Lecture 25 · 13.7.2021 · 14:15–16:00 via Zoom

This Friday's lecture: revision of the course. Please bring your questions/doubts.

Doob's h-transform (applications)

Recall what we have seen last week: given a markov chain $(\mathbb{P}_x)_{x\in E}$ on a canonical space with state space E and transition kernel P, a positive harmonic function $h: E\to\mathbb{R}_{\geqslant 0}$ and a point $x_0\in E$ such that $h(x_0)=1$. We can construct a new probability measure $\mathbb{Q}$ such that

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

The measure $\mathbb{Q}$ is the Doob's h-transform of $\mathbb{P}$ with h. Under $\mathbb{Q}$ the canonical process $(X_n)_{n\geqslant 0}$ is a Markov chain with transition kernel

$$P^h f = h^{-1}(Phf)$$

on the state space $Z^c=\{x\in E: h(x)>0\}\subseteq E$.

Let's now see how to use this construction to solve certain conditioning problem for Markov chains. Let us keep the setting as above and let $A\in\mathscr{E}$ and assume that $\mathbb{P}_{x_0}(T_A=+\infty)>0$. Then let

$$h(x)=\frac{\mathbb{P}_x(T_A=+\infty)}{\mathbb{P}_{x_0}(T_A=+\infty)}, \qquad x\in E.$$

By construction this is a positive bounded harmonic function in A^c, i.e. $Ph(x)=h(x)$ for all $x\in A^c$ with $h(x_0)=1$ and $h(x)=0$ on A. In particular the process $(h(X_n^{T_A}))_{n\geqslant 0}$ is a martingale with average 1 under $\mathbb{P}_{x_0}$ (use the martingale problem of X to check it). We can the proceed to construct the measure $\mathbb{Q}$ as we did last week (with a small change due to stopping time), i.e.

$$\mathbb{Q}(B)=\mathbb{E}_{x_0}[h(X_n^{T_A})\mathbb{1}_B], \qquad B\in\mathscr{F}_n.$$

What represents this measure $\mathbb{Q}$?

For any $B\in\mathscr{F}_n$ (for some $n\geqslant 0$) we have

$$\mathbb{P}_{x_0}(B|T_A=+\infty)=\frac{\mathbb{P}_{x_0}(B,T_A=+\infty)}{\mathbb{P}_{x_0}(T_A=+\infty)}=\frac{\mathbb{E}_{x_0}[\mathbb{1}_B\mathbb{E}_{x_0}[\mathbb{1}_{T_A=+\infty}|\mathscr{F}_n]]}{\mathbb{P}_{x_0}(T_A=+\infty)}$$

Now we use the Markov property to write

$$\mathbb{E}_{x_0}[\mathbb{1}_{T_A=+\infty}|\mathscr{F}_n]=\mathbb{E}_{x_0}[(\mathbb{1}_{T_A=+\infty}\circ\theta_n)\mathbb{1}_{T_A\geqslant n}|\mathscr{F}_n]=\mathbb{E}_{x_0}[(\mathbb{1}_{T_A=+\infty}\circ\theta_n)|\mathscr{F}_n]\mathbb{1}_{T_A\geqslant n}$$

$$=\mathbb{P}_{X_n}(T_A=+\infty)\mathbb{1}_{T_A\geqslant n}=\mathbb{P}_{x_0}(T_A=+\infty)h(X_n)\mathbb{1}_{T_A\geqslant n}=\mathbb{P}_{x_0}(T_A=+\infty)h(X_n^{T_A})$$

Therefore we have

$$\mathbb{P}_{x_0}(B|T_A=+\infty)=\mathbb{E}_{x_0}[\mathbb{1}_B h(X_n^{T_A})]=\mathbb{Q}(B)$$

which tells us that the measure $\mathbb{Q}$ gives the probabilities for X conditioned never to touch the set A, i.e. on the event $\{T_A=+\infty\}$.

Remark. These considerations can be generalized to more general events by allowing the harmonic function to depend explicitly on time, i.e. taking $h(n,X_n)$. In this case however the h-transformed chain will not be anymore time-homogeneous.

Invariant measures for irreducible chains

One important problem in the theory of Markov chains is their behaviour for long times. What happens if we wait a long time and then look at the Markov chain? Does it settle down in an "equilibrium" situation. Think about mixing a fluid, or mixing a deck of cards, or in general looking at the result of very many small random actions.

The idea of "equilibrium" for a homogenenous Markov chain is enconded in the concept of invariant measure.

Definition. *A positive measure μ (not necessarily a probability measure) on the state space* $(E, \mathscr{E})$ *is an* ***invariant measure*** *for the Markov chain with transition kernel P iff*

$$\mu P = \mu.$$

If μ is an invariant probability measure and we start the Markov chain with initial law μ, i.e. $X_0 \sim \mu$ then we have that

$$\mathbb{P}(X_n \in A) = \mu P^{(n)} \mathbb{1}_A = \mu \underbrace{P \cdots P}_{n} \mathbb{1}_A = \mu(A)$$

so in particular the law of X_n does not depends on A, and more generally we also have that

$$\mathbb{E}[F \circ \theta_k] = \mathbb{E}[F], \qquad k \geqslant 0$$

for any bounded measurable $F: \Omega \to \mathbb{R}$. In particular if $f: E^n \to \mathbb{R}$ bounded then

$$\mathbb{E}[f(X_k, X_{k+1}, \ldots, X_{k+n})] = \mathbb{E}[f(X_0, X_1, \ldots, X_n)]$$

for all $k \geqslant 0$.

The law of the Markov chain is invariant under time shift. This is the reason why the probability μ is called invariant, because it give rise to a Markov chain which is invariant in law under time translations.

Now a natural question to ask is: does the equation

$$\mu P = \mu \tag{1}$$

has solutions (in the space of positive measures and in the space of probabilities)? And if yes, does it have a unique solution?

Both questions are non-trivial in general. There could be no solutions, no probability measures solutions, a unique solution which is not a probability, and a unique probability.

Exercise: construct a chain which has no invariant measures.

We will address these questions in the case of discrete Markov chains, i.e. E is a discrete space.

Remark. Note that if (1) has a solution μ then $\lambda \mu$ for $\lambda > 0$ is also a solution because the equation is linear. Also if μ, ν are two solutions then $\mu + \nu$ is also a solution and the set of invariant probabilities is a convex set.

Theorem. *If a discrete chain is irreducible, then*

- *any invariant measure ρ is everywhere strictly positive, i.e.* $\rho(y) > 0$ *for all* $y \in E$.

- *any two invariant measures differ by a multiplicative constants.*
- *there exists at most one invariant probability.*

Proof. Let ρ be a non-trivial invariant measure, then there must be $x \in E$ such that $\rho(x) > 0$. By irreducibility for any $y \in E$ there exists $n > 0$ such that $P^{(n)}(x,y) > 0$, now we also have by invariance

$$\rho(y) = (\rho P)(y) = \cdots = (\rho \underbrace{P \cdots P}_{n})(y) = (\rho P^{(n)})(y) = \sum_{z \in E} \rho(z) P^{(n)}(z,y) \geqslant \rho(x) P^{(n)}(x,y) > 0$$

so we conclude that $\rho(y) > 0$ for all $y \in E$. Any non-trivial invariant measure must be strictly positive everywhere.

[Let's complete it on Friday]

□

Example. Take the simple random walk on $\mathbb{Z}$, with $P(x,y) = \frac{1}{2}$ with $|x-y| = 1$ and =0 otherwise. Then the measure $\rho(x) = 1$ is an invariant measure and the chain is irreducible. Therefore there cannot be any invariant probability measures. Any invariant measure is a constant measure.

If the chain is recurrent then for any $x \in E$ we can define a measure ν^x on E by the formula

$$\nu^x(y) := \mathbb{E}_x\left[\sum_{n=1}^{S_x} \mathbb{1}_{X_n = y}\right], \qquad y \in E$$

where $S_x = \inf\{n \geqslant 1 : X_n = x\}$, the return time to x.

Theorem. *If the chain is irreducible and recurrent then ν^x is an invariant measure for any $x \in E$.*

Proof. We need to show that $\nu^x P = \nu^x$: for all $y \in E$

$$\nu^x(y) = \mathbb{E}_x\left[\sum_{n=1}^{S_x} \mathbb{1}_{X_n = y}\right] = \sum_{z \in E} \mathbb{E}_x\left[\sum_{n=1}^{S_x} \mathbb{1}_{X_n = y, X_{n-1} = z}\right]$$

$$= \sum_{z \in E} \mathbb{E}_x\left[\sum_{n=1}^{\infty} \mathbb{1}_{S_x \geqslant n} \mathbb{E}[\mathbb{1}_{X_n = y, X_{n-1} = z} | \mathscr{F}_{n-1}]\right]$$

$$= \sum_{z \in E} \mathbb{E}_x\left[\sum_{n=1}^{\infty} \mathbb{1}_{S_x \geqslant n} \mathbb{1}_{X_{n-1} = z}\right] P(z,y)$$

$$= \sum_{z \in E} \mathbb{E}_x\left[\sum_{n=0}^{\infty} \mathbb{1}_{S_x - 1 \geqslant n} \mathbb{1}_{X_n = z}\right] P(z,y) = \sum_{z \in E} \mathbb{E}_x\left[\sum_{n=0}^{S_x - 1} \mathbb{1}_{X_n = z}\right] P(z,y)$$

$$= \sum_{z \in E} \mathbb{E}_x\left[\sum_{n=1}^{S_x} \mathbb{1}_{X_n = z}\right] P(z,y)$$

$$= \sum_{z \in E} \nu^x(z) P(z,y) = (\nu^x P)(y)$$

Where we use the Markov property and the fact that $X_0 = X_{S_x} = x$ since by recurrence $S_x < \infty$ a.s. We proved the claim. □

By irreducibility we must have for any $x, y \in E$

$$\nu^y = C_{x,y} \nu^x$$

for some constant $C_{x,y}$. Note that for all $y \in E$

$$\nu^y(E) = \mathbb{E}_y\left[\sum_{n=1}^{S_y} \mathbb{1}_{X_n \in E}\right] = \mathbb{E}_y[S_y], \qquad \nu^y(y) = 1,$$

therefore if the chain is positive recurrent, that is if $\mathbb{E}_x[S_x] < \infty$ for all $x \in E$ then

$$C_{x,y} = \frac{\nu^y(E)}{\nu^x(E)} = \frac{\mathbb{E}_y[S_y]}{\mathbb{E}_x[S_x]}$$

and we can define a probability measure

$$\pi^x(z) = \frac{\nu^x(z)}{\nu^x(E)} = \frac{\nu^x(z)}{\mathbb{E}_x[S_x]}$$

and note that by irreducibility $\pi^x = \pi^y = \pi$ for all $x, y \in E$ which gives that

$$\pi(z) = \pi^z(z) = \frac{\nu^z(z)}{\mathbb{E}_z[S_z]} = \frac{1}{\mathbb{E}_z[S_z]}$$

Corollary. *If the chain is positive recurrent and irreducible then the probability measure*

$$\pi(x) = \frac{1}{\mathbb{E}_x[S_x]}, \qquad x \in E,$$

is the only invariant probability measure of the chain.

[I will take 30 min on friday to close the open points]