Two Classical Markov semigroups

Throughout this chapter, $(\Omega, \mathcal{A}, \mathbb{P})$ denotes a probability space, so that $\mathcal{A}$ is a σ-algebra of subsets of Ω and $\mathbb{P} : \mathcal{A} \to [0,1]$ is a probability measure, and S denotes a topological space, with $\mathcal{S}$ its Borel σ-algebra (that generated by its open sets).

An *S-valued random variable* is a $\mathcal{A}$-$\mathcal{S}$-measurable mapping $X : \Omega \to S$, so that

$$X^{-1}(A) := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{A} \qquad \text{for all } A \in \mathcal{S}.$$

If X is an S-valued random variable then $\sigma(X)$ is the smallest sub-σ-algebra $\mathcal{A}_0$ of $\mathcal{A}$ such that X is $\mathcal{A}_0$-$\mathcal{S}$ measurable.

Similarly, if $(X_i)_{i \in I}$ is an indexed set of S-valued random variables then $\sigma(X_i : i \in I)$ is the smallest sub-σ-algebra $\mathcal{A}_0$ of $\mathcal{A}$ such that X_i is $\mathcal{A}_0$-$\mathcal{S}$ measurable for all $i \in I$.

2.1 Markov processes

Definition 2.1. Given a topological space $(S, \mathcal{S})$, let

$$B_b(S) := \{f : S \to \mathbb{C} \mid f \text{ is Borel measurable and bounded}\},$$

with vector-space operations defined pointwise and supremum norm

$$\|f\| := \sup\{|f(x)| : x \in S\}.$$

By *Borel measurable*, we mean that $f^{-1}(A) = \{x \in S : f(x) \in A\} \in \mathcal{S}$ for every Borel-measurable subset $A \subseteq \mathbb{C}$.

Exercise 2.2. Show that $B_b(S)$ is a Banach space. Show further that the norm $\|\cdot\|$ is submultiplicative, where multiplication of functions is defined pointwise, so that $B_b(S)$ is a *Banach algebra*. Show finally that the *C^* identity* holds:

$$\|f\|^2 = \|f^* f\| \qquad \text{for all } f \in B_b(S),$$

where the isometric involution $f \mapsto f^*$ is such that $f^*(x) := \overline{f(x)}$ for all $x \in S$.

Definition 2.3. (Provisional) A *Markov process* with *state space* S is a collection of S-valued random variables $X = (X_t)_{t\in\mathbb{R}_+}$ on a common probability space such that, given any $f \in B_b(S)$,

$$\mathbb{E}[f(X_t) \mid \sigma(X_r : 0 \leqslant r \leqslant s)] = \mathbb{E}[f(X_t) \mid \sigma(X_s)] \qquad \text{for all } s, t \in \mathbb{R}_+ \text{ such that } s \leqslant t.$$

A Markov process is *time homogeneous* if, given any $f \in B_b(S)$,

$$\mathbb{E}[f(X_t) \mid X_s = x] = \mathbb{E}[f(X_{t-s}) \mid X_0 = x] \quad \text{for all } s, t \in \mathbb{R}_+ \text{ such that } s \leqslant t \text{ and } x \in S. \tag{2.1}$$

The above is somewhat informal; equality of conditional expectations must be interpreted almost surely, but what is the proper meaning of (2.1)? To be more precise, we introduce the following notion.

Definition 2.4. A *transition kernel* on $(S, \mathcal{S})$ is a map $p : S \times \mathcal{S} \to [0, 1]$ such that

(i) the map $x \mapsto p(x, A)$ is Borel measurable for all $A \in \mathcal{S}$

and (ii) the map $A \mapsto p(x, A)$ is a probability measure for all $x \in S$.

We interpret $p(x, A)$ as the probability that the transition ends in A, given that it started at x.

Exercise 2.5. If p and q are transition kernels on $(S, \mathcal{S})$ then the *convolution* $p * q$ is defined by setting

$$(p * q)(x, A) := \int_S p(x, \mathrm{d}y) q(y, A) \qquad \text{for all } x \in S \text{ and } A \in \mathcal{S}.$$

Prove that $p * q$ is a transition kernel. Prove also that convolution is associative: if p, q and r are transition kernels then $(p * q) * r = p * (q * r)$.

Definition 2.6. A triangular collection $\{p_{s,t} : s, t \in \mathbb{R}_+,\ s \leqslant t\}$ of transition kernels is *consistent* if $p_{s,t} * p_{t,u} = p_{s,u}$ for all s, t, $u \in \mathbb{R}_+$ with $s \leqslant t \leqslant u$, that is,

$$p_{s,u}(x, A) = \int_S p_{s,t}(x, \mathrm{d}y) p_{t,u}(y, A) \qquad \text{for all } x \in S \text{ and } A \in \mathcal{S}.$$

This is the *Chapman–Kolmogorov equation*. We interpret $p_{s,t}(x, A)$ as the probability of moving from x at time s to somewhere in A at time t.

A family of S-valued random variables $X = (X_t)_{t\in\mathbb{R}_+}$ on a common probability space is a *Markov process* if there exists a consistent triangular collection of transition kernels such that

$$\mathbb{E}[1_A(X_t) \mid \sigma(X_r : 0 \leqslant r \leqslant s)] = p_{s,t}(X_s, A) \quad \text{almost surely}$$

for all $A \in \mathcal{S}$ and s, $t \in \mathbb{R}_+$ such that $s \leqslant t$.

Definition 2.7. Similarly, a one-parameter collection $\{p_t : t \in \mathbb{R}_+\}$ of transition kernels is consistent if $p_s \star p_t = p_{s+t}$ for all $s, t \in \mathbb{R}_+$. In this case, the Chapman–Kolmogorov equation becomes

$$p_{s+t}(x, A) = \int_S p_s(x, \mathrm{d}y) p_t(y, A) \qquad \text{for all } x \in S \text{ and } A \in \mathcal{S}. \tag{2.2}$$

We interpret $p_t(x, A)$ as the probability of moving from x into A in t units of time.

A family of S-valued random variables $X = (X_t)_{t \in \mathbb{R}_+}$ on a common probability space is a *time-homogeneous Markov process* if there exists a consistent one-parameter collection of transition kernels such that

$$\mathbb{E}[1_A(X_t) \mid \sigma(X_r : 0 \leqslant r \leqslant s)] = p_{t-s}(X_s, A) \quad \text{almost surely}$$

for all $A \in \mathcal{S}$ and $s, t \in \mathbb{R}_+$ such that $s \leqslant t$.

Definition 2.8. A *Markov semigroup* on $B_b(S)$ is a contraction semigroup T such that T_t is *positive* for all $t \in \mathbb{R}_+$: if $f \in B_b(S)$ is such that $f \geqslant 0$, that is, $f(x) \in \mathbb{R}_+$ for all $x \in S$, then $T_t f \geqslant 0$.

[Note that we impose no condition with respect to continuity at the origin.]

If T_t preserves the unit, that is, $T_t 1 = 1$, where the constant function $1 : S \to \mathbb{C};\ x \mapsto 1$, for all $t \in \mathbb{R}_+$, then the Markov semigroup T is *conservative*.

Theorem 2.9. Let $p = \{p_t : t \in \mathbb{R}_+\}$ be a family of transition kernels. Setting

$$(T_t f)(x) := \int_S p_t(x, \mathrm{d}y) f(y) \qquad \text{for all } f \in B_b(S) \text{ and } x \in S$$

defines a bounded linear operator on $B_b(S)$ which is positive, contractive and preserves the unit. Furthermore, the family $T = (T_t)_{t \in \mathbb{R}_+}$ is a Markov semigroup if and only if p is consistent.

Proof. If $f \in B_b(S)$, $x \in S$ and $s, t \in \mathbb{R}_+$ then the Chapman–Kolmogorov equation (2.2) implies that

$$\begin{aligned}(T_{s+t} f)(x) = \int_S p_{s+t}(x, \mathrm{d}z) f(z) &= \int_S \int_S p_s(x, \mathrm{d}y) p_t(y, \mathrm{d}z) f(z) \\ &= \int_S p_s(x, \mathrm{d}y)(T_t f)(y) \\ &= \big(T_s(T_t f)\big)(x).\end{aligned}$$

Verifying the remaining claims is left as an exercise. □

2.2 Feller semigroups

Definition 2.10. Let S be a locally compact Hausdorff space. Then

$$C_0(S) := \{f : S \to \mathbb{C} \mid f \text{ is continuous and vanishes at infinity}\} \subseteq B_b(S)$$

is a Banach space when equipped with pointwise vector-space operations and the supremum norm. [Recall that a function $f : S \to \mathbb{C}$ *vanishes at infinity* if, for all $\varepsilon > 0$, there exists a compact set $K \subseteq S$ such that $|f(x)| < \varepsilon$ for all $x \in S \setminus K$.]

Exercise 2.11. Prove that $C_0(S)$ lies inside $B_b(S)$ and is a Banach space, as claimed.

Definition 2.12. A Markov semigroup T is *Feller* if

(i) $T_t\big(C_0(S)\big) \subseteq C_0(S)$ for all $t \in \mathbb{R}_+$

and (ii) $\lim_{t\to 0+} \|T_t f - f\| = 0$ for all $f \in C_0(S)$.

Remark 2.13. If a time-homogeneous Markov process X has Feller semigroup T then

$$\mathbb{E}\big[f(X_{t+h}) - f(X_t) \mid X_t\big] = (T_h f - f)(X_t) = h(Af)(X_t) + o(h),$$

so A describes the change in X over an infinitesimal time interval.

Definition 2.14. An $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{t\in\mathbb{R}_+}$ is a *Lévy process* if and only if X

(i) has independent increments, so $X_t - X_s$ is independent of the past σ-algebra $\sigma(X_r : 0 \leqslant r \leqslant s)$ for all s, $t \in \mathbb{R}_+$ with $s \leqslant t$,

(ii) has stationary increments, so $X_t - X_s$ has the same distribution as $X_{t-s} - X_0$, for all s, $t \in \mathbb{R}_+$ with $s \leqslant t$

and (iii) is continuous in probability at the origin, so $\lim_{t\to 0+} \mathbb{P}\big(|X_t - X_0| \geqslant \varepsilon\big) = 0$ for all $\varepsilon > 0$.

Remark 2.15. Lévy processes are well behaved; they have cádlág modifications, and such a modification is a semimartingale, for example.

Exercise 2.16. Prove that if X is a stochastic process with independent and stationary increments, and with cádlág paths, then X is continuous at the origin in probability.

Theorem 2.17. Every Lévy process gives rise to a Feller semigroup.

Sketch proof. For all $t \in \mathbb{R}_+$, $x \in \mathbb{R}^d$ and Borel $A \subseteq \mathbb{R}^d$, let

$$p_t(x, A) := \mathbb{E}[1_A(X_t - X_0 + x)]$$

and note that p_t is a transition kernel. If $s \in \mathbb{R}_+$ then

$$\begin{aligned} p_t(x, A) = \mathbb{E}[1_A(X_t - X_0 + x)] &= \mathbb{E}[1_A(X_{s+t} - X_s + x)] \\ &= \mathbb{E}[1_A(X_{s+t} - X_s + x) \mid \mathcal{F}_s], \end{aligned} \tag{2.3}$$

where $\mathcal{F}_s := \sigma(X_r : 0 \leqslant r \leqslant s)$; the second equality holds by stationarity and the third by independence. In particular,

$$p_t(X_s, A) = \mathbb{E}[1_A(X_{s+t}) \mid \mathcal{F}_s],$$

so X is a Markov process with transition kernels $\{p_t : t \in \mathbb{R}_+\}$ if these are consistent. For consistency, we use Theorem 2.9; let T be defined as there and note that

$$(T_t f)(x) = \int_S p_t(x, \mathrm{d}y) f(y) = \mathbb{E}[f(X_t - X_0 + x)].$$

From the previous working, it follows that

$$(T_t f)(x) = \mathbb{E}[f(X_{s+t} - X_s + x) \mid \mathcal{F}_s],$$

and replacing x with the $\mathcal{F}_s$-measurable random variable $X_s - X_0 + x$ gives that

$$(T_{s+t} f)(x) = \mathbb{E}[f(X_{s+t} - X_0 + x)] = \mathbb{E}[(T_t f)(X_s - X_0 + x)] = \big(T_s(T_t f)\big)(x),$$

as required.

If $f \in C_0(\mathbb{R}^d)$ then $x \mapsto f(X_t - X_0 + x) \in C_0(\mathbb{R}^d)$ almost surely, whence $T_t f \in C_0(\mathbb{R}^d)$ by the Dominated Convergence Theorem.

For continuity, let $\varepsilon > 0$ and note that $f \in C_0(\mathbb{R}^d)$ is uniformly continuous, so there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$. Hence

$$\begin{aligned} \|T_t f - f\| &\leqslant \sup_{x \in \mathbb{R}^d} \mathbb{E}\big[|f(X_t - X_0 + x) - f(x)|\big] \\ &= \sup_{x \in \mathbb{R}^d} \Big(\mathbb{E}\big[1_{|X_t - X_0| < \delta} |f(X_t - X_0 + x) - f(x)|\big] \\ &\qquad\qquad + \mathbb{E}\big[1_{|X_t - X_0| \geqslant \delta} |f(X_t - X_0 + x) - f(x)|\big]\Big) \\ &\leqslant \varepsilon + 2\|f\| \mathbb{P}\big(|X_t - X_0| \geqslant \delta\big) \\ &\to \varepsilon \qquad \text{as } t \to 0+. \end{aligned}$$

□

Theorem 2.18. Let T be a conservative Feller semigroup. If the state space S is metrisable then there exists a time-homogeneous Markov process which gives rise to T.

Sketch proof. For all $t \in (0, \infty)$, let

$$p_t(x, A) := (T_t 1_A)(x) \quad \text{for all } x \in S \text{ and } A \in \mathcal{S}.$$

Then p_t is readily verified to be a transition kernel.

Let μ be a probability measure on S. If $t_n \geqslant \ldots \geqslant t_1 \geqslant 0$ and $A_1, \ldots A_n \in \mathcal{S}$ then

$$p_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \int_S \mu(\mathrm{d}x_0) \int_{A_1} p_{t_1}(x_0, \mathrm{d}x_1) \cdots \int_{A_n} p_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n).$$

By the Chapman–Kolmogorov equation (2.2), these finite-dimensional distributions form a projective family. The Daniell–Kolmogorov extension theorem now yields a probability measure on the product space

$$\Omega := S^{\mathbb{R}_+} = \{\omega = (\omega_t)_{t \in \mathbb{R}_+} : \omega_t \in S \text{ for all } t \geqslant 0\}$$

such the coordinate projections $X_t : \Omega \to S;\ \omega \mapsto \omega_t$ form a time-homogeneous Markov process X with associated semigroup T. □

Example 2.19. (Uniform motion) If $S = \mathbb{R}$ and $X_t = X_0 + t$ for all $t \in \mathbb{R}_+$ then

$$(T_t f)(x) = f(x+t) = \int_{\mathbb{R}} p_t(x, \mathrm{d}y) f(y) \qquad \text{for all } f \in C_0(\mathbb{R}) \text{ and } x \in \mathbb{R},$$

where the transition kernel $p_t : (x, A) \mapsto \delta_{x+t}(A)$. It follows that X gives rise to a Feller semigroup with generator A such that $Af = f'$ whenever $f \in \operatorname{dom} A$.

Example 2.20. (Brownian motion) If $S = \mathbb{R}$ and X is a standard Brownian motion then Itô's formula gives that

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\,\mathrm{d}s \qquad \text{for all } f \in C^2(\mathbb{R}).$$

It follows that the Lévy process X has a Feller semigroup with the generator A such that $Af = \frac{1}{2}f''$ for all $f \in C^2(\mathbb{R}) \subseteq \operatorname{dom} A$. [Informally,

$$t^{-1}\big(\mathbb{E}[f(X_t) \mid X_0 = x] - f(x)\big) = \frac{1}{2t}\int_0^t \mathbb{E}[f''(X_s)|X_0 = x]\,\mathrm{d}s \to \frac{1}{2}f''(x)$$

as $t \to 0+$.]

Example 2.21. (Poisson process) If $S = \mathbb{R}$ and X is a Poisson process with unit intensity and unit jumps then

$$\mathbb{E}[f(X_t)|X_0 = x] = e^{-t}\sum_{n=0}^{\infty} \frac{t^n}{n!} f(x+n) \qquad \text{for all } t \in \mathbb{R}_+.$$

Hence the Lévy process X has a Feller semigroup with the bounded generator A such that $(Af)(x) = f(x+1) - f(x)$ for all $x \in \mathbb{R}$ and $f \in C_0(\mathbb{R})$. [To see this, note that

$$\frac{(T_t f - f)(x)}{t} = \frac{e^{-t} - 1}{t} f(x) + e^{-t} f(x+1) + O(t)$$

as $t \to 0+$, uniformly for all $x \in \mathbb{R}$.]

2.3 The Hille–Yosida–Ray theorem

Throughout this section, S denotes a locally compact Hausdorff space. Let

$$C_0(S;\mathbb{R}) := \{f : S \to \mathbb{R} \mid f \in C_0(S)\}$$

denote the real subspace of $C_0(\mathbb{R})$ containing those functions which take only real values.

Definition 2.22. A linear operator A in $C_0(S)$ satisfies the *positive maximum principle* if, for all $f \in \operatorname{dom} A \cap C_0(S;\mathbb{R})$ and $x_0 \in S$ such that $f(x_0) = \sup\{f(x) : x \in S\}$, it holds that $(Af)(x_0) \leqslant 0$.

Theorem 2.23. (Hille–Yosida–Ray) A closed, densely defined operator A in $C_0(S)$ is the generator of a Feller semigroup on $C_0(S)$ if and only if A satisfies the positive maximum principle and $zI - A$ is surjective for some $z > 0$

Proof. Suppose first that A generates a Feller semigroup on $C_0(S)$. By the Lumer–Phillips theorem, Theorem 1.50, it suffices to show that A satisfies the positive maximum principle. Given $f \in \operatorname{dom} A \cap C_0(S;\mathbb{R})$, let $x_0 \in S$ be such that $f(x_0) = \sup_{x\in S} f(x)$. Setting $f^+ := x \mapsto \max\{f(x), 0\}$, we see that

$$(T_t f)(x_0) \leqslant (T_t f^+)(x_0) \leqslant \|T_t f^+\| \leqslant \|f^+\| = f(x_0).$$

Thus

$$(Af)(x_0) = \lim_{t\to 0+} \frac{(T_t f - f)(x_0)}{t} \leqslant 0,$$

as required.

Conversely, suppose A satisfies the positive maximum principle, and let $f \in \operatorname{dom} A$. There exists $x_0 \in S$ such that $|f(x_0)| = \|f\|$; without loss of generality, let us suppose that $f(x_0) \geqslant 0$. If $z > 0$ then

$$\|(zI - A)f\| \geqslant |zf(x_0) - (Af)(x_0)|,$$

and $(Af)(x_0) \leqslant 0$, by the positive maximum principle. Consequently,

$$\|(zI - A)f\| \geqslant zf(x_0) - Af(x_0) \geqslant zf(x_0) = z\|f\|.$$

Hence T is a strongly continuous contraction semigroup, by the Lumer–Phillips theorem and Remark 1.48.

For positivity, let $f \in C_0(S)$ be non-negative, and consider $g = (zI - A)^{-1} f \in C_0(S)$, where $z > 0$. Either g does not attain its infimum, in which case $g \geqslant 0$ because g vanishes at infinity, or there exists $x_0 \in S$ such that $g(x_0) = \inf g := \inf\{g(x) : x \in S\}$. Then

$$zg - Ag = (zI - A)g = f \iff zg - f = Ag,$$

so $zg(x_0) - f(x_0) = (Ag)(x_0) \geqslant 0$, by the positive maximum principle applied to $-g$. Thus if $x \in S$ then

$$zg(x) \geqslant z \inf g = zg(x_0) \geqslant f(x_0) \geqslant 0,$$

so $z(zI - A)^{-1}$ is positive and therefore so is $(zI - A)^{-1}$. Finally, since

$$T_t f = \lim_{n\to\infty} (I - tn^{-1}A)^{-n} f = \lim_{n\to\infty} (t^{-1}n)^n (t^{-1}nI - A)^{-n} f \qquad \text{for all } f \in X, \tag{2.4}$$

each T_t is positive also. □

Remark 2.24. Equation (2.4), which holds in full generality, shows how one may recover a semigroup solely from the family of resolvents. See [3, Problem 8.2.4].