Lecture 22 · 2.7.2021 · 10:15–12:00 via Zoom

Markov chains (continued)

In the last lecture we saw that

- the law of a Markov chain $(Y_n)_{n\geqslant 0}$ on a state space $(E,\mathscr{E})$ (i.e. a stochastic process satisfying the Markov property and with values in $(E,\mathscr{E})$) is characterized by giving its *initial law* $\mathrm{Law}(Y_0)\in\Pi(E,\mathscr{E})$ and its transition kernel (or transition probability) $P: E\to\Pi(E,\mathscr{E})\approx E\times\mathscr{E}\to[0,1]$

- any Markov chain can be canonically realised on the space $(E^{\mathbb{N}},\mathscr{E}^{\otimes\mathbb{N}})$ on which we also dispose of the shift operator $\theta_n: E^{\mathbb{N}}\to E^{\mathbb{N}}$ defined as $(\theta_n(\omega))_k=\omega_{n+k}$ and of the canonical process $X_n(\omega)=\omega_n$ so that $X_k\circ\theta_n=X_{k+n}$.

- for any $x\in E$ we can construct on $(E^{\mathbb{N}},\mathscr{E}^{\otimes\mathbb{N}})$ the probability $\mathbb{P}_x$ which is the law of the Markov chain with initial condition $X_0=x$ $\mathbb{P}_x$-a.s. Note that $\mathbb{P}_\bullet: E\to\Pi(E^{\mathbb{N}},\mathscr{E}^{\otimes\mathbb{N}})$.

Example. (Random walk on $\mathbb{R}^n$). Let $E=\mathbb{R}^n$ with the Borel σ-algebra $\mathscr{E}=\mathscr{B}(\mathbb{R}^n)$. For every $x\in E$, consider an homogeneous Markov chain $(X_n)_{n\geqslant 0}$ defined as $X_0=x$ and

$$X_{n+1}=X_n+Z_n$$

with $(Z_n)_{n\geqslant 1}$ an i.i.d. sequence with values in E and such that $\rho=\mathrm{Law}(Z_1)\in\Pi(\mathbb{R}^n)$. In this case the transition kernel P is given by

$$P(x,\mathrm{d}y)=(\rho*\delta_x)(\mathrm{d}y)$$

where $*$ denotes convolution of the two measures ρ and δ_x on E. More explicitly for any bounded measurable function $f: E\to\mathbb{R}$ we have

$$Pf(x)=\int_E f(x+z)\rho(\mathrm{d}z),\qquad x\in\mathbb{R}^n.$$

In this case

$$P^{(n)}(x,\mathrm{d}y)=\underbrace{P\cdots P}_{n\text{ times}}(x,\mathrm{d}y)=(\rho^{*n}*\delta_x)(\mathrm{d}y).$$

Martingale problems

First we look at an important relation of Markov processes and martingale theory.

Let $f: E\to\mathbb{R}$ be a bounded measurable function and $(X_n)_{n\geqslant 0}$ a Markov process with transition kernel P. Then by the Markov property:

$$\mathbb{E}[f(X_{n+1})|\mathscr{F}_n]=Pf(X_n)$$

which means that the new process $(f(X_n))_{n\geqslant 0}$ has Doob's decomposition given by

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

where $\mathscr{L}$ is a linear operator on the space of bounded measurable functions on E given by

$$\mathscr{L} f = P f - f,$$

and $(M_n)_{n \geqslant 0}$ is a martingale (explicitly $\Delta M_n = f(X_n) - \mathbb{E}[f(X_n)|\mathscr{F}_{n-1}]$, but this is rarely needed). We learn that for function of Markov processes the predictable component in Doob's decomposition has a very explicit adn simple form.

The linear operator $\mathscr{L}$ is called the *generator* of the process. The observation above motivates us to give the following definition.

Definition. *Let $\mathscr{L}$ be a linear operator on the bounded measurable functions on E to $\mathbb{R}$. An adapted process $(X_n)_{n \geqslant 0}$ with values in $(E, \mathscr{E})$ satisfies the* ***martingale problem*** *with respect to $\mathscr{L}$, to the filtration $(\mathscr{F}_n)_{n \geqslant 0}$ and with initial law $\nu \in \Pi(E, \mathscr{E})$ iff*

$$\mathrm{Law}(X_0) = \nu$$

and for any bounded measurable $f: E \to \mathbb{R}$ we have that the process $(M_n^f)_{n \geqslant 0}$

$$M_n^f := f(X_n) - f(X_0) - \sum_{k=0}^{n-1} \mathscr{L} f(X_k), \qquad n \geqslant 0,$$

is a martingale wrt. the filtration $(\mathscr{F}_n)_{n \geqslant 0}$.

Consider now the measure space $(\Omega, \mathscr{F})$ given by the canonical space $(E^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}})$ with canonical process $(X_n)_{n \geqslant 0}$.

Theorem. *On the space $(\Omega, \mathscr{F}) = (E^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}})$ a family of laws $(\mathbb{P}_x)_{x \in E}$ for which $\mathbb{P}_x(X_0 = x) = 1$ for all $x \in E$ is an homogeneous Markov process with transition kernel P iff it solves the martingale problem for a generator $\mathscr{L}$. In this case we have*

$$\mathscr{L} = P - \mathrm{Id}.$$

Proof. We have seen that a Markov process solves the martingale problem. So it remains to show that if the canonical process solve the martingale problem for an operator $\mathscr{L}$ then it must a Markov process with transition kernel $P = \mathrm{Id} + \mathscr{L}$ (btw, which is a strong constraint on $\mathscr{L}$). To prove the markov property we have to consider $\mathbb{E}_x[f(X_{n+1})|\mathscr{F}_n]$ for a bounded measurable $f: E \to \mathbb{R}$. Since X solve the martingale problem we have

$$f(X_{n+1}) - f(X_n) = M_{n+1}^f - M_n^f + \mathscr{L} f(X_n)$$

where M^f is a martingale. Taking cond. exp. we get

$$\mathbb{E}_x[f(X_{n+1})|\mathscr{F}_n] = \underbrace{\mathbb{E}_x[f(X_n)|\mathscr{F}_n]}_{f(X_n)} + \mathbb{E}_x\underbrace{[M_{n+1}^f - M_n^f|\mathscr{F}_n]}_{=0} + \underbrace{\mathbb{E}_x[\mathscr{L} f(X_n)|\mathscr{F}_n]}_{\mathscr{L} f(X_n)} = T f(X_n)$$

with $T f := f + \mathscr{L} f$. This holds for any $\mathbb{P}_x$, i.e. any $x \in E$. In particular taking $n = 0$ we have $\mathbb{P}_x$-a.s. that

$$T f(x) = T f(X_0) = \mathbb{E}_x[f(X_1)|\mathscr{F}_0] = \mathbb{E}_x[f(X_1)]$$

because X_0 is $\mathbb{P}_x$-a.s. constant and therefore if $A \in \mathscr{F}_0 = \sigma(X_0)$ then $\mathbb{P}_x(A) \in \{0,1\}$, which implies that $\mathbb{E}_x[f(X_1)|\mathscr{F}_0] = \mathbb{E}_x[f(X_1)]$ $\mathbb{P}_x - a.s.$. Therefore we deduce that for all $x \in E$ we have

$$Tf(x) = \mathbb{E}_x[f(X_1)] = \int f(y)(\mathbb{P}_x \circ (X_1)^{-1})(\mathrm{d}y)$$

which in particular shows that T is the probability kernel given by $P = \mathbb{P}_x \circ (X_1)^{-1} = \mathrm{Id} + \mathscr{L}$. □

Corollary. *If $(Y_n)_{n\geqslant 0}$ and $(Z_n)_{n\geqslant 0}$ are two processes on a general probability space $(\Omega, \mathscr{F}, \mathbb{P})$ which solve the martingale problem for the same generator $\mathscr{L}$ and such that*

$$\mathrm{Law}_{\mathbb{P}}(Y_0) = \mathrm{Law}_{\mathbb{P}}(Z_0)$$

then they are both Markov processes with the same transtition kernel and the same initial law and as consequence they have the same law, that is

$$\mathrm{Law}_{\mathbb{P}}((Y_n)_{n\geqslant 0}) = \mathrm{Law}_{\mathbb{P}}((Z_n)_{n\geqslant 0}) \in \Pi(E^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}}).$$

Proof. As said. □

So solutions to martingale problems are unique in law (provided the initial conditions agree)

Remark. Martingale problems are a method to define dynamics of stochastic processes, much like in the deterministic setting one would use finite difference equations. Existence and uniqueness of solutions to martingale problems reduce in discrete time to problems of Markov processes. All these constructions have powerful analogs also in continuous time.

Example. For the random walk on $\mathbb{R}^n$ introduced above we have that for all $x \in \mathbb{R}^n$ we have

$$\mathscr{L}f(x) = Pf(x) - f(x) = \int_{\mathbb{R}^n} f(x+z)\rho(\mathrm{d}z) - f(x) = \int_{\mathbb{R}^n} [f(x+z) - f(x)]\rho(\mathrm{d}z)$$

Now take $n = 1$, i.e. $E = \mathbb{R}$ and fix $\varepsilon > 0$ and small and take

$$\rho(\mathrm{d}y) = \frac{1}{2}\delta_{+\varepsilon} + \frac{1}{2}\delta_{-\varepsilon}.$$

Then

$$\mathscr{L}f(x) = \frac{1}{2}[f(x+\varepsilon) - f(x)] + \frac{1}{2}[f(x-\varepsilon) - f(x)]$$

$$= \frac{1}{2}[f(x+\varepsilon) + f(x-\varepsilon) - 2f(x)]$$

by Taylor expansion, assuming f is sufficiently smooth:

$$\mathscr{L}f(x) = \frac{1}{2}\left[f'(x)\varepsilon + \frac{1}{2}f''(x)\varepsilon^2 + O(\varepsilon^3) - f'(x)\varepsilon + \frac{1}{2}f''(x)\varepsilon^2 + O(\varepsilon^3)\right]$$

$$= \varepsilon^2 \frac{1}{2} f''(x) + O(\varepsilon^3)$$

which shows that for $\varepsilon \to 0$ the operator $\varepsilon^{-2}\mathscr{L}$ approximate the standard Laplacian on $\mathbb{R}$ (i.e. the second derivative.)

Note that if we consider the random walk $(X_n)_{n\geqslant 0}$ with this generator which depends on ε and define a new process

$$Y_n := X_{n\varepsilon^{-2}}, \qquad n\geqslant 0$$

then we have that for any bounded function (and working a bit formally)

$$\begin{aligned} f(Y_n) = f(X_{n\varepsilon^{-2}}) &= f(X_0) + M^f_{n\varepsilon^{-2}} + \sum_{k=0}^{n\varepsilon^{-2}} \mathscr{L}f(X_k) \\ &= f(Y_0) + M^f_{n\varepsilon^{-2}} + \varepsilon^2 \sum_{k=0}^{n\varepsilon^{-2}} \varepsilon^{-2}\mathscr{L}f(Y_{\varepsilon^2 k}) \end{aligned}$$

and just handwaving a little one could guess that as $\varepsilon\to 0$ the process Y converges to a process in continuous time $(Y_t)_{t\geqslant 0}$ such that

$$f(Y_t) = f(Y_0) + \hat{M}^f_t + \frac{1}{2}\int_0^t f''(Y_s)\mathrm{d}s$$

for any $f\in C^2(\mathbb{R})$ and for some martingale $(\hat{M}^f_t)_{t\geqslant 0}$ in continuous time. This is very heuristic here but it can be made rigorous (in Stochastic Analysis). This is the point of introducing martingale problems and the solution of the continuous time maringale problem with generator given by $\frac{1}{2}f''$ is what we will call the Brownian motion.

Let us not consider a martingale problem with generator $\mathscr{L}$ and special classes of fuctions:

Definition. *A function f is **harmonic** wrt. $\mathscr{L}$ if*

$$\mathscr{L}f(x) = 0, \qquad x\in E.$$

*It is called **superharmonic** iff $\mathscr{L}f\leqslant 0$ and **subharmonic** iff $\mathscr{L}f\geqslant 0$.*

Remark. Note that if all these cases, if $(X_n)_{n\geqslant 0}$ solve the martingale problem for $\mathscr{L}$ then the process $(f(X_n))_{n\geqslant 0}$ is either a martingale, a supermartingale or a submartingale.

The generator $\mathscr{L}$ of a Markov process satisfy the maximum priciple:

Theorem. *(Maximum principle) Let $\mathscr{L} = P - \mathrm{Id}$ be the generator of a Markov process on $(E,\mathscr{E})$. Let $(X_n)_{n\geqslant 0}$ to be the canonical Markov process for the generator $\mathscr{L}$ with family of laws given by $(\mathbb{P}_x)_{x\in E}$. Let $D\in\mathscr{E}$ and*

$$T_{D^c} = \inf\{n\geqslant 0 : X_n\in D^c\}.$$

Assume that T_{D^c} is finite almost surely under any $\mathbb{P}_x$ for all $x\in E$. Then if $f: E\to R$ is a bounded function which is subharmonic for $\mathscr{L}$, i.e.

$$\mathscr{L}f(x)\geqslant 0,$$

then we have

$$\sup_{x\in D} f(x) \leqslant \sup_{x\in D^c} f(x).$$

Proof. Then we know that the process $(f(X_n))_{n\geqslant 0}$ is a submartingale under every $\mathbb{P}_x$ for all $x\in E$. By the submartingale property, the optional stopping theorem, for all $n\geqslant 0$ and for all $x\in E$

$$f(x)=\mathbb{E}_x[f(X_0)]\leqslant\mathbb{E}_x[f(X_{n\wedge T_{D^c}})]$$

by dominated convergence we have

$$\lim_{n\to\infty}\mathbb{E}_x[f(X_{n\wedge T_{D^c}})]=\mathbb{E}_x[f(X_{T_{D^c}})]\leqslant\sup_{z\in D^c}f(z).$$

So in particular we proved that $f(x)\leqslant\sup_{z\in D^c}f(z)$ for all $x\in D$. □