Markov processes – Course note 2.

Martingale problems, recurrence properties of discrete time chains.

[version 1, 2017.11.1]

We introduce the notion of martingale problem for Markov chains and use it to discuss stochastic stability and recurrence properties of discrete state chains.

See also Chapter 1 of Prof. Eberle's "Markov Process" course.

1 Martingale problem

Let $(X_n)_n$ be a Markov chain on E and let $f: E \rightarrow \mathbb{R}$ be a bounded function. One way to describe the dynamics associated to the process $(X_n)_n$ is to use the decomposition of the process $f(X_n)$ as a martingale and a predictable process:

$$f(X_n) = M_n^f + A_{n-1}^f, \qquad n \geqslant 1.$$

There is only one such decomposition and it follows that $A_{n-1}^f = \sum_{k=0}^{n-1} (\mathcal{L}f)(X_k)$, where $\mathcal{L}f = (\pi - I)f$, $\pi = P_1$ the one step transition kernel and I the identity transition kernel. The operator $\mathcal{L}: \mathcal{F}_b(E) \rightarrow \mathcal{F}_b(E)$ is called the *generator* of the discrete time chains. The martingale property of $(M_n^f)_n$ and the generator characterise completely the Markov chain.

Theorem 1. *$(X_n)_n$ is a Markov chain with one–step transition kernel π iff for any $f \in \mathcal{F}_b(E)$ the process*

$$M_n^f := f(X_n) - \sum_{k=0}^{n-1} (\mathcal{L}f)(X_k), \qquad n \geqslant 0,$$

is a martingale.

The proof is left as an exercise.

We postpone to discuss the continuous time formulation of the martingale problem which requires some analytical sophistication. Now we concentrate to illustrate the connection between martingale and Markov properties and use martingale to establish basic criteria for stability and recurrence of Markov chains.

The *exterior boundary* ∂D of a set $D \in \mathcal{E}$ w.r.t. the Markov chain is given by

$$\partial D = \cup_{x \in D} \{\operatorname{supp} \pi(x, \cdot)\} \backslash D.$$

where the support $\operatorname{supp}\mu$ of the measure μ is defined as the smallest closed set such that $|\mu|(A^c)=0$. Open sets contained in $(D\cup\partial D)^c$ cannot be reached by the chain in one step starting from D.

Example 2. For the simple random walk on $\mathbb{Z}^d$: $\partial D=\{y\in\mathbb{Z}^d\colon \exists x\in D\colon |x-y|=1\}$.

Recall that we denote with T_D the hitting time of D: $T_D=\inf\{n\geqslant 0\colon X_n\in D\}$.

Interesting quantities of a Markov chain are:

a) The exit probability from D starting at $x\in D$: $\mathbb{P}_x(T_{D^c}<\infty)$;

b) The "law" of the exit point: $\mathbb{P}_x(X_{T_{D^c}}\in B, T_{D^c}<\infty)$;

c) The mean exit time: $\mathbb{E}_x[T_{D^c}]$;

d) The average occupation of B before exiting D:

$$G_D(x,B)=\mathbb{E}_x[\sum_{k=0}^{T_{D^c}-1}\mathbb{I}_B(X_k)]=\sum_{k\geqslant 0}\mathbb{P}_x(X_k\in B, k<T_{D^c}),$$

(Green kernel of D);

e) The Laplace transform of the exit time: $\mathbb{E}_x[e^{-\lambda T_{D^c}}]$;

f) The Laplace transform of the occupation time of B: $\mathbb{E}_x\Big[e^{-\lambda\sum_{k=0}^{T_{D^c}-1}\mathbb{I}_B(X_k)}\Big]$;

Let $v,c\in\mathcal{F}_+(E)$ and consider the process

$$M_n:=v(X_n)+\sum_{k=0}^{n-1}c(X_k)=M_n^v+\sum_{k=0}^{n-1}(\mathcal{L}v+c)(X_k)$$

where we made explicit its Doob's decomposition. If $\mathcal{L}v+c\leqslant 0$ we deduce that $(M_n)_n$ is a non-negative supermartingale, the same applies to the stopped process $M_n^T=M_{n\wedge T}$ where T is a stopping time. By the supermartingale convergence theorem the limit $M_\infty:=\lim_n M_n$ exists almost surely and in particular we have $M_n^T\to M_T$ a.s. everywhere, even on $\{T=+\infty\}$.

Let now $D\in\mathcal{E}$, $T=T_{D^c}$, $v,f\in\mathcal{F}_+(D)$ such that $\mathcal{L}v+c\leqslant 0$ on D and $v\geqslant f$ on ∂D. Then

$$M_T\geqslant v(X_T)\mathbb{I}_{T<\infty}+\sum_{k=0}^{T-1}c(X_k)>f(X_T)\mathbb{I}_{T<\infty}+\sum_{k=0}^{T-1}c(X_k)$$

and by Fatou we conclude

$$u(x):=\mathbb{E}_x\Bigg[f(X_T)\mathbb{I}_{T<\infty}+\sum_{k=0}^{T-1}c(X_k)\Bigg]\leqslant\mathbb{E}_x[M_T]\leqslant\mathbb{E}_x[M_0]=v(x),\qquad x\in D.$$

Then all non–negative solutions to

$$\begin{array}{ll} \mathcal{L}v + c = 0 & \text{on } D \\ v = f & \text{on } \partial D \end{array} \tag{1}$$

dominate u. Let now assume that $T < \infty$ a.s., then by the Markov property we have, for all $x \in D \backslash \partial D$

$$\mathbb{E}_x[f(X_T) + \sum_{k=0}^{T-1} c(X_k)|\mathcal{F}_1] = \mathbb{E}_x[f(X_T) + \sum_{k=1}^{T-1} c(X_k)|\mathcal{F}_1] + c(X_0)$$

$$= \mathbb{E}_x[(f(X_T) + \sum_{k=0}^{T-1} c(X_k)) \circ \theta_1|\mathcal{F}_1] + c(X_0)$$

$$= \mathbb{E}_{X_1}[f(X_T) + \sum_{k=0}^{T-1} c(X_k)] + c(X_0) = u(X_1) + c(X_0)$$

since $T > 1$ $\mathbb{P}_x - a.s.$ if $x \in D \backslash \partial D$. From this we conclude that

$$0 = \mathbb{E}_x[u(X_1) + c(x) - u(x)] = (\mathcal{L}u + c)(x), \qquad x \in D \backslash \partial D$$

and moreoever $u(x) = f(x)$ if $x \in \partial D$. So u is a solution of the problem (1). On the other hand, if v is a solution to (1) and f is a bounded function then we have $u = v$ since all the inequalities in the above supermartingale argument become equalitites. That is, solution of the above problem are unique.

Remark 3. Extension to absorbed chains are discussed in Eberle's notes.

2 Recurrence for countable Markov chains

These results hints to the link between superharmonic functions (i.e. $\mathcal{L}v \leqslant 0$) and asymptotic behaviour of the process. In order to understand better this connection we consider the case of Markov chains on a countable state space E.

Introduce the first return time to $A \in \mathcal{E}$ as

$$T_A^+ := \inf \{n \geqslant 1 \colon X_n \in A\}$$

and let $T_x^+ = T_{\{x\}}^+$ for $x \in E$.

Definition 4. *A state $x \in E$ is recurrent if $\mathbb{P}_x(T_x^+ < \infty) = 1$ otherwise is transient. Is positive recurrent if $\mathbb{E}_x[T_x^+] < +\infty$.*

Let $N_x = \#\{n \geqslant 1 \colon X_n = x\} = \sum_{n \geqslant 1} \mathbb{I}_{X_n = x}$ the number of visits of $x \in E$. By strong Markov, recurrence is equivalent to require that $\mathbb{P}_x(N_x = +\infty) = 1$. If x is transient we have $\mathbb{P}_x(N_x = +\infty) = 0$ (see below for the proof).

Introduce a sequence of passage times at $x \in E$: $T_x^0 = 0$, $T_x^n = \inf\{k > T_x^{n-1} : X_k = x\}$, $n \geqslant 1$. For $n \geqslant 1$, if $T_x^{n-1} < +\infty$ let $\tau_x^n := T_x^n - T_x^{n-1}$.

Proposition 5. (Regeneration) *Let $x \in E$ and $n \geqslant 1$. Conditionally on $\{T_x^n < +\infty\}$ the law of τ_x^{n+1} is independent of $(T_x^1, \ldots, T_x^n)$ and*

$$\mathbb{P}(\tau_x^{n+1} = k | T_x^n < +\infty) = \mathbb{P}_x(T_x = k), \qquad k \in \mathbb{N} \cup \{+\infty\}.$$

Proof. Exercise. □

Lemma 6. *For $n \geqslant 0$ we have that $\mathbb{P}_x(N_x \geqslant n) = f_x^n$ with $f_x := \mathbb{P}_x(T_x < +\infty)$.*

Proof. Use strong Markov. □

Remark 7. For any r.v. $X: \Omega \rightarrow \mathbb{N}$ we have

$$\mathbb{E}[X] = \mathbb{E}[\sum_{k \geqslant 0} 1_{k \leqslant X}] = \sum_{k \geqslant 0} \mathbb{P}(X \geqslant k) \tag{2}$$

Theorem 8. *There is the following dychotomy:*

i. $\mathbb{P}_x(T_x < \infty) = 1 \Rightarrow \mathbb{P}_x(N_x = \infty) = 1$ *and* $\sum_{n \geqslant 1} \pi^n(x, x) = +\infty$*;*

ii. $\mathbb{P}_x(T_x < \infty) < 1 \Rightarrow \mathbb{P}_x(N_x = \infty) = 0$ *and* $\sum_{n \geqslant 1} \pi^n(x, x) < +\infty$ *.*

Proof. If $f_x = \mathbb{P}_x(T_x < \infty) = 1$ then by Lemma 6 we have

$$\mathbb{P}_x(N_x = +\infty) = \lim_{n \rightarrow \infty} \mathbb{P}_x(N_x \geqslant n) = \lim_{n \rightarrow \infty} f_x^n = 1$$

and then $\mathbb{P}_x(N_x = \infty) = 1$ and $\infty = \mathbb{E}_x[N_x] = \mathbb{E}_x[\sum_{n \geqslant 1} 1_{X_n = x}] = \sum_{n \geqslant 1} \pi^n(x, x)$. On the other hand if $f_x < 1$ then by eq. (2) and Lemma 6,

$$\sum_{n \geqslant 1} \pi^n(x, x) = \mathbb{E}_x[N_x] = \sum_{n \geqslant 0} \mathbb{P}_x(N_x \geqslant n) = \sum_{n \geqslant 0} f_x^n = \frac{1}{1 - f_x} < +\infty,$$

which implies that $\mathbb{P}_x(N_x = +\infty) = 0$. □

Introduce a transitive relation $x \rightarrow y$ on states $x, y \in E$ when one of the following equivalent condition holds

a) $\mathbb{P}_x(T_y < +\infty) = 1$;

b) $\pi^n(x, y) > 0$ for some $n \geqslant 0$;

c) there exists a sequence of states $(x_k)_{k = 0, \ldots, n}$ such that $x_0 = x$, $x_n = y$ and $\pi(x_k, \{x_{k+1}\}) > 0$.

Let $x \leftrightarrow y$ the equivalence relation given by $x \rightarrow y$ and $y \rightarrow x$. This equivalence relation induce a partition of E into equivalence classes, usually called communication classes.

When there is only a class we say that the chain is *irreducible*.

Theorem 9. *All the states in the same class are of the same type (either transient or recurrent).*

Proof. If $x \leftrightarrow y$ then there are N, M such that $\pi^N(x, y) > 0$ et $\pi^M(y, x) > 0$. A simple bound gives

$$\pi^{2N+n+2M}(x,x) \geqslant \pi^N(x,y)\pi^{N+n+M}(y,y)\pi^M(y,x) \geqslant [\pi^N(x,y)\pi^M(y,x)]^2\pi^n(x,x)$$

for all $n \geqslant 1$. Let $\alpha = \pi^N(x,y)\pi^M(y,x) > 0$, then

$$\sum_{k\geqslant 0} \pi^k(x,x) \geqslant \sum_{k\geqslant 2N+2M} \pi^k(x,x) \geqslant \alpha \sum_{k\geqslant N+M} \pi^k(y,y) \geqslant \alpha^2 \sum_{k\geqslant 0} \pi^k(x,x)$$

and then the states x, y are both either transient or recurrent. □

Remark 10. An irreducible chain is either recurrent or transient.

Proposition 11. *A finite set $A \subseteq E$ such that $\pi(x, A^c) = 0$ for all $x \in A$ contains at least one recurrent state. A finite irreducible chain is recurrent.*

Proof. Let $\#A < +\infty$ and assume that for all $z \in A$, $\mathbb{P}_z(N_z = +\infty) = 0$. Fix $x \in A$, for all $z \in A$ eq. (?) gives

$$\mathbb{P}_x(N_z \geqslant r) = \mathbb{P}_x(T_y < +\infty)\mathbb{P}_z(N_z \geqslant r).$$

Taking limits for $r \rightarrow +\infty$ we get $\mathbb{P}_x(N_z = +\infty) = \mathbb{P}_x(T_z < +\infty)\mathbb{P}_z(N_z = +\infty) = 0$ for all $z \in A$ and as a consequence

$$1 = \mathbb{P}_x(\cap_{z\in A}\{N_z < +\infty\}) = \mathbb{P}_x(\sum_{z\in A} N_z < +\infty) = \mathbb{P}_x(\sum_{n\geqslant 0} 1_{X_n\in A} < +\infty)$$

since $\sum_{z\in A} N_z = \sum_{z\in A}\sum_{n\geqslant 0} 1_{X_n = z} = \sum_{n\geqslant 0} 1_{X_n\in A}$ is the time passed in A by the chain. Since the set A is closed we have $\mathbb{P}_x(X_n \in A) = 1$ for all $n \geqslant 0$ which should imply that the time spent in A is infinite. A contradiction. □

3 Forster–Lyapounouv criteria for recurrence

We assume that the chain is irreducible.

Theorem 12. *A discrete Markov chain is*

a) Transient iff there exists $V \in \mathcal{F}_+(E)$ and a set $A \subseteq E$ and a state $y \in E$ such that $\mathcal{L}V \leqslant 0$ on A^c and $V(y) < [\inf_A V]$;

b) *Recurrent iff there exists* $V \in \mathcal{F}_+(E)$ *such that* $\#\{\mathcal{L}V > 0\} < \infty$ *and* $\#\{V \leqslant N\} < \infty$ *for all* $N \geqslant 0$.

c) *Positive recurrent iff there exists* $V \in \mathcal{F}_+(E)$ *such that* $\#\{\mathcal{L}V > -1\} < \infty$.

Proof. Case (a). Consider the process $M_n = V(X_{n \wedge T_A})$ is a non–negative supermartingale. By optional stopping

$$V(y) \geqslant \mathbb{E}_y[V(X_{n \wedge T_A})] \geqslant \mathbb{E}_y[V(X_{T_A})\mathbb{I}_{T_A < \infty}] \geqslant \left[\inf_A V\right] \mathbb{P}_y(T_A < \infty),$$

so if $V(y) < [\inf_A V]$ we obtain $\mathbb{P}_y(T_A < \infty) < 1$, that is the chain is transient. Conversely, if the chain is transient, we can take $V(x) = \mathbb{P}_x(T_A < +\infty)$.

Case (b). Let $A = \{x: \mathcal{L}V > 0\}$ and let $D_N = \{x: V(x) > N\}$ which by assumption has finite complement for all N. Then $\mathbb{P}_x(T_{D_N} < \infty) = 1$ for all $x \in E$. By optional stopping, for all $x \in E$,

$$V(x) \geqslant \mathbb{E}_x[V(X_{T_{D_N} \wedge T_A})] \geqslant \mathbb{E}_x[V(X_{T_{D_N}})\mathbb{I}_{T_{D_N} \leqslant T_A}] \geqslant N\mathbb{P}_x(T_{D_N} \leqslant T_A) \geqslant N\mathbb{P}_x(T_A = +\infty),$$

and taking $N \to \infty$ we obtain $\mathbb{P}_x(T_A = +\infty) = 0$, so the chain is recurrent. Now assume that the chain is recurrent and consider a finite set A and a decreasing sequence of sets $(B_N)_N$ with finite complements and such that $\cap_N B_N = \varnothing$ and let $V_N(x) = \mathbb{P}_x(T_{B_N} < T_A)$ which is an harmonic function on $E \backslash (A \cup D_N)$ such that $V_N = 1$ on B_N and $V_N = 0$ on A. By recurrence we have $V_N(x) \searrow \mathbb{P}_x(+\infty = T_A) = 0$ for all $x \in E$ and we can find a sequence $(N_k)_k$ such that $V(x) = \sum_k V_{N_k}(x) < \infty$ for every x (by a diagonal argument). This function satisfies the required conditions since $V(x) \geqslant N_k$ on D_{N_k} and if $x \in D_N$ we have $\mathcal{L}V_N(x) \leqslant 1 - 1 = 0$ so $\mathcal{L}V \leqslant 0$ on A^c.

Case (c). We let $A = \{\mathcal{L}V > -1\}$ and consider

$$V(x) \geqslant \lim_n \mathbb{E}_x[V(X_{n \wedge T_A}) + (n \wedge T_A)] \geqslant \mathbb{E}_x[T_A]$$

which implies then that the chain is positive recurrent. On the other hand if the chain is positive recurrent, then let $V(x) = \mathbb{E}_x(T_A)$ for an arbitrary finite set A and check that V satisfies the assumptions. □

3.1 Recurrence and transience of random walks

The random walk on $\mathbb{Z}^d$ has generator

$$\mathcal{L}f(x) = \frac{1}{2d} \sum_{y: y \sim x} [f(y) - f(x)]$$

where $y \sim x \Leftrightarrow |x - y| = 1$. The chain is irreducible. We want to show that this chain is recurrent only if $d \leqslant 2$. If f varies slowly then we have $\mathcal{L}f(x) \simeq \Delta f(x)$ and if we let $V(x) = |x|^{2\alpha}$ we have

$$\mathcal{L}V(x) \simeq (2\alpha)(d + 2\alpha - 2)|x|^{2\alpha - 2}$$

so for $d>2$ we can find $\alpha<0$ such that $\mathcal{L}V\leqslant 0$ outside a ball $B(0,r_0)$ (taking accounts errors in the above approximations) and V is decreasing at infinity so for any x we can arrange to have $V(x)<\inf_{B(0,r_0)}V$. When $d=1$ we need to choose $\alpha\in(0,1/2)$ to ensure $V(x)\to+\infty$ as $|x|\to+\infty$. For $d=2$ the right function to look for is of the form $V(x)=(\log|x|^2)^\alpha$ so that

$$\Delta V(x)\simeq\frac{4\alpha(\alpha-1)}{|x|^2}(\log|x|^2)^{\alpha-2}$$

and taking $\alpha\in(0,1)$ gives a function which satisfies the recurrence criterion.

3.2 Harris recurrence of sets in general state spaces

We consider now the situation where the state space is a general Polish space (and $\mathcal{E}$ the Borel σ-algebra). In this case the Forster–Lyapounouv conditions are not as tight. However the existence of certain superharmonic functions implies Harris recurrence.

Definition 13. *A set $A\in\mathcal{E}$ is Harris recurrent if $\mathbb{P}_x(T_A^+<\infty)=1$ for all $x\in A$. Is positive recurrent if $\mathbb{E}_x[T_A^+]<\infty$ for all $x\in A$.*

Proposition 14.

a) *If there exists a function $V\in\mathcal{F}_+(E)$ such that $\mathcal{L}V\leqslant 0$ on A^c and $T_{\{V>c\}}<\infty$ $\mathbb{P}_x$–a.s. for all $x\in E$ and $c\geqslant 0$ then the set A is Harris recurrent.*

b) *If there exists a function $V\in\mathcal{F}_+(E)$ such that $\mathcal{L}V\leqslant -1$ on A^c and $\pi V<\infty$ on A then the set A is positive recurrent.*

Proof. For (a) the proof goes as in the discrete setting. For (b) we deduce first that $\mathbb{E}_x[T_A]\leqslant V(x)$ and then by a one–step computation deduce that $\mathbb{E}_x[T_A^+]=\mathbb{E}_x[\mathbb{E}_{X_1}[T_A^+]]=(\pi V)(x)<\infty$. □

Example 15. (State space model on $\mathbb{R}^d$) Consider the following Markov chain on $\mathbb{R}^d$

$$X_{n+1}=X_n+b(X_n)+W_n$$

where $(W_n)_n$ are iid with mean zero and covariance $\mathrm{Cov}(W^i,W^j)=\delta_{ij}$. We consider the function $V(x)=|x|^2/\varepsilon$. Then

$$\varepsilon(\pi V)(x)=\mathbb{E}[|x+b(x)+W_1|^2]-|x|^2=2\langle x,b(x)\rangle+|b(x)|^2+d.$$

By choosing ε small enough we can see that the sufficient condition for positive recurrence holds for $A=B(0,r)$ and r sufficiently large.