, \sum_{n=1}^\infty \prod_{j=1}^n A_j\,, \tag{25}$$

exists and it is a.s. finite. In fact, by the law of large numbers $(1/n)\log A_1A_2\cdots A_n \longrightarrow \mathbb{E}\log A<0$, so for every $\beta\in(\exp(\mathbb{E}\log A),1)$ there exists $C(\omega)$ (C is an a.s. finite random variable) such that for every $n\in\mathbb{N}$

$$A_1(\omega)A_2(\omega)\cdots A_n(\omega)\;\le\; C(\omega)\beta^n\,. \tag{26}$$

This suffices to show that $(Y_n(\omega))$ is a.s. a Cauchy sequence and therefore the limit $\lim_n Y_n =: Y_\infty$ exists a.s.. Note that (26) yields also that the limit is independent of the value of x. Note moreover that Y_∞ is supported on $[0,\infty)$. And now by exploiting that $X_n\sim Y_n$ for every n we can complete (*Exercise*) the proof of

PROPOSITION 6.1. *The random coefficient autoregressive process defines by (22) has a unique invariant probability ν and for every initial condition $x\in\mathbb{R}$ we have (X_n) converges in law to Y_∞.*

This result can be easily generalized almost (verbatim) to the case in which $B=B_1$ is random with $\mathbb{E}[(\log|B|)_+]<\infty$ and to the case in which the A and B variables assume also negative values. Of course the invariant probability will no longer be supported on the positive semi-axis.

REMARK 6.2. *The invariant probability is more interesting than it looks at first. Note, for example, that even if $A(=B)\ge 0$ is a bounded random variable, Y_∞ may not even be in $\mathbb{L}^1$. In fact*

$$\mathbb{E}[Y_\infty]\;=\;\sum_{n=1}^{\infty}(\mathbb{E}[A])^n\,, \tag{27}$$

so $\mathbb{E}[Y_\infty]<\infty$ if and only if $\mathbb{E}[A]<1$. With some work (but not too much) it is also possible to see that an analogous result holds for all moments: $\mathbb{E}[Y_\infty^k]<\infty$ if and only if $\mathbb{E}[A^k]<1$. Therefore Y_∞ has a heavy tail unless $\mathbb{E}[A^k]<1$ for every k, which requires $\mathbb{P}(A<1)=1$ (i.e., that the process is contractive for every ω, not just in a probabilistic sense).

Sticking for simplicity to the case (22), let us address the question of whether this process is a Harris MC. This requires conditions on law of A. We will not try to look for optimal conditions and we start by observing that if A is a continuous random variable – we denote by f its density – then measure $p(x,\cdot)$, p is the transition kernel, has density $y\mapsto f(y/(1+x))/(1+x)$. In particular for $x=1$ the map is $y\mapsto f(y/2)/2$ Therefore if $f(1/2)>0$, by continuity we can find $\varepsilon_0>0$ and $\delta>0$ such that $f(y/(1+x))/(1+x)\ge\varepsilon_0$ uniformly in $x,y\in[1-\delta,1+\delta]$. Therefore we can choose $\mathbf{A}=\mathbf{B}=[1-\delta,1+\delta]$, ρ the uniform mesure on $\mathbf{B}$, and $\varepsilon=2\delta\varepsilon_0$ and satisfy the second of the Harris requirement.

But also the first Harris requirement is fulfilled. In fact it suffices to show that for every $x\in\mathbb{R}$ we can find a value of n and $a_1,\ldots,a_n\in I_\eta:=(-\eta+1/2,\eta+1/2)$ (we are choosing η so that $\inf_{I_\eta}f>0$) we have that

$$a_n+a_na_{n-1}+\ldots+\ldots+a_na_{n-1}\cdots a_2+a_na_{n-1}\cdots a_1(1+x)\in(1-\delta,1+\delta)\,. \tag{28}$$

Once this is established, the result is obtained because (28) holds also in an open neighborhood of $(a_1, \ldots, a_n)$. The requirement (28) may appear difficult to establish, but it is not the case. In fact, it suffices to remark that if $x < 1$ then $x < (1+x)/2 < 1$, so even by choosing simply $a_j = 1/2$ for every j we will hit the target (of becoming larger than $1-\delta$) in a finite number of steps. Analogous reasoning for $x > 1$.

As a matter of fact, the argument we just developed shows that the chain is Harris with $\mathbf{A} = \mathbf{B} = [1-\delta, 1+\delta]$ and suitable choice of ε (ρ is the uniform probability on $\mathbf{B}$) under the assumption that the law of A is bounded below by a measure with a density f that is continuous and $f(1/2) > 0$.

Remark 6.3. *A similar argument can be developed if $f(x) > 0$ for $x \in (1/2, 1]$.*

Remark 6.4. *Another interesting point is to notice that the minimum of the support of the invariant probability can be determined with precision: if we call a, $a < 1$ by hypothesis, the minimum of the support of the law of A, then by iterating from $x = 0$ we see that the support of the invariant probability does not go below $a + a^2 + \ldots = a/(1-a)$. Just a slightly more involved argument gives the bound in the other direction. On the other hand the supremum of the support of the invariant probability is $+\infty$, since we are assuming that the maximum of the support of A is larger than one.*

In the cases in which we are able to prove that the MC is Harris, we can apply the general theory and conclude that , under the assumption that $\mathbb{E}[\log A] < 0$, the only invariant measure is the invariant probability and that $p_n(x, \cdot)$ converges to the invariant probability in total variation distance for every x such that $\mathbb{P}_x(T_{\mathbf{A}} < \infty) = 1$. But we know by the weak convergence result (obtained by exploiting the convergence in law result in Proposition 6.1) that $\lim_n \mathbb{E}_x[h(X_n)] = \int h \, \mathrm{d}\nu$ for every x (ν is the invariant probability). By choosing h to be a smoothed version of $\mathbf{1}_{\mathbf{A}}$ we see that the chain that starts from x visits $\mathbf{A}$ a.s., that is $\mathbb{P}_x(T_{\mathbf{A}} < \infty) = 1$ for every x.

7. A Foster-Lyapunov argument for (null) recurrence, with applications

Roughly, we would like to say that if we can find V bounded below such that $\lim_{x:\,|x|\to\infty} V(x) = \infty$ and such that $pV \le V$ outside of a compact set, then the p-MC is recurrent. In order to minimize introducing definitions, we give a minimal version of this result that is tailored to the two applications we give.

Proposition 7.1. *Let us consider $E = [l, \infty)$, $l \in \mathbb{R}$, and a p-MC on this space. We assume that there exists $V : E \to [0, \infty)$ which outside of a bounded set satisfies two properties:*

1. *V is (strictly) increasing and* $\lim_{x\to\infty} V(x) = \infty$*;*
2. *$pV \le V$.*

We also assume that for every x outside of a bounded set there exists $\varepsilon > 0$ such that $\inf_{y \in B_\varepsilon(x)} \mathbb{P}_y(T_{B_\varepsilon(x)^\complement} < \infty) > 0$ *with* $B_\varepsilon(x) := \{z : |z - x| < \varepsilon\}$*. Then there exists $r > l$ such that for every $x \in E$ we have* $\mathbb{P}_x(T_{[l,r]} < \infty) = 1$.

The condition on $\mathbb{P}_y(T_{B_\varepsilon(x)^\complement} < \infty) > 0$ is a very weak requirement (see applications) in order to avoid that the process gets stuck at some x for arbitrarily large x.

The proof is given as a guided exercise.

Proof. First of all note (*Exercise*) that the hypothesis involving $B_\varepsilon(x)$ implies that $\mathbb{P}_x(\sum_n \mathbf{1}_{B_\varepsilon(x)^\complement}(X_n) = \infty) = 1$. Moreover, we can choose $r > l$ such that the hypotheses hold for $x \geq r$. We set for conciseness $T = T_{[l,r]}$ (the hitting time of $[l, r]$) and we introduce $Y_n := V(X_{n\wedge T})$ for $n = 0, 1, \ldots$. The process $(Y_n)_{n=0,1,\ldots}$ is a super-martingale for every choice of $Y_0 = x \in E$ (*Exercise*: note that by iterating $pV \leq V + b\mathbf{1}_{[l,r]}$, that holds by hypothesis for a positive b, we obtain $Y_n \in \mathbb{L}^1$ for every n). Since $Y_n \geq 0$ we have that $(Y_n(\omega))$ converges to a limit that we call $Y_\infty(\omega) < \infty$ for every $\omega \in G$, with $\mathbb{P}_x(G) = 1$.

If there exists x such that $\mathbb{P}_x(T_{[l,r]} = \infty) > 0$ then for $\omega \in \{T_{[l,r]} = \infty\} \cap G$ we have that $X_n(\omega) > r$ for every n and therefore $Y_n(\omega) = V(X_n(\omega)) \longrightarrow Y_\infty(\omega) < \infty$. Since V diverges at infinity and since it is a bijection on the region we consider, we obtain that $(X_n(\omega))$ converges to a limit in $X_\infty(\omega) \in [r, \infty)$. But this means that $X_n(\omega) \in B_\varepsilon(X_\infty(\omega))$ for every $\varepsilon > 0$ and all n large and this is incompatible with our hypothesis on the exit probability. Therefore $\mathbb{P}_x(T_{[l,r]} = \infty) = 0$ for every x. $\square$

Application to the Lindley MC with centered drift. We consider the Lindley MC with $\mathbb{E}[\xi] = 0$ (we recall that we exclude the trivial case of $\xi \equiv 0$). We treat only the case in which there exists L such that $\mathbb{P}(\xi < -L) = 0$. This is an assumption that simplifies (a lot!) the analysis: it is possible to generalize this result at least to the case in which the variance of ξ is finite (with the very same choice of V). We choose $V(x) = \log(1+x)$, $x \in [0, \infty) = E$. We remark that the following elementary bound holds: for $y > -1$

$$\log(1 + y) \,\leq\, y - \frac{1}{2} y^2 \mathbf{1}_{y<0}\,. \tag{29}$$

Then

$$pV(x) \,=\, \mathbb{E}\left[\log\left(1 + (x+\xi)_+\right)\right] \,=\, V(x) + \mathbb{E}\left[\log\left(1 + \frac{(x+\xi)_+ - x}{1+x}\right)\right]. \tag{30}$$

If $x > L$ then a.s. $(x + \xi)_+ = x$. Therefore for $x > L$

$$pV(x) - V(x) \,=\, \mathbb{E}\left[\log\left(1 + \frac{\xi}{1+x}\right)\right] \,\leq\, -\frac{1}{2(1+x)^2}\mathbb{E}[\xi^2;\, \xi < 0]\,. \tag{31}$$

Therefore $pV(x) < V(x)$ for every $x > L$. Finally, in this case and outside $[0, L]$, the evolution is just a random walk and since ξ is centered and nontrivial, we have that $p(y, B_y(2\varepsilon)^\complement) = \mathbb{P}(\xi \in B_0(2\varepsilon)^\complement) =: p_\varepsilon > 0$ for $y > L$ and $\varepsilon > 0$, so $p(y, B_x(\varepsilon)) \geq p_\varepsilon$ for every $y \in B_x(\varepsilon)$ and $x > L + \varepsilon$. Therefore we can apply Proposition 7.1 and the set $[0, L+\varepsilon]$ is visited infinitely often by the MC. From this one easily extracts that also 0 is visited infinitely often (so 0, and the whole MC since 0 is accessible from every x, is recurrent).

Application to the RCA MC with $\mathbb{E}[\log A] = 0$*.* Also in this case we simplify our life by making a strong assumption on the support of $\log A$: there exists L such that $\mathbb{P}(\log A > -L) = 1$.

The first step is to work with $Z_n := \log X_n$:

$$Z_{n+1} \;=\; \log A_{n+1} + \log\left(1 + \exp(Z_n)\right) \;=\; \log A_{n+1} + Z_n + \log\left(1 + \exp(-Z_n)\right), \quad (32)$$

which makes clear that Z behaves almost as a random walk when it is positive and far from the origin. On the other hand, Z has a lot of difficulty to enter the negative semi axis. In fact, since $\xi \geq -L$, if $Z_0 < -L$ then $Z_1 \geq -L$ and $Z_n \in [-L, \infty)$ for every $n \geq 1$. So we can choose $E = [-L, \infty)$.

We then choose $V(x) = \log_+(x)$ for $x > 0$ and $V(x) = 0$ for $x \in [-L, 0]$. Since $\xi \geq -L$ for $x \geq L + 1$

$$pV(z) \;=\; \log z + \mathbb{E}\left[\log\left(1 + \frac{\log A}{z} + \frac{\log(1+\exp(-z))}{z}\right)\right], \quad (33)$$

so by (29)

$$\begin{aligned} pV(z) \;=&\; V(z) + \frac{\log(1+\exp(-z))}{z} - \\ &\frac{1}{2z^2}\mathbb{E}\left[\left(\log A + \log(1+\exp(-z))\right)^2; \log A + \log(1+\exp(-z)) < 0\right] \\ &\overset{z\to\infty}{=} V(z) + O\left(\frac{\exp(-z)}{z}\right) - \frac{1}{2z^2}\left(\mathbb{E}\left[(\log A)^2; \log A < 0\right] + o(1)\right). \quad (34) \end{aligned}$$

Therefore there exists $M > L + 1$ such that $pV(z) \leq V(z)$ for every $z \geq M$. The argument to show that Z cannot visit infinitely many times a neighborhood of a point goes pretty much as for the Lindley case. Therefore we can apply Proposition 7.1 and conclude that $[-L, M]$ is a.s. visited by Z and therefore $[\exp(-L], \exp(M)]$ is a.s. visited by the RCA MC X.

Bibliography

[1] S. Asmussen, *Applied probability and queues*, Second edition, Springer-Verlag, New York, 2003.

[2] J.-F. Le Gall, *Intégration, Probabilités er Processus Aléatoires*, *google* the title, otherwise https://www.math.u-psud.fr/ jflegall/IPPA2.pdf

[3] D. Williams, *Probability with Martingales*, Cambridge Mathematical textbooks, Cambridge University Press, 1991.