Next friday's lecture: revision of the course. Please bring your questions/doubts.

Last lecture: Markov chains in discrete state space: recurrence/transience.

Intermezzo: computations of probabilities related to a Markov chain

The Markov property (also in the strong) version is a powerful tool to compute interesting quantities related to a Markov chain.

We will assume in this part to be on the canonical space of a Markov chain $(X_n)_{n\geqslant 0}$ on a general state space $(E,\mathscr{E})$ and we will denote as usual with $(\mathbb{P}_x)_{x\in E}$ the family of laws of the process starting at $x\in E$.

Take $A,B\in\mathscr{E}$ with $A\cap B\neq\emptyset$. We would like to compute the probability

$$\mathbb{P}_x(X \text{ reaches } A \text{ before } B).$$

Let T_A be the **hitting time** of the set A, i.e. the stopping time

$$T_A=\inf\{n\geqslant 0: X_n\in A\}$$

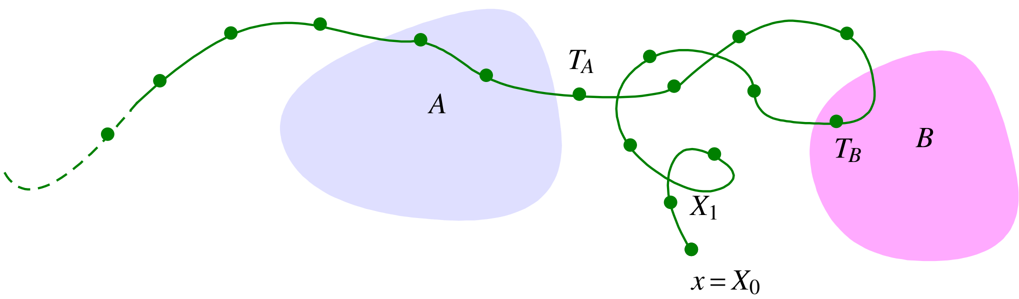
and similarly T_B, both can be $+\infty$ in general. The event of interest is $\{T_A<T_B\}$. The way we can study this probability for a given initial state $x\in E$ is to generalize the problem and try to solve it for all initial states $x\in E$ so we let

$$u(x)=\mathbb{P}_x(T_A<T_B), \qquad x\in E$$

and we look it as an unknown function $u:E\to[0,1]$. What we know about it?

- If $x\in A$ then $u(x)=1$;
- If $x\in B$ then $u(x)=0$;

If $x\in(A\cup B)^c$ then we can reason as follows. Under $\mathbb{P}_x$ we will have $T_A\geqslant 1$ and $T_B\geqslant 1$. Let's make a drawing:



Provided $x\in(A\cup B)^c$ after the first step, we will be in X_1 and therefore have probability $u(X_1)$ to reach A before B, so we guess that we should have the following equation

$$u(x)=\mathbb{E}_x[u(X_1)]=\int_E u(y)P(x,\mathrm{d}y)=(Pu)(x)$$

This was an heuristic explanation, we can prove it via the Markov property. Indeed note that

$$\mathbb{1}_{T_A<T_B} = \mathbb{1}_{T_A<T_B} \circ \theta_1$$

provided $X_0 \in (A\cup B)^c$. Recall that $\theta_1(\omega) = (\omega_1, \omega_2, \dots)$, i.e. the trajectory of the process without the initial point. Taking expectation wrt. $\mathbb{P}_x$ we get

$$u(x) = \mathbb{E}_x[\mathbb{1}_{T_A<T_B}] = \mathbb{E}_x[\mathbb{1}_{T_A<T_B} \circ \theta_1] = \mathbb{E}_x[\mathbb{E}_x[\mathbb{1}_{T_A<T_B} \circ \theta_1 | \mathscr{F}_1]]$$

$$\underset{\text{Markov}}{=} \mathbb{E}_x[\mathbb{E}_{X_1}[\mathbb{1}_{T_A<T_B}]] = \mathbb{E}_x[u(X_1)]$$

which is what we claimed. This is called a one-step computation. A very useful technique for Markov chains. Recall that the generator $\mathscr{L}$ is defined as $\mathscr{L} = P - \mathrm{Id}$, therefore we have that

$$\mathscr{L}u(x) = Pu(x) - u(x) = 0, \qquad x \in (A\cup B)^c.$$

Summarizing we have found that the function $u: E \to [0,1]$ is a solution of the problem:

$$\begin{cases} \mathscr{L}u(x) = 0 & x \in (A\cup B)^c; \\ u(x) = 0 & x \in B; \\ u(x) = 1 & x \in A. \end{cases} \tag{1}$$

In many cases we can actually find a unique solution to this linear equation, which usually boils down to solve some finite difference equation with boundary conditions. For example think about this problem for the simple random walk on $\mathbb{Z}$.

Note that the first condition implies that u is harmonic in $(A\cup B)^c$ for the operator $\mathscr{L}$.

In general the system (1) does not have unique solutions and therefore it is not enough to determine the probabilities u.

Let $v: E \to [0,1]$ be another solution of (1). Since v is harmonic in $(A\cup B)^c$ then we know that the process

$$v(X_n) = v(X_0) + M_n^v + \sum_{k=0}^{n-1} (\mathscr{L}v)(X_k)$$

is a martingale by the martingale problem solved by X. By computing this at $n\wedge T_A\wedge T_B$ we have that $X_{n\wedge T_A\wedge T_B} = X_n^{T_A\wedge T_B}$ and

$$v(X_n^{T_A\wedge T_B}) = v(X_0^{T_A\wedge T_B}) + M_n^{v, T_A\wedge T_B} + \sum_{k=0}^{n\wedge T_A\wedge T_B - 1} (\mathscr{L}v)(X_k) = v(X_0^{T_A\wedge T_B}) + M_n^{v, T_A\wedge T_B}$$

since $\sum_{k=0}^{n\wedge T_A\wedge T_B-1} (\mathscr{L}v)(X_k) = 0$ because $X_k \in (A\cup B)^c$ for all $k = 0, \dots, n\wedge T_A\wedge T_B - 1$. So we conclude that $v(X_n^{T_A\wedge T_B})$ is a martingale. Then

$$v(x) = \mathbb{E}_x[v(X_0^{T_A\wedge T_B})] = \mathbb{E}_x[v(X_n^{T_A\wedge T_B})]$$

$$= \mathbb{E}_x[v(X_n^{T_A\wedge T_B})\mathbb{1}_{T_A<T_B}] + \mathbb{E}_x[v(X_n^{T_A\wedge T_B})\mathbb{1}_{T_A=T_B=+\infty}] + \mathbb{E}_x[v(X_n^{T_A\wedge T_B})\mathbb{1}_{T_A>T_B}]$$

$$\geqslant \mathbb{E}_x[v(X_n^{T_A\wedge T_B})\mathbb{1}_{T_A<T_B}] + \mathbb{E}_x[v(X_n^{T_A\wedge T_B})\mathbb{1}_{T_A>T_B}]$$

Letting $n\to\infty$ this quantity converges to

$$v(x) \geqslant \mathbb{E}_x[v(X_{T_A})\mathbb{1}_{T_A<T_B}] + \mathbb{E}_x[v(X_{T_B})\mathbb{1}_{T_A>T_B}] = \mathbb{E}_x[\mathbb{1}_{T_A<T_B}] = u(x)$$

since $X_{T_A} \in A$ if $T_A < \infty$ and similarly for T_B. We conclude that

$$v(x) \geqslant u(x).$$

This holds actually for any solution v which is positive and bounded. This means that u is singled out among all the solutions to (1) as the smallest positive and bounded solution.

We can packaged these considerations in a small theorem.

Theorem. *Let $\mathscr{L}$ be the generator of a Markov process X. Then the problem (1) has a unique smallest positive and bounded solution given by*

$$u(x) = \mathbb{P}_x(T_A < T_B).$$

We can use the same approach for other quatities, not necessarily probabilities. For example in the setting above consider the quantity

$$u(x) = \mathbb{E}_x[T_A]$$

which is now a function $u: E \to \mathbb{R}_+$. The one-step analysis for this quantity goes as follows: if $X_0 \notin A$ then certainly $T_A > 1$ and therefore

$$T_A(\omega) = 1 + T_A(\theta_1(\omega))$$

which means that we have now for $x \in A^c$

$$u(x) = \mathbb{E}_x[T_A] = \mathbb{E}_x[1 + T_A \circ \theta_1] = 1 + \mathbb{E}_x[\mathbb{E}_{X_1}[T_A]] = 1 + \mathbb{E}_x[u(X_1)] = 1 + (Pu)(x)$$

moreover if $x \in A$ then $u(x) = 0$. We have again a linear system for u:

$$\begin{cases} \mathscr{L}u(x) = -1 & x \in A^c; \\ u(x) = 0 & x \in A. \end{cases} \tag{2}$$

and again we can prove that $u(x) = \mathbb{E}_x[T_A]$ is the smallest positive solution (not necessarily bounded) for this equation, indeed if $v: E \to \mathbb{R}_+$ is another solution then the process

$$M_n = v(X_n^{T_A}) + (n \wedge T_A)$$

is now a positive martingale and then

$$v(x) = \mathbb{E}_x[v(X_0^{T_A})] = \mathbb{E}_x[M_0] = \mathbb{E}_x[M_n] = \mathbb{E}_x[v(X_n^{T_A}) + (n \wedge T_A)] \geqslant \mathbb{E}_x[(n \wedge T_A)]$$

and taking $n \to \infty$ by monotone convergence one has $u(x) \leqslant v(x)$.

Note that $T_A = \sum_{k=0}^{T_A - 1} 1$. More generally we can introduce the quantity

$$u(x) = \mathbb{E}_x\left[\sum_{k=0}^{T_A - 1} g(X_k) + q(X_{T_A}) \mathbb{1}_{T_A < \infty}\right]$$

for given positive functions $g: E \to \mathbb{R}_+$ and $q: E \to \mathbb{R}_+$ and a set $A \in \mathscr{E}$. It is easy to see that this general case comprises the above two special cases for specific choices of g, q, A.

For this general form the equation is

$$\begin{cases} \mathscr{L}u(x) = -g(x) & x \in A^c; \\ u(x) = q(x) & x \in A. \end{cases} \tag{3}$$

Which is left an exercise to show. For any solution ν of this system the process

$$M_n = \nu(X_n^{T_A}) + \sum_{k=0}^{n-1} g(X_k^{T_A})$$

is a positive martingale and $\nu \geqslant u$.

Doob's *h*-transform

Let $h: E \to \mathbb{R}_{\geqslant 0}$ be a positive harmonic function such that $h(x_0) = 1$ for a given $x_0 \in E$. Then the process $M_n = h(X_n)$ is a positive martingale under $\mathbb{P}_{x_0}$ with $M_0 = 1$ and therefore $\mathbb{E}_{x_0}[M_n] = 1$ for all n. On $\mathscr{F}_n$ we can define the probability $\mathbb{Q}_n$ by

$$\mathbb{Q}_n(A) = \mathbb{E}_{x_0}[\mathbb{1}_A h(X_n)] \qquad A \in \mathscr{F}_n.$$

The family $(\mathbb{Q}_n)_{n\geqslant 0}$ is a consistent family of probabilities on $\cup_k \mathscr{F}_k$, i.e. $A \in \mathscr{F}_k$ then for any $n \geqslant k$ we have $A \in \mathscr{F}_n$

$$\mathbb{Q}_n(A) = \mathbb{E}_{x_0}[\mathbb{1}_A h(X_n)] = \mathbb{E}_{x_0}[\mathbb{1}_A \mathbb{E}_x[h(X_n)|\mathscr{F}_k]] = \mathbb{E}_{x_0}[\mathbb{1}_A h(X_k)] = \mathbb{Q}_k(A).$$

Then by Caratheodory extension theorem there exists a unique probability measure $\mathbb{Q}$ on $\mathscr{F}_\infty := \sigma(\mathscr{F}_k: k \geqslant 0)$ such that $\mathbb{Q}(A) = \mathbb{Q}_n(A)$ for $A \in \mathscr{F}_n$. Since we assume to be in the canonical space of the Markov chain we have $\mathscr{F}_\infty = \mathscr{F}$.

The measure $\mathbb{Q}$ defined in this way is called the Doob's h-transform of $\mathbb{P}$. We have

$$\frac{\mathrm{d}\mathbb{Q}|_{\mathscr{F}_n}}{\mathrm{d}\mathbb{P}_{x_0}|_{\mathscr{F}_n}} = \frac{\mathrm{d}\mathbb{Q}_n}{\mathrm{d}\mathbb{P}_{x_0}|_{\mathscr{F}_n}} = h(X_n), \qquad n \geqslant 0.$$

Let T_Z the hitting time of the zero set $Z = \{x \in E: h(x) = 0\}$ of h. Then

$$\mathbb{Q}(T_Z < \infty) = \lim_{n\to\infty} \mathbb{Q}(T_Z \leqslant n) = \lim_{n\to\infty} \mathbb{Q}_n(T_Z \leqslant n)$$

$$= \lim_{n\to\infty} \mathbb{E}_{x_0}[h(X_n) 1_{T_Z \leqslant n}] = \lim_{n\to\infty} \mathbb{E}_{x_0}[\mathbb{E}[h(X_n)|\mathscr{F}_{T_Z \wedge n}] 1_{T_Z \leqslant n}]$$

(by optional stopping)

$$= \lim_{n\to\infty} \mathbb{E}_{x_0}[h(X_{T_Z \wedge n}) 1_{T_Z \leqslant n}] = \lim_{n\to\infty} \mathbb{E}_{x_0}[h(X_{T_Z}) 1_{T_Z \leqslant n}] = 0$$

since $h(X_{T_Z}) = 0$! So $\mathbb{Q}(T_Z < \infty) = 0$ which means that under the measure $\mathbb{Q}$ the process never reaches Z. That could not be true under $\mathbb{P}$. Under the measure $\mathbb{Q}$ the way the process moves around is different! It tries to avoid Z.

Under the measure $\mathbb{Q}$ the process $(X_n)_{n\geqslant 0}$ is still Markov process (!!!) and actually the solution of a martingale problem with generator

$$\mathscr{L}^h f = h^{-1} \mathscr{L}(hf)$$

where $h^{-1}(x) = 1/h(x)$. Let's show it: Take $f = h^{-1} g$

$$M_n^f = f(X_n) - f(X_0) - \sum_{k=0}^{n-1} \mathscr{L}^h f(X_k) = h^{-1}(X_n) g(X_n) - h^{-1}(X_0) g(X_0) - \sum_{k=0}^{n-1} h^{-1}(X_k)(\mathscr{L}g)(X_k)$$

We need to prove to be a martingale under $\mathbb{Q}$. Take $A \in \mathscr{F}_n$,

$$\mathbb{E}_{\mathbb{Q}}[(M^f_{n+1} - M^f_n)\mathbb{1}_A] = \mathbb{E}_{x0}[h(X_{n+1})(M^f_{n+1} - M^f_n)\mathbb{1}_A]$$

$$= \mathbb{E}_{x0}[h(X_{n+1})(h^{-1}(X_{n+1})g(X_{n+1}) - h^{-1}(X_n)Pg(X_n))\mathbb{1}_A]$$

$$= \mathbb{E}_{x0}[(g(X_{n+1}) - h(X_{n+1})h^{-1}(X_n)Pg(X_n))\mathbb{1}_A]$$

$$= \mathbb{E}_{x0}[(\mathbb{E}[g(X_{n+1})|\mathscr{F}_n] - \mathbb{E}[h(X_{n+1})|\mathscr{F}_n]h^{-1}(X_n)Pg(X_n))\mathbb{1}_A]$$

$$= \mathbb{E}_{x0}[(Pg(X_n) - Pg(X_n))\mathbb{1}_A] = 0$$

which shows the martingale property.

Note that

$$\mathscr{L}^h f = h^{-1}(P - \mathrm{Id})(hf) = h^{-1}(P(hf) - hf) = h^{-1}P(hf) - f$$

which means also that the transition kernel of the Markov chain $(X_n)_{n\geqslant 0}$ under the measure $\mathbb{Q}$ is given by P^h with

$$(P^h f)(x) = (h^{-1}P(hf))(x) = \frac{1}{h(x)}\int_E h(y)f(y)P(x, \mathrm{d}y)$$

for all $x \in Z^c$. We can take as state space the set $E \setminus Z = Z^c$, i.e the set where $h(x) > 0$.