Markov processes – Course notes 1.

The Markov property

`[version 1.2, 2017.10.26]`

We introduce the Markov property and the basic objects in the theory of Markov processes: the transition kernel and the construction of a canonical setup. We discuss basic examples of Markov processes and some technical issues necessary to "tame" the continuum of time, cadlag paths, right–continuous filtrations and the Feller property. A fundamental result is the strong Markov property.

1 The Markov property

Let E be a locally compact separable metric space (hence Polish) and $\mathcal{E}$ the Borel σ–field (generated by open sets). We denote by $\mathcal{F}(E)$ (resp. $\mathcal{F}_b(E), \mathcal{F}_+(E)$) the space of $(E,\mathcal{E})$ measurable real functions (resp. the bounded functions, the positive functions). $C(E)$ denotes the Banach space of bounded continuous functions with supremum norm and $C_0(E)$ those vanishing at infinity, that is those for which $\{x\in E: |f(x)|>\varepsilon\}$ is compact for all $\varepsilon>0$. Denote with $\mathcal{P}(E)$ the probabilities on $(E,\mathcal{E})$ and with $\mathcal{M}(E)$ the signed measures.

Let us be given a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_t,\mathbb{P})$ where $t\in I=\mathbb{R}_+$ or $\mathbb{N}$. and let $(X_t)_t$ be a stochastic process with values in E. It is adapted if $X_t \,\hat{\in}\, \mathcal{F}_t$ (meaning that X_t is $\mathcal{F}_t$ measurable) for all t.

A Markov process (in particular, Markov chain if $t\in\mathbb{N}$) is a process whose future can be predicted from its current state, disregarding its past behaviour.

Definition 1. *The process $(X_t)_t$ is a Markov process if*

$$\mathbb{E}[f(X_t)|\mathcal{F}_s]=\mathbb{E}[f(X_t)|X_s], \qquad s\leqslant t,\ f\in\mathcal{F}_b(E). \tag{1}$$

(Markov property)

A probability kernel on E is a map $\rho: E\to\mathcal{P}(E)$ such that $x\mapsto\rho(x,B)$ is measurable for every $B\in\mathcal{E}$. Note that a probability kernel ρ can also be seen as a map $\rho:\mathcal{F}_b(E)\to\mathcal{F}_b(E)$ or as a map $\rho:\mathcal{M}(E)\to\mathcal{M}(E)$ via

$$x\in E\mapsto\rho f(x):=\int_E f(y)\rho(x,\mathrm{d}y), \qquad B\in\mathcal{E}\mapsto\mu\rho(B):=\int_E\rho(x,B)\mu(\mathrm{d}x).$$

Assuming that E is a Polish space (metrizable, separable, complete) and $\mathcal{E} = \mathcal{B}(E)$ its Borel σ-field guarantees that there exists a regular version of the conditional probability $\mathbb{P}(X_t \in B | X_s)$ namely a probability kernel $P_{s,t}$ such that

$$\mathbb{E}[f(X_t)|X_s] = \int_E f(x) P_{s,t}(X_s, \mathrm{d}x), \qquad \mathbb{P}\text{-a.s.}, s \leqslant t.$$

The family of kernels $(P_{s,t})_{s,t}$ satisfies

$$P_{s,u} P_{u,t} = P_{s,t}, \qquad \mathbb{P} \circ X_s^{-1}\text{-a.s.}$$

Using the Markov property (1) iteratively we deduce that, for all $n \geqslant 1$, $f \in \mathcal{F}_b(E^n)$ and $s \leqslant t_1 \leqslant \cdots \leqslant t_n$,

$$\mathbb{E}[f(X_{t_1}, \ldots, X_{t_n}) | \mathcal{F}_s] = \int_{E^n} f(x_1, \ldots, x_n) P_{s,t_1}(X_s, \mathrm{d}x_1) \prod_{k=1}^{n} P_{t_k, t_{k+1}}(x_k, \mathrm{d}x_{k+1})$$

and in particular

$$\mathbb{E}[f(X_0, X_{t_1}, \ldots, X_{t_n})] = \int_{E^{n+1}} f(x_0, x_1, \ldots, x_n) \nu_0(\mathrm{d}x_0) P_{0,t_1}(x_0, \mathrm{d}x_1) \prod_{k=1}^{n} P_{t_k, t_{k+1}}(x_k, \mathrm{d}x_{k+1}) \qquad (2)$$

for all $f \in \mathcal{F}_b(E^{n+1})$ and $0 \leqslant t_1 \leqslant \cdots \leqslant t_n$, where $\nu_0 = \mathbb{P} \circ X_0^{-1}$ (the law of X_0). This shows that the law of the process $(X_t)_t$ (as a measure on the product space E^I) is uniquely determined by the family $(P_{s,t})_{s,t}$ and ν_0.

To easily go backward, from the kernels to the law of the process we introduce the notion of transition kernel.

Definition 2. *A transition kernel (or function) is a family of probability kernels $P = (P_{s,t})_{s,t}$ such that*

a) $P_{s,u} P_{u,t}(x, B) = P_{s,t}(x, B)$ for all $x \in E$, $B \in \mathcal{E}$, $s \leqslant t \in I$,

b) $P_{s,s}(x, \cdot) = \delta_x$ for all $s \in I$.

Property (a) is called Chapman–Kolmogorov equation.

A transition kernel is *homogeneous* if $P_{s,t} = P_{s+h,t+h}$ for all $h \geqslant 0$. In this case the kernel depends only on $t - s$ and we denote it by $P_t = P_{0,t} = P_{s,s+t}$. This does not implies that the Markov process is stationary.

Any non-homogeneous Markov process can be transformed into an homogeneous one by introducing an extended state space $\hat{E} = \mathbb{R}_+ \times E$ and considering the new process $\hat{X}_t = (t, X_t)$.

Exercise 1. Show that the transition kernel $\hat{P}$ for $(\hat{X}_t)_t$ is

$$\hat{P}_{s,t}((x, u), \mathrm{d}y \mathrm{d}r) = P_{u,u+t-s}(x, \mathrm{d}y) \delta_{u+t-s}(\mathrm{d}r).$$

From now on we will consider only homogeneous Markov processes (that is with homogeneous transition function) unless specifies otherwhise.

2 Some examples

Example 3. (Constant speed) The deterministic process $X_t = x + t$ is an (homogeneous) Markov process with transition kernel

$$P_t(x, \cdot) = \delta_{x+t}, \qquad t \geqslant 0.$$

Small modification of this example can produce already interesting behaviour for which $P_0(x, \cdot) = \delta_x$ does not hold. Take $E = \{0\} \cup (-\infty, -1] \cup [1, +\infty)$ and let

$$P_0(0, \cdot) = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}, \qquad P_t(0, \cdot) = \frac{1}{2}\delta_{1+t} + \frac{1}{2}\delta_{-1-t}, \quad t > 0;$$

$$P_t(x, \cdot) = \mathbb{I}_{x>0}\delta_{x+t} + \mathbb{I}_{x<0}\delta_{x-t}, \qquad x \in E \backslash \{0\}.$$

Then we have $P_t P_0 = P_0 P_t = P_t$, but there is a "branching" at zero.

Example 4. Countable state Markov chain. The easiest situation is where E is countable and time is discrete, in this case we consider the matrix

$$p_{i,j}(t) = P_t(i, \{j\}), \qquad i, j \in E$$

which satisfy the equation

$$p_{i,j}(t+s) = \sum_{k \in E} p_{i,k}(t) p_{k,j}(s), \qquad p_0(i,j) = \delta_{i,j}, \qquad i, j \in E.$$

Definition 5. *A standard Brownian motion $(B_t)_t$ started in x is a real valued process such that*

a) *it has independent increments;*

b) $B_t - B_s \sim \mathcal{N}(0, |t-s|)$ *for all $t > s$;*

c) *it satisfies $B_0 = x$ a.s.*

Definition 6. *A Poisson process $(N_t)_t$ starting at $n \in \mathbb{N}$ with intensity $\lambda > 0$ is an integer valued process such that*

a) *it has independent increments;*

b) $N_t - N_s \sim \text{Poisson}(\lambda(t-s))$ *for all $t > s$;*

c) *it satisfies $N_0 = n$ a.s.*

Definition 7. *A compound Poisson process $(X_t)_t$ with jump probability $\nu \in \mathcal{P}(\mathbb{R})$, intensity λ and starting at $x \in \mathbb{R}$ is a real valued process such that*

$$X_t = x + \sum_{n=1}^{N_t} Y_n,$$

where $(Y_n)_{n\geqslant 1}$ *is an iid sequence with law* ν *and* $(N_t)_t$ *a Poisson process with intensity* λ.

All these processes are Markov processes. Compute their transition functions.

3 Canonical realization

The canonical space for the Markov process is the space $\Omega_{\text{can}} = E^I$ endowed with the product σ-field $\mathcal{F}_{\text{can}}$. In this setting $X_t(\omega) = \omega_t$ is the canonical process on Ω_{can} and $(\mathcal{F}_t^X)_t$ the filtration generated by $(X_t)_t$, namely $\mathcal{F}_t^X = \sigma(X_s: s \leqslant t)$. Then we have

Theorem 8. *For any* $\nu \in \mathcal{P}(E)$ *and transition kernel* P *there exists a unique probability* $\mathbb{P}_\nu$ *on* $(\Omega_{\text{can}}, \mathcal{F}_{\text{can}})$ *such that* $(X_t)_t$ *is a Markov process with initial distribution* ν *and*

$$\mathbb{E}_\nu[f(X_t)|\mathcal{F}_s^X] = P_{s,t} f(X_s), \qquad \mathbb{P}-a.s., s \leqslant t, f \in \mathcal{F}_b(E).$$

Proof. Show that the family of laws given by the expression (2) is consistent and then apply Kolmogorov's extension theorem. □

Usefulness of this canonical realization is clear by the following pathwise version of the Markov property. Let us introduce the shift operator $\theta_s: \Omega_{\text{can}} \to \Omega_{\text{can}}$ defined as $\theta_s \omega(t) = \omega(t+s)$ and denote $\mathbb{P}_{s,x} = \mathbb{P}_{\delta_x}$ the canonical Markov process started in $x \in E$.

Remark 9. Note that, for any $F \in \mathcal{F}_b(\Omega_{\text{can}})$ the map $x \mapsto \mathbb{E}_x[F]$ is measurable (Prove it using a monotone class argument, see the proof below for an hint).

Theorem 10. *The family of probabilities* $(\mathbb{P}_x)_x$ *satisfies the Markov property wrt.* $(\mathcal{F}_t^X)_t$ *if and only if*

$$\mathbb{E}_x[F \circ \theta_s | \mathcal{F}_s] = \mathbb{E}_{X_s}[F], \qquad \mathbb{P}_x\text{-}a.s., \qquad F \in \mathcal{F}_b(\Omega_{\text{can}}), s \geqslant 0. \tag{3}$$

Proof. The "if" direction is clear. Let us consider the "only if". Let $F = f_1(X_{t_1}) \cdots f_n(X_{t_n})$ then $F \circ \theta_s = f_1(X_{s+t_1}) \cdots f_n(X_{s+t_n})$ and the claim reads

$$\mathbb{E}_x[f_1(X_{s+t_1}) \cdots f_n(X_{s+t_n}) | \mathcal{F}_s] = \mathbb{E}_{X_s}[f_1(X_{t_1}) \cdots f_n(X_{t_n})]$$

$$= P_{t_1}(f_1 \cdots P_{t_{n-2}-t_{n-1}} f_{n-1}(P_{t_n - t_{n-1}} f_n))(X_s)$$

which can be checked directly from the one step Markov property. In order to extend the equivalence to all $\mathcal{F}_b(\Omega_{\text{can}})$ we use the standard machinery of the monotone class theorem. We take Π as the family of sets of the form $\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\}$ which is closed under intersection and $\mathcal{H}$ as the vector space of function $F \in \mathcal{F}_b(\Omega_{\text{can}})$ satisfying (3). It is clear that it is closed under monotone limits. But we already shown that functions of the form $f_1(X_{t_1}) \cdots f_n(X_{t_n})$ are in $\mathcal{H}$ so $\mathcal{H}$ contains also $\mathbb{I}_A$ for $A \in \Pi$. This implies that $\mathcal{H} = \sigma(\Pi) = \mathcal{F}_b(\Omega_{\text{can}})$. □

The canonical process constructed via this theorem is very general but not very useful since the σ-field $\mathcal{F}_{\text{can}}$ is too small to describe interesting events (e.g. jumps of the process, or its continuity properties).

Example 11. Think about the homogeneous Markov process for which $P_0(x,\cdot)=\delta_x$ and $P_t(x,\cdot)=\mu$ for $t>0$ where $\mu\in\mathcal{P}(E)$ is given. This process corresponds to a family of iid variables and there is no relation whatsoever between states at different times. So in general we cannot expect any regularity from the sample paths of the process.

A function $f\colon\mathbb{R}_{\geqslant 0}\to E$ is *càdlàg* if at every point it admit limit from the left and it is continuous on the right, namely $f_{t-}=\lim_{s\uparrow t}f_t$ exists and $f_{t+}=\lim_{s\downarrow t}f_t=f_t$. We denote $\Delta f_t=f_{t+}-f_{t-}$ the jump of f in t.

Remark 12. The reader can prove the following properties of càdlàg functions:

a) If $f\colon\mathbb{R}_{\geqslant 0}\to\mathbb{R}$ is càdlàg and $I\subseteq\mathbb{R}_{\geqslant 0}$ is a compact interval, then $\{s\in I\colon|\Delta f_s|>\varepsilon\}$ is finite for any $\varepsilon>0$. This implies that f has at most countably many jumps.

b) A càdlàg function is bounded on a compact interval.

c) A uniform limit of càdlàg functions is càdlàg.

Under some additional mild regularity properties for the transition kernel it is possible to realise a canonical version of the Markov process on the space Ω of *càdlàg* paths on E endowed with the filtration $(\mathcal{F}_t^X)_t$ (See Rogers & Williams, vol 1).

An additional improvement consists in taking the filtration $(\mathcal{F}_t)_t$ defined as $\mathcal{F}_t:=\cap_{s>t}\mathcal{F}_s^X$. This new filtration has the property of being right continuous, namely $\mathcal{F}_t=\mathcal{F}_{t+}:=\cap_{s>t}\mathcal{F}_s$. Peeking just a bit in the future does not give additional informations.

Note that $(\mathcal{G}_{t+})_t$ is always a right continuous filtration for any filtration $(\mathcal{G}_t)_t$. We can similarly define a left continuous filtration $(\mathcal{G}_{t-})_t$ by $\mathcal{G}_{t-}:=\vee_{s<t}\mathcal{G}_s$. (Recall the meaning of $\vee$).

We need to prove that a $(\mathcal{F}_t^X)_t$ Markov process is still Markov under $(\mathcal{F}_t)_t$.

Lemma 13. *Assume that $(X_t)_t$ is Markov wrt. $(\mathcal{F}_t^X)_t$, then if $P_t\colon C_0(E)\to C_0(E)$ for all $t>0$, the process is also Markov wrt. $(\mathcal{F}_t)_t$.*

Proof. Take $s,t\geqslant 0$, $h>0$, $A\in\mathcal{F}_s\subseteq\mathcal{F}_{s+h}^X$ for all $h>0$ and use the $(\mathcal{F}_t^X)_t$ Markov property to have

$$\mathbb{E}[f(X_{t+h+s})\mathbb{I}_A]=\mathbb{E}[P_tf(X_{h+s})\mathbb{I}_A].$$

Now observe that if $f\in C_b(E)$ then by assumption $P_tf(X_{h+s})\in C_b(E)$ and by right continuity of paths we have $f(X_{t+h+s})\to f(X_{t+s})$ and $P_tf(X_{h+s})\to P_tf(X_s)$ as $h\searrow 0$. By dominated convergence we conclude that $\mathbb{E}[f(X_{t+s})\mathbb{I}_A]=\mathbb{E}[P_tf(X_s)\mathbb{I}_A]$. □

If E is locally compact it suffices to verify that P_t sends continuous compactly supported functions into continuous functions since on locally compact spaces we can approximate continuous functions with continuous compactly supported ones.

Definition 14. *A transition kernel such that*

a) $P_t: C_0(E) \to C_0(E)$ *for all* $t \geqslant 0$*;*

b) $\mathrm{limt}_{t\downarrow 0} \| P_t f - f \| \to 0$ *for all* $f \in C_0(E)$*,*

is said to satisfy the Feller property and the associated Markov process is called Feller.

Remark 15. Property b) is equivalent to require the pointwise limit $f(x) = \lim_{t\downarrow 0} P_t f(x)$ for all $x \in E$ and $f \in C_0(E)$. Property b) can fail if we replace $C_0(E)$ by $C(E)$.

Exercise 2. Check that Brownian motion in $\mathbb{R}^n$ and compound Poisson process on a locally compact space are Feller processes.

When a Markov process is realised as a family of probabilities $(\mathbb{P}_x)_x$ on $(\Omega_{\text{cadlag}}, \mathcal{F}, (\mathcal{F}_t)_t)$ where Ω_{cadlag} is the space of cadlag paths, $(\mathcal{F}_t)_t$ is the right continuous filtration considered above, then we say that it is realised on the canonical setup. From now on, unless stated otherwise, we will alway assume that we can construct a given Markov processes on the canonical setup. This is a main technical assumption. Feller processes can be constructed on the canonical setup.

To understand the reason for the canonical setup note that we can now consider very interesting events, for example, the r.v.

$$\limsup_{t\downarrow 0} \frac{X_t}{\varphi(t)}$$

is $\mathcal{F}_0$ measurable but it is not $\mathcal{F}_0^X$ measurable, and on the product space $E^{\mathbb{R}_{\geqslant 0}}$ this r.v. is not even measurable at all.

The transition from $(\mathcal{F}_t^X)_t$ to $(\mathcal{F}_t)_t$ does not add much, events in the immediate future of the present are trivial, as stated by Blumenthal's 0-1 law:

Theorem 16. *(Blumenthal's 0-1 law) If* $(\mathbb{P}_x)_x$ *is a Markov process wrt.* $(\mathcal{F}_t)_t$ *and* $A \in \mathcal{F}_0$ *then*

$$\mathbb{P}_x(A) \in \{0, 1\}.$$

Proof.

$$\mathbb{I}_A = \mathbb{E}_x[\mathbb{I}_A | \mathcal{F}_0] = \mathbb{P}_{X(0)}(A) = \mathbb{P}_x(A), \qquad \mathbb{P}_x\text{-a.s.}$$

□

Example 17. Consider standard Brownian motion on $\mathbb{R}$ (starting in 0) and define

$$\tau := \inf \{t > 0: X_t > 0\}, \qquad \sigma := \inf \{t > 0: X_t < 0\}.$$

By symmetry we have $\mathbb{P}_0(\tau = 0) = \mathbb{P}_0(\sigma = 0)$. Moreover $\mathbb{P}_0(\tau \leqslant t) \geqslant \mathbb{P}_0(B_t > 0) = 1/2$ for all $t > 0$ so $\mathbb{P}_0(\tau = 0) \geqslant 1/2$. But now note that $\{\tau = 0\} \in \mathcal{F}_0$ so by Blumenthal's 0-1 law we deduce that $\mathbb{P}_0(\tau = 0) = \mathbb{P}_0(\sigma = 0) = 1$. Using continuity of paths we conclude that any interval $(0, \varepsilon)$ contains a zero of BM almost surely and from this that BM is monotonic on no interval.

4 Stopping times

The canonical setup is a first device to "tame" the continuum of time. Another important tool are stopping times which select of us interesting moments in the life of a Markov process.

Let $\mathcal{F}_\infty = \vee_{t \geqslant 0} \mathcal{F}_t$. Recall that $(\mathcal{F}_t := \mathcal{F}_{t+}^X)_t$ is defined to be right continuous.

Definition 18. *A r.v. $T: \Omega \to [0, \infty]$ is a stopping time relative to $(\mathcal{F}_t)_t$ if $\{T \leqslant t\} \in \mathcal{F}_t$ for all t.*

Several properties of stopping times are easy exercise to verify:

a) T is a stopping time iff $\{T < t\} \in \mathcal{F}_t$ for all t;

b) for any open $G \subset E$ the r.v. $\tau_G = \inf \{t \geqslant 0: X_t \in G\}$ is a stopping time;

c) if τ, σ are stopping times, then $\tau \wedge \sigma$, $\tau \vee \sigma$ and $\tau + \sigma$ are stopping times;

d) if $(\tau_n)_n$ is a sequence of stopping times, then $\inf_n \tau_n, \sup_n \tau_n, \operatorname{liminf}_n \tau_n, \operatorname{limsup}_n \tau_n$ are stopping times;

If $G \subset E$ is not open it is more difficult to decide if τ_G is a stopping time. We postpone this problem to later stages of development of the theory.

Definition 19. *If τ is a stopping time we let*

$$\mathcal{F}_\tau := \{A \in \mathcal{F}: A \cap \{\tau \leqslant t\} \in \mathcal{F}_t \textit{ for all } t \geqslant 0\}.$$

Lemma 20. *We have the following properties:*

a) $\mathcal{F}_\tau$ is a σ-algebra and $\tau \hat{\in} \mathcal{F}_\tau$;

b) if $\tau \leqslant \sigma$ then $\mathcal{F}_\tau \subseteq \mathcal{F}_\sigma$, if $\tau_n \downarrow \tau$ then $\mathcal{F}_\tau = \cap_n \mathcal{F}_{\tau_n}$;

c) if $(Z_t)_t$ is $(\mathcal{F}_t)_t$ adapted and right continuous then $Z_\tau \mathbb{I}_{\tau < \infty} \hat{\in} \mathcal{F}_\tau$.

Proof. We will prove only the last one. Assume first that τ takes only countably many values $(t_n)_n$ then

$$\{Z_\tau \mathbb{I}_{\tau < \infty} \leqslant a, \tau \leqslant t\} = \cup_n \{Z_{t_n} \leqslant a, \tau = t_n \leqslant t\} \in \mathcal{F}_t$$

for all $t \geqslant 0$, which implies that $\{Z_\tau \mathbb{I}_{\tau < \infty} \leqslant a\} \in \mathcal{F}_\tau$, hence measurability follows. For an arbitrary stopping time τ we approximate it from above with a sequence of stopping times $(\tau_n)_n$ taking countably many values. Specifically, we let

$$\tau_n = \sum_{k \geqslant 0} \frac{k+1}{2^n} \mathbb{I}_{k 2^{-n} \leqslant \tau < (k+1) 2^{-n}} + (+\infty) \mathbb{I}_{\tau = +\infty}.$$

Check that this is a stopping time. Why we approximated from above and not below? Then $\tau_n \downarrow \tau$ and $Z_{\tau_n} \to Z_\tau$ on $\{\tau < \infty\}$ by right continuity of the paths. From the first part of the proof we know already that $\{Z_{\tau_n} \mathbb{I}_{\tau_n < \infty} \leqslant a\} \in \mathcal{F}_{\tau_n} \subseteq \mathcal{F}_{\tau_m}$ for every $m \geqslant n$. Finally we can write

$$Z_\tau \mathbb{I}_{\tau < \infty} = \lim_n Z_{\tau_n} \mathbb{I}_{\tau_n < \infty} \hat{\in} \cap_n \mathcal{F}_{\tau_n} = \mathcal{F}_\tau. \qquad \square$$

Remark 21. The condition $\tau < \infty$ is a bit ugly to carry on. Sometimes it is convenient to add a "cemetery" state Δ to the state space E and define $E_\Delta = E \cup \{\Delta\}$ where Δ is seen as an isolated point if E is compact or as the point at infinity for non-compact E (i.e. its neighborhood basis is the class of all complements of compact subsets of E). E_Δ is a compact separable metric space. In this case we denote by $C_0(E_\Delta)$ the functions vanishing at Δ. The space of cadlag paths Ω over E_Δ is assumed to have the property that if $\omega(s) = \Delta$ for some s then $\omega(t) = \Delta$ for all $t \geqslant s$. This ensures that Δ is an absorbing state for the Markov process and we can set $\omega(\infty) = \Delta$. Transition kernels on E can be naturally extended to E_Δ by making Δ absorbing, i.e. $P_t(\{\Delta\}, \cdot) = \delta_\Delta$ for all $t \geqslant 0$.

In this way it holds that $Y_\tau \hat{\in} \mathcal{F}_\tau$ for all E_Δ valued cadlag paths $(Y_t)_t$. However such conventions should be looked with care.

5 The strong Markov property

Definition 22. *A Markov process $(\mathbb{P}_x)_x$ satisfies the strong Markov property if for any $(\mathcal{F}_t)_t$ stopping time τ and any $F\colon [0, \infty) \times \Omega \to \mathbb{R}$ jointly measurable bounded function we have*

$$\mathbb{E}[Y_\tau \circ \theta_\tau | \mathcal{F}_\tau] = \mathbb{E}_{X(\tau)}[Y_\tau] \qquad \text{on } \{\tau < \infty\} \ \mathbb{P}_x\text{-a.s.}$$

Remark 23. The notation in this definition is a bit ambigous. On the left $Y_\tau \circ \theta_\tau(\omega) = Y(\tau(\omega), \theta_{\tau(\omega)}\omega)$, while on the right we really mean $\mathbb{E}_{X(\tau)}[Y_\tau] = \phi(\tau, X(\tau))$ with $\phi(t, x) = \mathbb{E}_x(Y_t)$.

Proof. If τ takes a countable number of values, by conditioning we reduce the proof to the proof of the standard Markov property. In the general case, approximate τ from above by discrete stopping times as seen before. $\square$

5.1 More on the Brownian zeros via strong Markov

In this part we will discuss the zeros of the Brownian motion exploiting the strong Markov property. We will assume that the Brownian motion is realised on the space $\Omega_{\text{cont}} = C(\mathbb{R}_{\geqslant 0}; \mathbb{R})$ of continuous paths on $\mathbb{R}$. This is always possible and in general certain properties of the transition functions can guarantee that a Feller process does not make jumps. For the moment we will just assume this.

Let $\mathcal{Z}$ be the random set

$$\mathcal{Z}(\omega) = \{t \geqslant 0 \colon \omega(t) = 0\}.$$

This set is closed since we assume ω to be a continuous function. We want to show that, similarly to a Cantor set, it is perfect (all the points are accumulation points), in particular it is uncountable.

Let us begin by noting that the Lebesgue measure $\lambda(\mathcal{Z})$ of $\mathcal{Z}$ is 0 almost surely. By Fubini's theorem

$$\mathbb{E}_x\lambda(\mathcal{Z}) = \mathbb{E}_x\int_0^\infty \mathbb{I}_{\mathcal{Z}}(s)\mathrm{d}s = \int_0^\infty \mathbb{E}_x\mathbb{I}_{\mathcal{Z}}(s)\mathrm{d}s = \int_0^\infty \mathbb{P}_x(X_s = 0)\mathrm{d}s = 0.$$

Proposition 24. $\mathbb{P}_x$*–a.s. the set* $\mathcal{Z}$ *is perfect.*

Proof. We need to show that every point of $\mathcal{Z}$ is a limit point of $\mathcal{Z}$. Consider the stopping time $\tau_a = \inf\{t \geqslant a: X_t = 0\}$ for every $a \geqslant 0$. (We need to assume $\tau_a < \infty$ a.s., how to get rid of this?). Let $A := \{\omega: \omega(t_n) = 0 \text{ for some sequence } t_n \downarrow 0\}$ and observe that $\mathbb{I}_A \circ \theta_{\tau_a} = \mathbb{I}_{A_a}$ where $A_a := \{\omega: \omega(t_n) = 0 \text{ for some sequence } t_n \downarrow \tau_a\}$. The strong Markov property then gives

$$\mathbb{E}_x[\mathbb{I}_A \circ \theta_{\tau_a} | \mathcal{F}_{\tau_a}] = \mathbb{E}_{X_{\tau_a}}[\mathbb{I}_A] = \mathbb{E}_0[\mathbb{I}_A] = 1,$$

since $X_{\tau_a} = 0$ by path continuity and $\mathbb{P}_0(A) = 1$ by Blumenthal's 0/1 law and the fact that we already showed that every interval $(0, \varepsilon)$ contains zeros of the BM. Taking expectation of this equality we conclude that $\mathbb{P}_x(A_a) = 1$. So every τ_a is an accumulation point of zeros from the right. On the other hand, a point $s \in \mathcal{Z} \backslash \{\tau_a: a \in \mathbb{Q}, a \geqslant 0\}$ is such that $a \leqslant \tau_a < s$ for $a \in \mathbb{Q} \cap [0, s]$ so it has to be an accumulation point of zeros from the left. □

Remark 25. Think how to remove the assumption that $\tau_a < \infty$ a.s. in the proof above.

Remark 26. About uncountability of the perfect set $\mathcal{Z}$. Assume that it is countable $\{s_i\} = \mathcal{Z}$ then do the following construction. Let U_1 a neighborhood of s_1. Now look for the next number in the sequence which is in U_1, let it be s_k, define $U_2 \subseteq U_1$ to be a neighborhood of s_k whose closure does not contain s_1. Continue in this way to construct a sequence of closed sets $(\bar{U}_n)_n$ which are non-empty, compact and do not contain the initial part of our sequence $\{s_i\}$. Form $V = \cap_n \bar{U}_n$ which should be non–empty since it is the intersection of a nested sequence of compact sets. But then the elements of V are accumulation points of a subsequence of $\mathcal{Z}$ but surely it cannot be in the list $\{s_i\}$.

Remark 27. The complement $\mathcal{Z}^c$ is the union of countably many disjoint open intervals. The strong Markov property says that starting from the right of one of these intervals, the process is a Brownian motion, however the same cannot be true starting from the left, since in this case there is a finite amount of time in which the process will not touch zero, which we know is impossible from Brownian motion.

6 Some non-strong Markov processes.

Here are some examples where the strong Markov property fails.

1. **Waiting, then constant speed.** On the state space $\mathbb{R}_{\geqslant 0}$ define $\mathbb{P}_x$ as follows. For $x > 0$ let it be the law of the process $X_t = x + t$. For $x = 0$ let $\mathbb{P}_x$ be the measure under which X_t waits at zero for an exponential time after which it grows at speed 1, namely $X_t = (t - T)_+$ where $T \sim \mathcal{E}(1)$. It can be checked directly that this is a Markov process with respect to the right–continuous filtration $(\mathcal{F}_t)_t$. Let $\tau = \inf\{t \geqslant 0: X_t > 0\}$. This is a stopping time, but if we consider the function $F = X_1$ we have that $(X_t \circ \theta_\tau) = X_{t+\tau} = t$ with probability 1 while $\mathbb{E}_{X_\tau}(X_1) < 1$, so the strong Markov property does not hold. Indeed this is not a Feller process.

2. **Brownian motion with a twist.** Let $\mathbb{P}_x$ be a Brownian motion for all $x \in \mathbb{R} \backslash \{0\}$ and let $\mathbb{P}_0 = \delta_{\omega^0}$ be the law concentrated on the constant path $\omega^0(t) = 0$. This is a Markov process and $\tau = \inf \{t \geqslant 0: X_t = 0\}$ a stopping time. But now $X_\tau = 0$ and $\mathbb{E}_{X_\tau}[F] = F(\omega^0)$ while for $x \neq 0$ $\mathbb{E}_x[F \circ \theta_\tau]$ is just a probability for a standard Brownian motion.

3. **Càgdàl paths.** Finally, we give an example which does satisfy the Feller property for the transition kernel, but which still fails the strong Markov property. In this case, the failure is caused by the process not having right–continuous paths (recall that right–continuous paths is one of our standing assumptions, so it's implicitly in our proof of the strong Markov property). Let X be a left–continuous version of the Poisson process which can be constructed by taking $(T_n)_{n \geqslant 1}$ $\mathcal{E}(1)$ iid random variables and letting

$$X_t = \sum_{k \geqslant 1} \mathbb{I}_{T_1 + \cdots + T_n < t}, \qquad t \geqslant 0.$$

Let $\tau = \inf \{t: X_t > 0\}$, then $\tau = T_1$ and it is a stopping time with respect to $(\mathcal{F}_t = \mathcal{F}_{t+}^X)_t$ but not with respect to $(\mathcal{F}_t^X)_t$. In this case the strong Markov property fails: the l.h.s. is equal to $\mathbb{E}_1[F_\tau]$ while the r.h.s. is $\mathbb{E}_0[F_\tau]$ because $X_\tau = 0$.