Lecture 23 · 6.7.2021 · 14:15–16:00 via Zoom

Discrete Markov chains

This week we are going to give a more detailed look at Markov chains in discrete state spaces, i.e. we take E to be a discrete set (maybe infinite) and $\mathscr{E} = \mathscr{P}(E)$.

This is a technically simpler setting where however many of the properties and the typical behaviour of Markov chains are already appeerent.

For a discrete E the transition kernel $P: E \times \mathscr{E} \to [0, 1]$ is equivalente to a *transition matrix* P: $E \times E \to [0, 1]$ which we denote with the same simbol and such that $P(x, y) = P(x, \{y\})$. It represents the probability to see the Markov chain jumping to y conditioned to be in x. We have $\sum_{y \in E} P(x, y) = 1$.

Let $(X_n)_{n \geqslant 0}$ be a Markov chain with state space E and denote with

For any $x \in E$ we denote

$$T_x = T_{\{x\}} = \inf \{n > 0 : X_n = x\}$$

the **return time** to the point x.

We want to study the general behaviour of the chain $(X_n)_{n \geqslant 0}$ for long times. In this context it is useful to make an additional assumption on the structure of the Markov chain. We will assume that we work in the canonical measure space $(\Omega, \mathscr{F}) = (E^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}})$ for the chain with $(\mathbb{P}_x)_{x \in E}$ the family of its laws with starting point $x \in E$. In this case $(X_n)_{n \geqslant 0}$ will be the canonical process in $(\Omega, \mathscr{F})$.

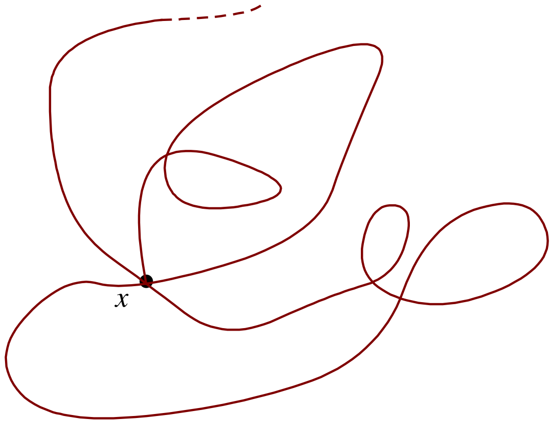
Definition. *We say that the chain $(X_n)_{n \geqslant 0}$ is **irreducible** if there exists a positive probability to go from any state $x \in E$ to any other state $y \in E$: i.e. for all $x, y \in E$, $x \neq y$ $\mathbb{P}_x(T_y < \infty) > 0$ in which case we write $x \to y$. Equivalently x is connected to y iff there exists $n = n(x, y) \in \mathbb{N}$ such that $P^{(n)}(x, y) > 0$.*

Definition 1. *Given an irreducible chain we say that it is*

- ***transient** if $\mathbb{P}_x(T_x < \infty) < 1$ for all $x \in E$; (i.e. there is positive probability not to ever come back to some state x)*
- ***recurrent** if $\mathbb{P}_x(T_x < \infty) = 1$ for all $x \in E$; (i.e. we always come back to the starting point)*
- ***positive recurrent** if $\mathbb{E}_x[T_x] < \infty$ for all $x \in E$; (i.e. we come back to the starting point after finite time in average)*

Remark. We can establish an equivalence relations among the states of a genera discrete chain by saying that $x \sim y$ (x and y are connected) if $x \to y$ and $y \to x$ (in the sense above, i.e. $\mathbb{P}_x(T_y < \infty) > 0$ and viceversa). This is an equivalence relation on E (verify this!) which induces a partition of the state space E in equivalence classes. It is not difficult to show that transience and recurrence are properties of the equivalence classes and that equivalence classes are either transient or recurrent (not other possibilities are given). If the chain is finite (i.e. E is finite) then at least of the classes is recurrent. We will focus only to the chains that are irreducible which corresponds to the case where we have only one class.

We focus now on criteria which tells us whether an irreducible chain is transient or recurrent.



The chain forget its past every time it comes back to x. It starts afresh, like it was the first time.

Theorem. *For a (non-necessarily irreducible) chain and for all $x \in E$ it holds*

$$\mathbb{P}_x(X_n = x \text{ infinitely often}) = 0 \Leftrightarrow \mathbb{E}_x\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = x}\right] < +\infty \Leftrightarrow \mathbb{P}_x(T_x < \infty) < 1,$$

$$\mathbb{P}_x(X_n = x \text{ infinitely often}) = 1 \Leftrightarrow \mathbb{E}_x\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = x}\right] = +\infty \Leftrightarrow \mathbb{P}_x(T_x < \infty) = 1.$$

Moreover if the chain is irreducible then for any $x, y \in E$

$$\mathbb{E}_x\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = x}\right] < \infty \Leftrightarrow \mathbb{E}_y\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = y}\right] < \infty.$$

As a consequence an irreducible Markov chain is either recurrent or transient that that it is decided by whether

$$\mathbb{E}_x\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = x}\right] = \sum_{n \geqslant 1} P^{(n)}(x, x)$$

is finite or not for some state (and then all of them).

Proof. Let $\lambda > 0$. Then the quantity $\sum_{n \geqslant 1} e^{-\lambda n} \mathbb{1}_{X_n = x}$ is always well defined.

We either come back to x at least once, at time T_x or never, i.e. $T_x = +\infty$, in both cases we have

$$\sum_{n \geqslant 1} e^{-\lambda n} \mathbb{1}_{X_n = x} = e^{-\lambda T_x} + \sum_{n > T_x} e^{-\lambda n} \mathbb{1}_{X_n = x} = e^{-\lambda T_x} + \sum_{n \geqslant 1} e^{-\lambda (n + T_x)} \mathbb{1}_{X_{T_x + n} = x}$$

$$= e^{-\lambda T_x} + e^{-\lambda T_x} \sum_{n \geqslant 1} e^{-\lambda n} (\mathbb{1}_{X_n = x} \circ \theta_{T_x})$$

By strong Markov property we have:

$$
\begin{aligned}
\mathbb{E}_x\left[\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{1}_{X_n=x}\right] &= \mathbb{E}_x[e^{-\lambda T_x}] + \mathbb{E}_x\left[e^{-\lambda T_x}\sum_{n\geqslant 1} e^{-\lambda n}(\mathbb{1}_{X_n=x}\circ\theta_{T_x})\right] \\
&= \mathbb{E}_x[e^{-\lambda T_x}] + \mathbb{E}_x\left[\mathbb{E}_x\left[e^{-\lambda T_x}\sum_{n\geqslant 1} e^{-\lambda n}(\mathbb{1}_{X_n=x}\circ\theta_{T_x})\middle|\mathscr{F}_{T_x}\right]\right] \\
&= \mathbb{E}_x[e^{-\lambda T_x}] + \mathbb{E}_x\left[e^{-\lambda T_x}\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{E}_x[(\mathbb{1}_{X_n=x}\circ\theta_{T_x})|\mathscr{F}_{T_x}]\right] \\
&= \mathbb{E}_x[e^{-\lambda T_x}] + \mathbb{E}_x\left[e^{-\lambda T_x}\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{E}_{X_{T_x}}[\mathbb{1}_{X_n=x}]\right] \\
&= \mathbb{E}_x[e^{-\lambda T_x}] + \mathbb{E}_x[e^{-\lambda T_x}]\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{E}_x[\mathbb{1}_{X_n=x}] \\
&= \mathbb{E}_x[e^{-\lambda T_x}]\left(1+\mathbb{E}_x\left[\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{1}_{X_n=x}\right]\right)
\end{aligned}
$$

where we used also that $X_{T_x}=x$ by definition of T_x if $T_x<\infty$ (In case $T_x=+\infty$ all the expression is 0 due to $e^{-\lambda T_x}=0$). We have now that

$$
\mathbb{E}_x\left[\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{1}_{X_n=x}\right] = \frac{\mathbb{E}_x[e^{-\lambda T_x}]}{1-\mathbb{E}_x[e^{-\lambda T_x}]}
$$

since $\mathbb{E}_x[e^{-\lambda T_x}]\leqslant e^{-\lambda}<1$ due to the fact that $T_x\geqslant 1$. Now we want to take $\lambda\to 0$, assume that $\mathbb{P}_x(T_x<\infty)<1$, then by monotone convergence

$$
\mathbb{E}_x\left[\sum_{n\geqslant 1}\mathbb{1}_{X_n=x}\right] = \lim_{\lambda\downarrow 0}\mathbb{E}_x\left[\sum_{n\geqslant 1} e^{-\lambda n}\mathbb{1}_{X_n=x}\right] = \lim_{\lambda\downarrow 0}\frac{\mathbb{E}_x[e^{-\lambda T_x}]}{1-\mathbb{E}_x[e^{-\lambda T_x}]} = \frac{\mathbb{P}_x(T_x<\infty)}{1-\mathbb{P}_x(T_x<\infty)}<\infty
$$

and therefore $\sum_{n\geqslant 1}\mathbb{1}_{X_n=x}<\infty$ $\mathbb{P}_x$-a.s. which means that under $\mathbb{P}_x$ the set $\{n\geqslant 1 : X_n=x\}$ is almost surely finite, i.e. $\mathbb{P}_x(X_n=x \text{ infinitely often})=0$.

On the other hand if $\mathbb{P}_x(X_n=x \text{ infinitely often})>0$ then certainly $\mathbb{E}_x[\sum_{n\geqslant 1}\mathbb{1}_{X_n=x}]=+\infty$ which implies that $\mathbb{P}_x(T_x<\infty)=1$.

We need then to show that if $\mathbb{P}_x(T_x<\infty)=1$ then $\mathbb{P}_x(X_n=x \text{ infinitely often})=1$. This is done using again the strong Markov property as follows. Let

$$
A=\{X_n=x \text{ infinitely often}\}
$$

so

$$
A^c=\{X_n=x \text{ only finitely many times}\}=\cup_{L\geqslant 0}\{X_n=x\ L \text{ times}\}
$$

Let now $B_L=\{X_n=x\ L \text{ times}\}$, then for all $L\geqslant 1$

$$
B_L=\{T_x<\infty\}\cap\{X_n\circ\theta_{T_x}=x\ L-1 \text{ times}\}
$$

$$
\mathbb{1}_{B_L}=\mathbb{1}_{T_x<\infty}\cdot\mathbb{1}_{B_{L-1}}\circ\theta_{T_x}
$$

where we understand that B_0 is the event that I never come back i.e. $B_0=\{T_x=+\infty\}$.

Taking average and using the strong Markov property we have

$$\mathbb{P}_x(B_L) = \mathbb{E}_x[\mathbb{1}_{B_L}] = \mathbb{E}_x[\mathbb{E}_x[\mathbb{1}_{T_x<\infty} \cdot \mathbb{1}_{B_{L-1}} \circ \theta_{T_x} | \mathscr{F}_{T_x}]]$$

$$= \mathbb{E}_x[\mathbb{1}_{T_x<\infty} \mathbb{E}_x[\mathbb{1}_{B_{L-1}} \circ \theta_{T_x} | \mathscr{F}_{T_x}]] = \mathbb{E}_x[\mathbb{1}_{T_x<\infty} \mathbb{E}_{X_{T_x}}[\mathbb{1}_{B_{L-1}}]] = \mathbb{E}_x[\mathbb{1}_{T_x<\infty}] \mathbb{E}_x[\mathbb{1}_{B_{L-1}}]$$

$$= \mathbb{P}_x(T_x < \infty) \mathbb{P}_x(B_{L-1}).$$

This shows that

$$\mathbb{P}_x(B_L) = (\mathbb{P}_x(T_x < \infty))^L \mathbb{P}_x(T_x = +\infty) = 0,$$

since $\mathbb{P}_x(T_x < \infty) = 1$ which implies $\mathbb{P}_x(T_x = \infty) = 0$. This proves that

$$\mathbb{P}_x(A^c) = \sum_{L \geqslant 0} \mathbb{P}_x(B_L) = 0.$$

From these chain of implications we can deduce those in the statement of the theorem. For the last part under the assumption of irreducibility we need to show that

$$\mathbb{E}_x\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = x}\right] < \infty \Rightarrow \mathbb{E}_y\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = y}\right] < \infty,$$

for any two states $x, y \in E$. By irreducibility there exist n_0 and n_1 such that $P^{(n_0)}(x, y) > 0$ and equally that $P^{(n_1)}(y,x) > 0$. Therefore we also have that

$$P^{(n_0+m+n_1)}(x,x) = \sum_{z,z' \in E} P^{(n_0)}(x,z) P^{(m)}(z,z') P^{(n_1)}(z',x)$$

$$\geqslant P^{(n_0)}(x,y) P^{(m)}(y,y) P^{(n_1)}(y,x).$$

Now note that (check)

$$\mathbb{E}_x\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = x}\right] = \sum_{n \geqslant 1} \mathbb{E}_x[\mathbb{1}_{X_n = x}] = \sum_{n \geqslant 1} P^{(n)}(x,x)$$

$$\geqslant \sum_{m \geqslant 1} P^{(n_0+m+n_1)}(x,x) \geqslant P^{(n_0)}(x,y) \sum_{m \geqslant 1} P^{(m)}(y,y) P^{(n_1)}(y,x)$$

$$= P^{(n_0)}(x,y) P^{(n_1)}(y,x) \mathbb{E}_y\left[\sum_{n \geqslant 1} \mathbb{1}_{X_n = y}\right]$$

which implies the claim since $P^{(n_0)}(x,y) P^{(n_1)}(y,x) > 0$. □