Finite state Markov chains in exercises

June 26, 2021

Contents

1 Conditional measure and expectation

Conditional measure

Definition 1.1. Let X be a finite set. Any function $\mu : X \to \mathbb{R}_+$ is called a *non-negative measure* on X. If, in addition,

$$\sum_{x \in X} \mu(x) = 1, \tag{1.1}$$

then μ is called a *probability measure* on X. If X is equipped with a probability measure, then a function $\xi : X \to \mathbb{R}$ is often called a *random variable*. For any subset $\Gamma \subset X$ and any random variable $\xi : X \to \mathbb{R}$, we define

$$\mu(\Gamma) := \sum_{x \in \Gamma} \mu(x), \quad \mathbb{E}\,\xi := \int_X \xi \,\mathrm{d}\mu \equiv \int_X \xi(x)\mu(\mathrm{d}x) \equiv \sum_{x \in X} \xi(x)\mu(x).$$

We call $\mu(\Gamma)$ the *measure* of Γ and $\mathbb{E}\,\xi$ the *mean value* of ξ.

Let us denote by $\mathcal{P}(X)$ the set of all probability measures on X. The *support* of μ is defined as $\operatorname{supp}\mu = \{x \in X : \mu(x) \neq 0\}$. A property depending $x \in X$ is said to be valid for μ*-almost everywhere* x (μ-a.e.) or μ*-almost surely* (μ-a.s.) if it is true on $\operatorname{supp}\mu$. For instance, a random variable $\xi : X \to \mathbb{R}$ is positive μ-a.s. if $\xi(x) > 0$ whenever $\mu(x) > 0$. In what follows, all abstract spaces are assumed to be finite, unless specified otherwise.

Exercise 1. Let μ be a probability measure on a finite set X and $\xi : X \to \mathbb{R}$ be a random variable.

Chebyshev inequality. Prove that if ξ is non-negative μ-almost everywhere, then

$$\mu\big(\{x \in X : \xi(x) \geq a\}\big) \leq a^{-1} \int_X \xi \,\mathrm{d}\mu \quad \text{for any } a > 0. \tag{1.2}$$

Jensen inequality. Recall that a function $\varphi : \mathbb{R} \to \mathbb{R}$ is said to be *convex* if, for any positive numbers $p_1, \dots, p_k$ whose sum is equal to 1 and any points $y_1, \dots, y_k \in \mathbb{R}$, we have

$$\varphi(p_1 y_1 + \cdots + p_k y_k) \leq p_1 \varphi(y_1) + \cdots + p_k \varphi(y_k). \tag{1.3}$$

Prove that if $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function, then

$$\varphi\left(\int_X \xi \,\mathrm{d}\mu\right) \leq \int_X \varphi(\xi)\,\mathrm{d}\mu. \tag{1.4}$$

Given a measure $\mu \in \mathcal{P}(X \times Y)$, where Y is another finite set, we write μ_X and μ_Y for the *marginals* of μ defined by the relations

$$\mu_X(x) = \sum_{y \in Y} \mu(x, y), \quad \mu_Y(y) = \sum_{x \in X} \mu(x, y).$$

We shall sometimes write $\Pi_X \mu$ and $\Pi_Y \mu$ for the marginals of μ.

Exercise 2. Let μ be a probability measure on a product space $X \times Y$. Prove that the marginals of μ are probability measures.

Given a measure $\mu \in \mathcal{P}(X \times Y)$, we define the *conditional measure* of μ given $x \in X$ by the relation

$$\mu^x(y) = \frac{\mu(x, y)}{\mu_X(x)} \quad \text{for } x \in \operatorname{supp} \mu_X. \tag{1.5}$$

Thus, the conditional measure μ^x is defined for μ-a.e. $x \in X$.

Exercise 3. Let μ be a probability measure on a product space $X \times Y$ and let μ^x be its conditional measure given $x \in X$.

(a) Prove that μ^x is a probability measure on Y for μ_X-a.e. $x \in X$.

(b) Prove that, for any random variable $\xi : X \times Y \to \mathbb{R}$,

$$\int\limits_{X \times Y} \xi \, \mathrm{d}\mu = \int_X \left\{ \int_Y \xi(x, y) \mu^x(\mathrm{d}y) \right\} \mu_X(\mathrm{d}x). \tag{1.6}$$

Given two measures $\mu_1 \in \mathcal{P}(X)$ and $\mu_2 \in \mathcal{P}(Y)$, we define their *direct product* $\mu_1 \otimes \mu_2$ by

$$(\mu_1 \otimes \mu_2)(x, y) = \mu_1(x)\mu_2(y) \quad \text{for } (x, y) \in X \times Y.$$

We call $\mu_1 \otimes \mu_2$ the *product measure* of μ_1 and μ_2.

Exercise 4. Prove that a probability measure μ on $X \times Y$ is the product of its marginals if and only if μ^x does not depend on $x \in \operatorname{supp} \mu_X$. Show also that, in this case, $\mu^x = \mu_Y$ for $x \in \operatorname{supp} \mu_X$.

Exercise 5. Let $X = \{1, \dots, N\}$ and let $\mu \in \mathcal{P}(X \times X)$ be defined by

$$\mu(i, j) = m^{-1}(\delta_{ij} + \delta_{i,j+1}), \quad i, j \in X,$$

where $m > 0$ is a normalising factor. Determine m and find the marginals of μ.

Conditional expectation

Let us fix a measure $\mu \in \mathcal{P}(X \times Y)$ and denote by $\mathcal{F}_X$ the family of all subsets of $X \times Y$ that have the form $A \times Y$ for some $A \subset X$.

Exercise 6. Prove that, for any random variable $\xi : X \times Y \to \mathbb{R}$, there is $\eta : X \to \mathbb{R}$ such that

$$\mathbb{E}(\xi \mathbb{1}_A) = \mathbb{E}(\eta \mathbb{1}_A) \quad \text{for any } A \subset X, \tag{1.7}$$

where $\mathbb{1}_A$ stands for the *indicator function* of A (that is, $\mathbb{1}_A(x) = 1$ for $x \in A$ and $\mathbb{1}_A(x) = 0$ otherwise), and both η and $\mathbb{1}_A$ are regarded as functions on $X \times Y$. Show also that η is uniquely determined on $\operatorname{supp} \mu_X$.

The random variable η constructed in Exercise 6 is denoted by $\mathbb{E}(\xi \,|\, \mathcal{F}_X)$ and called the *conditional expectation of ξ given $\mathcal{F}_X$*. In the case when X is a singleton, $\mathbb{E}(\xi \,|\, \mathcal{F}_X)$ is equal to the mean value of ξ. The following exercise establishes a relation between conditional expectation and conditional measure.

Exercise 7. Let $\mu \in \mathcal{P}(X \times Y)$ be a measure and let $\xi : X \times Y \to \mathbb{R}$ be a random variables. Prove that, for any $x \in \operatorname{supp} \mu_X$,

$$\mathbb{E}(\xi \,|\, \mathcal{F}_X)(x) = \int_Y \xi(x,y)\mu^x(\mathrm{d}y). \tag{1.8}$$

In particular, $\mu^x(y) = \mathbb{E}(\mathbb{1}_{\{y\}} \,|\, \mathcal{F}_X)(x)$ for $x \in \operatorname{supp} \mu_X$.

Exercise 8. Prove the following properties of the conditional expectation.

Linearity. For any numbers $a, b \in \mathbb{R}$ and random variables $\xi, \eta : X \times Y \to \mathbb{R}$,

$$\mathbb{E}(a\xi + b\eta \,|\, \mathcal{F}_X) = a\,\mathbb{E}(\xi \,|\, \mathcal{F}_X) + b\,\mathbb{E}(\eta \,|\, \mathcal{F}_X) \quad \text{on } \operatorname{supp} \mu_X. \tag{1.9}$$

Positivity. If $\xi \geq 0$ on $\operatorname{supp} \mu$, then $\mathbb{E}(\xi \,|\, \mathcal{F}_X) \geq 0$ on $\operatorname{supp} \mu_X$.

Exercise 9. **(a)** If a random variable $\eta : X \times Y \to \mathbb{R}$ does not depend on $y \in Y$, then

$$\mathbb{E}(\xi\eta \,|\, \mathcal{F}_X) = \eta\,\mathbb{E}(\xi \,|\, \mathcal{F}_X) \quad \text{on } \operatorname{supp} \mu_X. \tag{1.10}$$

(b) Prove that if the measure $\mu \in \mathcal{P}(X \times Y)$ is the product of its marginals, ξ is a random variable not depending on $x \in X$, and η depends only on $x \in X$, then for any $f : \mathbb{R}^2 \to \mathbb{R}$ we have

$$\mathbb{E}\big(f(\xi,\eta) \,|\, \mathcal{F}_X\big) = \big(\mathbb{E} f(\xi, z)\big)\Big|_{z=\eta(x)} \quad \text{for } x \in \operatorname{supp} \mu_X. \tag{1.11}$$

Exercise 10. Let $\xi : X \times Y \to \mathbb{R}$ be a random variable. Prove that $\mathbb{E}(\xi \,|\, \mathcal{F}_X)$ minimises the function $\eta \mapsto \mathbb{E}\,(\xi - \eta)^2$ over all $\eta : X \to \mathbb{R}$.

Exercise 11. Let μ be a probability measure on the space $S = X \times Y \times Z$ and let $\xi : S \to \mathbb{R}$ be a random variable. Prove that

$$\mathbb{E}\left(\mathbb{E}(\xi \,|\, \mathcal{F}_{X \times Y}) \,|\, \mathcal{F}_X\right) = \mathbb{E}(\xi \,|\, \mathcal{F}_X) \quad \text{on } \operatorname{supp} \mu_X. \tag{1.12}$$

Exercise 12 (Jensen inequality)**.** Prove that if $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function, then

$$\varphi\big(\mathbb{E}(\xi \,|\, \mathcal{F}_X)\big) \leq \mathbb{E}\big(\varphi(\xi) \,|\, \mathcal{F}_X\big) \quad \text{on } \operatorname{supp} \mu_X. \tag{1.13}$$

The L^p*-norm* of a random variable $\xi : X \to \mathbb{R}$ is defined by

$$|\xi|_p = \left(\int_X |\xi|^p \mathrm{d}\mu\right)^{1/p}.$$

Exercise 13. Prove that $|\mathbb{E}(\xi \,|\, \mathcal{F}_X)|_p \leq |\xi|_p$ for any $p \geq 1$. Formulate and prove a similar result for $p = \infty$.

2 Definition of a Markov process

Markov measures

Let $\mathcal{A} = \{1, \dots, N\}$ be a finite alphabet and let Ω be the set of all sequences $\omega = (\omega_j)_{j \geq 0}$ whose elements belong to $\mathcal{A}$. For any integer $t \geq 0$, we write Ω_t for the set of all finite sequences $\boldsymbol{\omega}^t := (\omega_0, \dots, \omega_t)$ such that $\omega_j \in \mathcal{A}$ for any $j \in [\![0, t]\!]$.

Definition 2.1. A *probability measure* $\mathbb{P}$ on Ω is defined as a sequence of measures $\mathbb{P}_t \in \mathcal{P}(\Omega_t)$ such that

$$\Pi_{\Omega_t} \mathbb{P}_{t+1} = \mathbb{P}_t \quad \text{for any integer } t \geq 0. \tag{2.1}$$

A measure $\mathbb{P}$ on Ω is said to be *Markovian* if there is a function $p : \mathcal{A} \times \mathcal{A} \to \mathbb{R}_+$ such that

$$\mathbb{P}_{t+1}^{\boldsymbol{\omega}^t}(\omega_{t+1}) = p(\omega_t, \omega_{t+1}) \tag{2.2}$$

for any $t \geq 0$ and any $\omega_0, \dots, \omega_{t+1} \in \mathcal{A}$ satisfying $\mathbb{P}_t(\omega_0, \dots, \omega_t) > 0$.

We shall write Π_t instead of Π_{Ω_t} to simplify notation, and the conditional measure on the left-hand side of (2.2) is sometimes denoted by $\mathbb{P}(\omega_{t+1} \,|\, \omega_0, \dots, \omega_t)$. The set of all measures on Ω will be denoted by $\mathcal{P}(\Omega)$. For an event $\Gamma \subset \Omega_t$ and a random variable $\xi : \Omega_t \to \mathbb{R}$, we denote $\mathbb{P}(\Gamma) := \mathbb{P}_t(\Gamma)$ and $\mathbb{E}\,\xi := \mathbb{E}_t \xi$, where $\mathbb{E}_t$ stands for the mean value with respect to $\mathbb{P}_t$.

Exercise 14. Let $\mathbb{P}$ be a Markov measure and let $i \in \mathcal{A}$ be such that $\mathbb{P}_0(i) > 0$. Prove the following relations:

$$p(i, j) = \frac{\mathbb{P}_1(i, j)}{\mathbb{P}_0(i)} \quad \text{for any } j \in \mathcal{A}, \qquad \sum_{j \in \mathcal{A}} p(i, j) = 1. \tag{2.3}$$

In what follows, we shall often write p_{ij} instead of $p(i, j)$ and call $P := (p_{ij})$ the *transition matrix* of the Markov measure $\mathbb{P}$. Note that $p_{ij} \geq 0$ and $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ stands for a vector of length N all of whose components are 1. Any matrix with these properties will be called a *stochastic matrix*.

Exercise 15. Let $\mathbb{P} \in \mathcal{P}(\Omega)$ be a measure such that $\mathbb{P}_{t+1} = \mathbb{P}_t \otimes \mu$ for any $t \geq 0$, where $\mu = \mathbb{P}_0$. Prove that $\mathbb{P}$ is a Markovian measure and find the corresponding transition matrix.

The measure described in the previous exercise is called an *i.i.d. measure*, in view of the following property.

Exercise 16. Under the hypotheses of Exercise 15, prove that, for any integers $s \geq 1$ and $t \geq 2$, and any functions $f, f_1, \dots, f_t : \mathcal{A} \to \mathbb{R}$, we have

$$\mathbb{E} \prod_{k=1}^{t} f_k(\omega_k) = \prod_{k=1}^{t} \mathbb{E}\, f_k(\omega_k), \quad \mathbb{E}\, f(\omega_s) = \mathbb{E}\, f(\omega_t). \tag{2.4}$$

The next two exercises show that a Markovian measure is uniquely defined, once an initial measure and a transition matrix are fixed.

Exercise 17. Let $\mathbb{P}$ be a Markov measure. Prove that, for any $t \geq 1$ and $\omega_0, \dots, \omega_t \in \mathcal{A}$, we have

$$\mathbb{P}_t(\omega_0, \dots, \omega_t) = \mathbb{P}_0(\omega_0)\, p(\omega_0, \omega_1) \dots p(\omega_{t-1}, \omega_t). \tag{2.5}$$

Exercise 18. Let $\mu \in \mathcal{P}(\mathcal{A})$ and let (p_{ij}) be an $N \times N$ stochastic matrix. Prove that there is a Markov measure $\mathbb{P}$ on Ω such that $\mathbb{P}_0 = \mu$ and relation (2.2) holds with $p(i,j) = p_{ij}$.

Markov processes

Definition 2.2. Let $\{\xi_t : \Omega \to \mathcal{A}, t \geq 0\}$ be a sequence of functions. We shall say that $(\xi_t)_{t\geq 0}$ is an *adapted random process* (or simply a *random process*) if for any $t \geq 0$ the function ξ_t depends only on $\boldsymbol{\omega}^t = (\omega_0, \dots, \omega_t) \in \Omega_t$. In this case, we shall often consider ξ_t as a random variable on Ω_t.

Given any integers $s \geq 0$ and $t \in [\![0, s]\!]$, we denote by $\mathcal{F}_{=t}$ the family of all subsets of Ω_s that have the form $\pi_t^{-1}(A)$, where $\pi_t : \Omega_s \to \mathcal{A}$ is the projection to the t^{th} component. We write $\mathbb{P}_{=t}$ for the marginal of $\mathbb{P}_s$ corresponding to the t^{th} component. Recall that $\mathcal{F}_{\Omega_t} =: \mathcal{F}_t$ denotes the family of all subsets of Ω_s that can be represented in the form $\Pi_t^{-1}(\Gamma)$ with $\Gamma \subset \Omega_t$.

Exercise* 19. Let $\mathbb{P}$ be a probability measure on Ω and let $\xi_t : \Omega \to \mathcal{A}$ be an adapted random process defined by

$$\xi_t(\omega) = \omega_t \quad \text{for } \omega = (\omega_j)_{j\geq 0} \in \Omega. \tag{2.6}$$

Assuming that $\mathbb{P}$ is a Markovian measure, prove that the following properties hold and give an informal description for them.[1]

(a) For any integers $0 \leq t < s$ and any function $f : \Omega_{s-t} \to \mathbb{R}$, we have

$$\mathbb{E}\big(f(\xi_t, \dots, \xi_s) \,|\, \mathcal{F}_t\big) = \mathbb{E}\big(f(\xi_t, \dots, \xi_s) \,|\, \mathcal{F}_{=t}\big) \quad \text{on } \operatorname{supp} \mathbb{P}_t. \tag{2.7}$$

(b) For any integers $s > t \geq 0$ and any functions $f : \Omega_{s-t} \to \mathbb{R}$ and $g : \Omega_t \to \mathbb{R}$, we have

$$\begin{aligned}\mathbb{E}\big(f(\xi_t, \dots, \xi_s)\, g(\xi_0, \dots, \xi_t) \,|\, \mathcal{F}_{=t}\big)& \\ = \mathbb{E}\big(f(\xi_t, \dots, \xi_s) \,|\, \mathcal{F}_{=t}\big)\, &\mathbb{E}\big(g(\xi_0, \dots, \xi_t) \,|\, \mathcal{F}_{=t}\big),\end{aligned} \tag{2.8}$$

where the equality holds on the support of $\mathbb{P}_{=t}$.

Show also that properties (a) and (b) are equivalent and imply that relation (2.2) holds on $\operatorname{supp} \mathbb{P}_t$ for any $t \geq 0$ and some function $p_t : \mathcal{A} \times \mathcal{A} \to \mathbb{R}_+$.

[1] Recall that $\mathbb{E} = \mathbb{E}_s$ for random variables defined on Ω_s.

The process defined by (2.6) is called the *canonical process associated with* $\mathbb{P}$. The following definition provides a central object of the theory.

Definition 2.3. Let $\mathcal{A} = \{1, \dots, N\}$ be a finite alphabet, let $P = (p_{ij})_{i,j\in\mathcal{A}}$ be a stochastic matrix, and let $\delta_i \in \mathcal{P}(\mathcal{A})$ be the Dirac mass at $i \in \mathcal{A}$. A discrete-time *Markov process with transition matrix* P is the collection of the following objects:

- **(i)** the canonical process $\xi_t : \Omega \to \mathcal{A}$ defined by (2.6);
- **(ii)** a family of Markov measures $\mathbb{P}^i$ parametrised by $i \in \mathcal{A}$ such that $(\mathbb{P}^i)_0 = \delta_i$, and the transition matrix of $\mathbb{P}^i$ is equal to P.

We write $(\xi_t, \mathbb{P}^i)_{i\in\mathcal{A}}$ for the Markov process with Markov measures $\mathbb{P}^i$.

Exercise 20. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process whose transition matrix has the form $p_{ij} = p_j$, where $\{p_j, j \in \mathcal{A}\}$ are non-negative numbers whose sum is equal to 1. Describe the Markov process $(\xi_t, \mathbb{P}^i)$.

Exercise 21. Prove that for any stochastic matrix P there is a unique Markov process with transition matrix P. Show also that the transition matrix of a Markov process $(\xi_t, \mathbb{P}^i)$ is uniquely defined.

Exercise 22. Let $\mathbb{P}$ be a Markov measure on Ω with transition matrix P and let $\mu := \mathbb{P}_0$. Prove that

$$\mathbb{P} = \sum_{i\in\mathcal{A}} \mu(i)\, \mathbb{P}^i, \tag{2.9}$$

where $(\xi_t, \mathbb{P}^i)$ is the Markov process with transition matrix P.

Exercise 23. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process and let $\mathbb{P}$ be a Markov measure. Prove the following assertions.

(a) For any integers $0 \le t < s$ and any function $f : \mathcal{A} \to \mathbb{R}$, we have

$$\mathbb{E}\big(f(\xi_s) \,|\, \mathcal{F}_t\big) = \mathbb{E}^{\omega_t} f(\xi_{s-t}) \quad \text{on } \operatorname{supp} \mathbb{P}_t. \tag{2.10}$$

(b) For any integers $0 \le t < s$ and any function $f : \Omega_{s-t} \to \mathbb{R}$, we have

$$\begin{aligned} \mathbb{E}\big(f(\xi_t, \dots, \xi_s) \,|\, \mathcal{F}_t\big) &= \mathbb{E}^{\omega_t} f(\xi_0, \dots, \xi_{s-t}) \\ &= \sum_{i_1,\dots,i_{s-t}\in\mathcal{A}} f(i_0, \dots, i_{s-t}) \prod_{l=1}^{s-t} p_{i_{l-1}i_l}, \end{aligned} \tag{2.11}$$

where $i_0 = \omega_t$, and the equality holds on $\operatorname{supp} \mathbb{P}_t$.

Exercises 19 and 23 establish the *Markov property* for the canonical process. An important generalisation of this property will be obtained in Section 4.

Markov semigroups

Many properties of a Markov process can be described in terms of Markov semigroups introduced in the following definition. We denote by $L(\mathcal{A})$ the space of functions $f : \mathcal{A} \to \mathbb{R}$ and endow it with the norm

$$|f| = \max_{i\in\mathcal{A}} |f(i)|. \tag{2.12}$$

Definition 2.4. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process with a finite state space $\mathcal{A}$ and a transition matrix $P = (p_{ij})$. The *Markov operators* $\mathfrak{P}_t : L(\mathcal{A}) \to L(\mathcal{A})$ and $\mathfrak{P}_t^* : \mathcal{P}(\mathcal{A}) \to \mathcal{P}(\mathcal{A})$ are defined by the relations

$$(\mathfrak{P}_t f)(i) = \mathbb{E}^i f(\xi_t), \tag{2.13}$$
$$(\mathfrak{P}_t^* \mu)(i) = \mathbb{P}^\mu\{\xi_t = i\}, \tag{2.14}$$

where $f \in L(\mathcal{A})$, $\mu \in \mathcal{P}(\mathcal{A})$, $i \in \mathcal{A}$, and $\mathbb{P}^\mu$ stands for the Markov measure (2.9).

Exercise 24. Describe the Markov operators for the Markov process defined in Exercise 20.

Exercise 25. Let $p_{ij}^{(t)}$ denote the entries of the matrix P^t. Prove that

$$(\mathfrak{P}_t f)(i) = \sum_{j\in\mathcal{A}} p_{ij}^{(t)} f(j), \tag{2.15}$$
$$(\mathfrak{P}_t^* \mu)(i) = \sum_{k\in\mathcal{A}} \mu(k)\, p_{ki}^{(t)}. \tag{2.16}$$

It is useful to introduce the *transition function*

$$P_t(i,\Gamma) := \sum_{j\in\Gamma} p_{ij}^{(t)}, \quad i \in \mathcal{A}, \quad \Gamma \subset \mathcal{A}, \tag{2.17}$$

which gives the probability of transition from the state i to the set Γ at time t. Relation (2.15) says that $\mathfrak{P}_t f$ is the integral of f against $P_t(i,\cdot)$, while (2.16) implies that $\mathfrak{P}_t^*\mu(\Gamma)$ is the integral of $P_t(\cdot,\Gamma)$ against μ.

Exercise 26. Check the following properties of the transition function and Markov operators.

Kolmogorov–Chapman relation. For any $i \in \mathcal{A}$, $\Gamma \subset \mathcal{A}$, and $t, s \geq 0$, we have

$$P_{t+s}(i,\Gamma) = \int_{\mathcal{A}} P_s(x,\Gamma) P_t(i,\mathrm{d}x) = \sum_{j\in\mathcal{A}} p_{ij}^{(t)} P_s(j,\Gamma). \tag{2.18}$$

Semigroup property. The sequences $\{\mathfrak{P}_t\}_{t\geq 0}$ and $\{\mathfrak{P}_t^*\}_{t\geq 0}$ are semigroups. That is, for any integers $t, s \geq 0$,

$$\mathfrak{P}_0 = \mathrm{Id}, \quad \mathfrak{P}_{t+s} = \mathfrak{P}_t \circ \mathfrak{P}_s, \qquad \mathfrak{P}_0^* = \mathrm{Id}, \quad \mathfrak{P}_{t+s}^* = \mathfrak{P}_t^* \circ \mathfrak{P}_s^*, \tag{2.19}$$

where Id denotes the identity map in the corresponding space.

Duality. For any $f \in L(\mathcal{A})$ and $\mu \in \mathcal{P}(\mathcal{A})$, we have

$$\mathbb{E}^\mu f(\xi_t) = \int_{\mathcal{A}} \mathfrak{P}_t f \, \mathrm{d}\mu = \int_{\mathcal{A}} f \, \mathrm{d}(\mathfrak{P}_t^* \mu). \tag{2.20}$$

In what follows, we call $\mathfrak{P}_t$ and $\mathfrak{P}_t^*$ the *Markov semigroups* for the Markov process $(\xi_t, \mathbb{P}^i)$. The following exercise provides yet another form of the Markov property.

Exercise 27. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process and let $f : \mathcal{A} \to \mathbb{R}$ be a function. Prove that, for any integers $t, s \geq 0$,

$$\mathbb{E}^i\big(f(\xi_{t+s}) \,|\, \mathcal{F}_t\big) = (\mathfrak{P}_s f)(\xi_t) \quad \text{on } \operatorname{supp} \mathbb{P}_t^i. \tag{2.21}$$

3 Stationary and ergodic measures

Definition 3.1. A measure $\mu \in \mathcal{P}(\mathcal{A})$ is said to be *stationary* (or *invariant*) for a Markov process $(\xi_t, \mathbb{P}^i)$ if $\mathfrak{P}_1^* \mu = \mu$.

Given a measure $\mathbb{P}$ on Ω and integers $0 \leq t \leq s$, we write $\mathbb{P}_{[t,s]}$ for the marginal of $\mathbb{P}_s$ corresponding to the components $(\omega_j, t \leq j \leq s)$.

Exercise 28. Let $\mathbb{P}$ be a Markov measure with a transition matrix P and an initial distribution $\mu := \mathbb{P}_0$. Prove that μ is stationary for the Markov process with the transition matrix P if and only if $\mathbb{P}_{[t,s]} = \mathbb{P}_{[0,s-t]}$ for any $0 \leq t \leq s$. In particular, if μ is stationary, then $\mathbb{P}_{=t} = \mu$ for any $t \geq 0$.

Exercise 29. Prove that any Markov process has a stationary measure. Show also that the stationary measures form a convex subset of $\mathcal{P}(\mathcal{A})$.

Exercise 30. Find all the stationary measures for the Markov processes with the following transition matrices:

$$P_1 = \begin{pmatrix} 0.3 & 0.7 \\ 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \end{pmatrix}.$$

In what follows, we denote by $\mathcal{P}_s(\mathcal{A}, P)$ the set of stationary measures for the Markov process with transition matrix P. We shall often write $\mathcal{P}_s(\mathcal{A})$ if the context implies the choice of the transition matrix.

Definition 3.2. A subset $\mathcal{C} \subset \mathcal{A}$ is said to be *invariant* for the Markov process $(\xi_t, \mathbb{P}^i)$ if

$$P_1(i, \mathcal{C}) = 1 \quad \text{for any } i \in \mathcal{C}. \tag{3.1}$$

In other words, a subset $\mathcal{C} \subset \mathcal{A}$ is invariant if a trajectory starting in $\mathcal{C}$ remains there for all $t \geq 1$ with probability 1.

Exercise 31. Prove the following properties.

(a) For any invariant subset $\mathcal{C} \subset \mathcal{A}$ there is a stationary measure whose support is a subset of $\mathcal{C}$.

(b) The support of a stationary measure is an invariant subset.

Definition 3.3. A stationary measure $\mu \in \mathcal{P}_s(\mathcal{A}, P)$ is said to be *ergodic* if for any $\nu \in \mathcal{P}_s(\mathcal{A}, P)$ either $\operatorname{supp}\mu \cap \operatorname{supp}\nu = \varnothing$ or $\operatorname{supp}\mu \subset \operatorname{supp}\nu$. The set of ergodic invariant measures is denoted by $\mathcal{E}(\mathcal{A}, P)$ or $\mathcal{E}(\mathcal{A})$.

Two measures $\mu, \nu \in \mathcal{P}(\mathcal{A})$ are said to be *singular* if $\operatorname{supp}\mu \cap \operatorname{supp}\nu = \varnothing$. In this case, we shall write $\mu \perp\!\!\!\perp \nu$.

Exercise 32. Prove that if two different measures $\mu, \nu \in \mathcal{P}_s(\mathcal{A}, P)$ have the same support, then neither of them is ergodic. Use this to check that any two ergodic measures $\mu, \nu \in \mathcal{E}(\mathcal{A}, P)$ are singular.

Exercise 33. Determine which of the stationary measures of Example 30 are ergodic.

Given a subset $\mathcal{A}^0 \subset \mathcal{A}$, we define the *attainable set* from $\mathcal{A}^0$ by the formula

$$S(\mathcal{A}^0) = \Big\{ j \in \mathcal{A} : \sum_{t=0}^{\infty} p_{ij}^{(t)} > 0 \text{ for some } i \in \mathcal{A}^0 \Big\}.$$

In the case $\mathcal{A}^0 = \{i\}$, we shall write S^i.

Exercise 34. Prove that $S(\mathcal{A}^0)$ is an invariant subset for any $\mathcal{A}^0 \subset \mathcal{A}$.

Exercise 35. Let $\mu \in \mathcal{P}_s(\mathcal{A})$. Prove that the following properties are equivalent:

(a) μ is ergodic;

(b) $S^i = \operatorname{supp}\mu$ for any $i \in \operatorname{supp}\mu$.

Exercise 36. Prove that for any invariant subset $\mathcal{C} \subset \mathcal{A}$ there is an ergodic measure whose support is included in $\mathcal{C}$.

Exercise 37 (ergodic decomposition)**.** Prove that any invariant measure is a convex combination of ergodic measures. That is, for any $\mu \in \mathcal{P}_s(\mathcal{A}, P)$ there are ergodic measures $\mu_1, \dots, \mu_k \in \mathcal{P}_s(\mathcal{A}, P)$ and positive numbers $p_1, \dots, p_k$ such that

$$\mu = \sum_{l=1}^{k} p_l \mu_l, \quad \sum_{l=1}^{k} p_l = 1. \tag{3.2}$$

Exercise 38. Let $\mathcal{A} = \{1, 2\}$. In the setting of the first two examples of Exercise 30, study the following limits for a function $f : \mathcal{A} \to \mathbb{R}$ satisfying $f(1) \neq f(2)$ and any measure $\nu \in \mathcal{P}(\mathcal{A})$:

$$\lim_{t\to\infty} \mathbb{E}^{\nu} f(\xi_t), \quad \lim_{t\to\infty} \mathbb{P}^{\nu} \Big\{ \Big| \frac{1}{t} \sum_{k=0}^{t-1} f(\xi_k) - L \Big| < \varepsilon \Big\}, \tag{3.3}$$

where $L \in \mathbb{R}$ is a fixed number and $\varepsilon > 0$ is smaller than $|f(1) - f(2)|$.

4 Stopping times and strong Markov property

Stopping times

Definition 4.1. A subset $\Gamma \subset \Omega$ is said to be *cylindrical* if there is an integer $t \geq 0$ and a set $\Gamma^0 \subset \Omega_t$ such that $\Gamma = \Pi_t^{-1}(\Gamma^0)$.

The family of all cylindrical subsets is denoted by $\mathscr{C}$. The set Γ^0 entering the definition of a cylindrical set Γ is called *base*. We shall write $\mathscr{C}_t$ to denote those cylindrical sets for which one can choose a base in Ω_t. Note that if $\Gamma \in \mathscr{C}_t$, then $\Gamma \in \mathscr{C}_s$ for any $s > t$.

Exercise 39. Prove that $\mathscr{C}$ is an algebra. That is, for any $\Gamma_1, \Gamma_2 \in \mathscr{C}$, we have $\Gamma_1 \cap \Gamma_2, \Gamma_1 \cup \Gamma_2, \Gamma_1^c \equiv \Omega \setminus \Gamma_1 \in \mathscr{C}$.

If $\mathbb{P} \in \mathcal{P}(\Omega)$, then we define the probability of a cylindrical set $\Gamma = \Pi_t^{-1}(\Gamma_0)$ by $\mathbb{P}(\Gamma) := \mathbb{P}_t(\Gamma_0)$.

Exercise 40. Prove that $\mathbb{P} : \mathscr{C} \to [0,1]$ is a well-defined additive function vanishing on the empty set.

Definition 4.2. Let $\overline{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{+\infty\}$. A function $\tau : \Omega \to \overline{\mathbb{Z}}_+$ is called a *stoping time* if

$$\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathscr{C}_t \quad \text{for any integer } t \geq 0. \tag{4.1}$$

In what follows, we shall write $\{\tau \leq t\}$ for the set in (4.1).

Exercise 41. Prove that if τ_1 and τ_2 are stopping times, then $\tau_1 \vee \tau_2$, $\tau_1 \wedge \tau_2$, and $\tau_1 + \tau_2$ are also stopping times.

Exercise 42. Let $A \subset \mathcal{A}$. Determine which of the following functions are stopping times:

$$\tau_A(\omega) = \min\{t \geq 0 : \xi_t(\omega) \in A\}, \tag{4.2}$$

$$\tau_A^+(\omega) = \min\{t \geq 0 : \xi_s(\omega) \in A \text{ for } s \geq t\}, \tag{4.3}$$

$$\tau_A^-(\omega) = \max\{t \geq 0 : \xi_s(\omega) \in A \text{ for } s \leq t\}. \tag{4.4}$$

Given an integer $s \geq 0$, define the *shift operator* $\theta_s : \Omega \to \Omega$ by the formula

$$(\theta_s \omega)_k = \omega_{k+s} \quad \text{for } k \geq 0. \tag{4.5}$$

Exercise 43. Prove that $\theta_s^{-1}(\mathscr{C}_t) \subset \mathscr{C}_{t+s}$ for any integers $t, s \geq 0$.

Exercise 44. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process with a state space $\mathcal{A}$. Prove that $\mu \in \mathcal{P}_s(\mathcal{A})$ if and only if $\mathbb{P}^\mu(\theta_t^{-1}(\Gamma) = \mathbb{P}^\mu(\Gamma)$ for any $t \geq 0$ and $\Gamma \in \mathscr{C}$.

Exercise 45. Let $\mathbb{P} \in \mathcal{P}(\Omega)$ be a Markovian measure. Prove that, for any integers $t, s \geq 0$ and any function $f : \mathcal{A}^{s+1} \to \mathbb{R}$, we have

$$\mathbb{E}\big(f(\xi_0 \circ \theta_t, \dots, \xi_s \circ \theta_t) \,|\, \mathcal{F}_t\big) = \mathbb{E}^{\xi_t(\omega)} f(\xi_0, \dots, \xi_s), \tag{4.6}$$

where the equality holds on $\operatorname{supp} \mathbb{P}_t$. Use this to calculate the conditional expectation $\mathbb{E}(\mathbb{1}_{\{\tau \circ \theta_t \geq s\}} \,|\, \mathcal{F}_t)$.

Exercise 46. Let τ and σ be two stopping times. Prove that the function $T = \tau + \sigma \circ \theta_\tau$ defined by

$$T(\omega) = \tau(\omega) + \sigma(\theta_{\tau(\omega)}\omega) \tag{4.7}$$

is a stopping time.

Definition 4.3. Let $\mathbb{P} \in \mathcal{P}(\Omega)$ and let $\Gamma \subset \Omega$. We say that Γ is an *event of probability zero* and write $\mathbb{P}(\Gamma) = 0$ if there is a decreasing sequence of cylindrical sets $(\Gamma_k)_{k\geq 1}$ such that

$$\Gamma \subset \bigcap_{k=1}^{\infty} \Gamma_k, \quad \lim_{k\to\infty} \mathbb{P}(\Gamma_k) = 0. \tag{4.8}$$

If Γ is the complement of an event of probability zero, then we say that Γ is an *event of probability one* or that Γ *is* $\mathbb{P}$-*almost surely true*, and we write $\mathbb{P}(\Gamma) = 1$.

Exercise 47. Prove that $\Gamma \subset \Omega$ is an event of probability one if and only there is an increasing sequence of cylindrical sets $(\Gamma_k)_{k\geq 1}$ such that

$$\Gamma \supset \bigcup_{k=1}^{\infty} \Gamma_k, \quad \lim_{k\to\infty} \mathbb{P}(\Gamma_k) = 1. \tag{4.9}$$

The following result is related to *Kolmogorov's extension theorem* and establishes a *continuity property* for measures.

Exercise* 48. Let $\Gamma \subset \Omega$ be an event of $\mathbb{P}$-probability zero and let $(B_k)_{k\geq 1}$ be a decreasing sequence of cylindrical sets whose intersection coincides with Γ. Prove that $\mathbb{P}(B_k) \to 0$ as $k \to \infty$. Formulate and prove a similar property for an event of probability one.

Exercise* 49. Let $(\xi_t, \mathbb{P}^i)$, $i \in \mathcal{A}$, be a Markov process and let τ and σ be two stopping times that are $\mathbb{P}^i$-almost surely finite for any $i \in \mathcal{A}$. Prove that the stopping time T defined by (4.7) is $\mathbb{P}^\mu$-a.s. finite for any initial measure $\mu \in \mathcal{P}(\mathcal{A})$.

Strong Markov property

Exercise 50. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process, let $\tau : \Omega \to \overline{\mathbb{Z}}_+$ be a stopping time, and let $\sigma = \varphi(\tau)$, where $\varphi : \mathbb{Z}_+ \to \mathbb{Z}_+$ is an arbitrary function. Prove that, for any integer $t \geq 0$, any initial measure $\mu \in \mathcal{P}(\mathcal{A})$, and any functions $f : \mathcal{A} \to \mathbb{R}$ and $g : \mathbb{Z}_+ \to \mathbb{R}$, we have

$$\mathbb{E}^\mu\big(\mathbb{1}_{\{\tau\leq t\}} f(\xi_{\tau+\sigma}) g(\tau)\big) = \mathbb{E}^\mu\big(\mathbb{1}_{\{\tau\leq t\}} (\mathfrak{P}_\sigma f)(\xi_\tau) g(\tau)\big), \tag{4.10}$$

where the left- and right hand sides are defined with the help of the representation $\{\tau \leq t\} = \{\tau = 0\} \cup \cdots \cup \{\tau = t\}$.

Relation (4.10) remains true if we replace the indicator function of $\{\tau \leq t\}$ by that of the set $\{\tau < \infty\}$, provided that g is a bounded function:

$$\mathbb{E}^\mu\big(\mathbb{1}_{\{\tau<\infty\}} f(\xi_{\tau+\sigma}) g(\tau)\big) = \mathbb{E}^\mu\big(\mathbb{1}_{\{\tau<\infty\}} (\mathfrak{P}_\sigma f)(\xi_\tau) g(\tau)\big). \tag{4.11}$$

However, in this case the mean values on the left- and right-hand sides need to be defined. This can be done in a "natural" way, using the following relation and interchanging the integral $\mathbb{E}^\mu$ and the sum in t:

$$\mathbb{1}_{\{\tau<\infty\}} = \sum_{t=0}^{\infty} \mathbb{1}_{\{\tau=t\}}. \tag{4.12}$$

Exercise 51. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process and let $\tau : \Omega \to \mathbb{Z}_+$ be a stopping time such that $\tau \leq t$ with $\mathbb{P}^i$-probability 1. Prove the following properties.

(a) For any $i, j \in \mathcal{A}$, we have

$$p_{ij}^{(t)} = \mathbb{E}^i p_{\xi_\tau j}^{(t-\tau)}. \tag{4.13}$$

(b) For any $i \in \mathcal{A}$ and any function $f : \mathcal{A} \to \mathbb{R}$, we have

$$\mathfrak{P}_t f(i) = \mathbb{E}^i\big((\mathfrak{P}_{t-\tau} f)(\xi_\tau)\big). \tag{4.14}$$

5 Perron–Frobenius theorem and strongly recurrent Markov processes

Perron–Frobenius theorem for stochastic matrices

Recall that an $N \times N$ matrix $P = (p_{ij})$ with non-negative entries is said to be *stochastic* if $\sum_j p_{ij} = 1$ for any $i \in [\![1, N]\!]$. We denote by P^* the adjoint matrix of P.

Exercise 52. Prove the following properties for a stochastic matrix P.

(a) All the eigenvalues of P and P^* belong to the closed unit circle centred at origin.

(b) The matrices P and P^* possess real eigenvectors with the eigenvalue $\lambda = 1$.

(c) The matrices P and P^* possess eigenvectors with the eigenvalue $\lambda = 1$ that have non-negative coordinates.

Definition 5.1. A stochastic matrix P is said to be *strongly recurrent* if there are $m \geq 1$ and $j_0 \in \mathcal{A}$ such that

$$p_{ij_0}^{(m)} > 0 \quad \text{for any } i \in \mathcal{A}. \tag{5.1}$$

A Markov process $(\xi_t, \mathbb{P}^i)$ is said to be *strongly recurrent* if so is its transition matrix.

In other words, a Markov process is strongly recurrent if there is a state $j_0 \in \mathcal{A}$ and a time $m \geq 1$ such that the probability of transition from any state $i \in \mathcal{A}$ to j_0 is positive at time m. The goal of Exercises 53–62 is to prove that, for strongly recurrent stochastic matrices, $\lambda = 1$ is a simple eigenvalue, and it is the only one on the boundary of the unit circle. The first step is Banach's *contraction mapping principle*.

Exercise 53. Let $K \subset \mathbb{R}^N$ be a closed set and let $F : K \to K$ be a map satisfying the inequality

$$|F(p) - F(q)| \leq \varkappa \, |p - q| \quad \text{for } p, q \in K, \tag{5.2}$$

where $|\cdot|$ is a norm on $\mathbb{R}^N$ and $\varkappa \in (0,1)$ is a number. Prove the following properties.

Fixed point. There is a unique point $\hat{p} \in K$ such that $F(\hat{p}) = \hat{p}$.

Convergence. For any $p \in K$ and $n \geq 0$, we have

$$|F^n(p) - \hat{p}| \leq \varkappa^n |p - \hat{p}|, \tag{5.3}$$

where F^n denotes the n^{th} iteration of F.

Exercise 54. With the notation of the previous exercise, suppose there are numbers $k \geq 2$ and $\varkappa \in (0,1)$ such that

$$|F^k(p) - F^k(q)| \leq \varkappa \, |p - q| \quad \text{for } p, q \in K. \tag{5.4}$$

Prove that F has a unique fixed point $\hat{p} \in K$, and for any $p \in K$ the sequence $(F^n(p))_{n \geq 1}$ converges to $\hat{p}$ exponentially fast, uniformly with respect to $p \in K$.

Recall that the space $L(\mathcal{A})$ of functions from $\mathcal{A}$ to $\mathbb{R}$ is endowed with the maximum norm $|\cdot|$ given by (2.12). Considering $\mathcal{P}(\mathcal{A})$ as a subset of $\mathbb{R}^N$, we define the *total variation metric* on it as the restriction to $\mathcal{P}(\mathcal{A})$ of the norm dual to $|\cdot|$:

$$\|\mu_1 - \mu_2\| = \frac{1}{2} \sup_{|f| \leq 1} \left| \int_{\mathcal{A}} f \mathrm{d}\mu_1 - \int_{\mathcal{A}} f \mathrm{d}\mu_2 \right|. \tag{5.5}$$

Exercise 55. Prove that, for any $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{A})$,

$$\|\mu_1 - \mu_2\| = \frac{1}{2} \sum_{j \in \mathcal{A}} |\mu_1(j) - \mu_2(j)| = 1 - \sum_{j \in \mathcal{A}} \mu_1(j) \wedge \mu_2(j), \tag{5.6}$$

where $a \wedge b$ stands for the minimum of a and b. Show also that $\mathcal{P}(\mathcal{A})$ is a closed subset of $\mathbb{R}^N$.

Exercise 56. Prove that two measures $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{A})$ are singular if and only if $\|\mu_1 - \mu_2\| = 1$.

The next four exercises provides a "probabilistic proof" of the celebrated *Perron–Frobenius theorem*. Given a stochastic matrix P, we denote by $\mathfrak{P}_t$ and $\mathfrak{P}_t^*$ the corresponding Markov semigroups; see (2.15) and (2.16).

Exercise 57. Let P be a stochastic matrix. Prove that

$$\|\mathfrak{P}_1^*\mu_1 - \mathfrak{P}_1^*\mu_2\| \leq \|\mu_1 - \mu_2\| \quad \text{for any } \mu_1, \mu_2 \in \mathcal{P}(\mathcal{A}). \tag{5.7}$$

Exercise 58 (Dobrushin decomposition)**.** Prove that, for any $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{A})$, there are three measures $\nu, \mu_1', \mu_2' \in \mathcal{P}(\mathcal{A})$ such that $\mu_1' \perp\!\!\!\perp \mu_2'$ and

$$\mu_i = (1-d)\nu + d\mu_i', \quad i = 1, 2, \tag{5.8}$$

where $d = \|\mu_1 - \mu_2\|$.

Exercise 59. Let P be a stochastic matrix and let $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{A})$ be two measures such that $\mu_i(j_0) \geq \varepsilon$ for $i = 1, 2$, where $\varepsilon > 0$ is a number and $j_0 \in \mathcal{A}$ is an index. Prove that

$$\|\mathfrak{P}_1^*\mu_1 - \mathfrak{P}_1^*\mu_2\| \leq 1 - \varepsilon. \tag{5.9}$$

Exercise 60. Let $(\xi_t, \mathbb{P}^i)$ be a strongly recurrent Markov process and let $m \geq 1$ be the integer entering Definition 5.1. Prove the following properties.

Contraction. The map $\mathfrak{P}_{m+1}^* : \mathcal{P}(\mathcal{A}) \to \mathcal{P}(\mathcal{A})$ is a contraction.

Stationary measure. There is a unique stationary measure $\mu \in \mathcal{P}(\mathcal{A})$.

Exponential mixing. There are positive numbers C and γ such that, for any $f \in L(\mathcal{A})$, $\nu \in \mathcal{P}(\mathcal{A})$, and $t \geq 0$,

$$|\mathfrak{P}_t f - \langle f, \mu\rangle \mathbb{1}| \leq C\, e^{-\gamma t}|f|, \tag{5.10}$$

$$\|\mathfrak{P}_t^*\nu - \mu\| \leq C\, e^{-\gamma t}, \tag{5.11}$$

where $\langle f, \mu\rangle$ denotes the integral of f against μ.

Given a stationary measure $\mu \in \mathcal{P}(\mathcal{A})$ for a Markov process $(\xi_t, \mathbb{P}^i)$, we can define the corresponding *path measure* $\mathbb{P}^\mu$ by (2.5) with $\mathbb{P}_0 = \mu$. We shall sometimes denote it by $\boldsymbol{\mu}$ to emphasise the difference between a probability measure and a measure on the state space.

Exercise 61. Under the hypotheses of Exercise 60, prove that, for any integer $s \geq 1$, any $i \in \mathcal{A}$, and any function $f : \mathcal{A}^{s+1} \to \mathbb{R}$,

$$|\mathbb{E}^i f(\xi_t, \ldots, \xi_{t+s}) - \langle f, \boldsymbol{\mu}\rangle| \leq C_s e^{-\gamma t}|f|, \quad t \geq 0, \tag{5.12}$$

where $\gamma > 0$ is the same as Exercise 60, and $C_s > 0$ depends only on s.

Exercise 62 (Perron–Frobenius theorem for strongly recurrent matrices)**.** Let P be a strongly recurrent matrix. Prove the following assertions.

- **(a)** The point $\lambda = 1$ is an algebraically simple eigenvalue of P.[2]
- **(b)** The eigenvalues of P that are different from 1 are strictly inside the unit circle centred at origin.

[2]That is, $\lambda = 1$ is a simple root of the characteristic polynomial for P. In particular, there is only one eigenvector corresponding to $\lambda = 1$.

Perron–Frobenius theorem for general positive matrices

Let P be a matrix with non-negative entries. Let us emphasise that we do not assume in this subsection that P is stochastic, so that Exercise 62 is not applicable. The goal of Exercises 65–68 is to prove that the conclusion is nevertheless true, if the hypothesis of strong recurrence is replaced by strong irreducibility (see Definition 5.2 below). The results of this subsection will be important in the theory of large deviations.

Exercise 63. Construct a 2×2 matrix with non-negative entries satisfying (5.1) for some integers $j_0 \in \{1,2\}$ and $m \geq 1$ such that the conclusion of Exercise 62 (a) is not true (with $\lambda = 1$ replaced by the spectral radius of P).

Exercise* **64** (Brouwer theorem)**.** Let $B \subset \mathbb{R}^N$ be a closed ball and $F : B \to B$ be a continuous map. Prove that there is at least one point $\hat{x} \in B$ such that $F(\hat{x}) = \hat{x}$. Is the result true when B is the boundary of a ball ?

Exercise 65. Let P be a matrix with non-negative entries.

- **(a)** Prove that both P and P^* have at least one eigenvector with non-negative coordinates and a non-negative eigenvalue λ and λ'.
- **(b)** Let $h \in \mathbb{R}^N$ and $\mu \in \mathbb{R}^N$ be eigenvectors of P and P^*, respectively, constructed in (a). Prove that if $\lambda \neq \lambda'$, then $\langle h, \mu\rangle = 0$, where $\langle \cdot, \cdot \rangle$ stands for the Euclidean scalar product.
- **(c)** Construct an example when λ and λ' are different, and one of them is equal to zero?

Definition 5.2. We shall say that a matrix P with non-negative entries is *strongly irreducible* if there is an integer $m \geq 1$ such that all the entries of P^m are strictly positive.

Exercise 66. Let P be a strongly irreducible matrix. Prove that the eigenvectors h and μ constructed in Exercise 65 have positive coordinates, and the corresponding eigenvalues λ and λ' are equal and positive.

Exercise* **67** (Perron–Frobenius theorem for matrices with positive entries)**.** Let P be a matrix with positive entires. Prove that there is a number $\lambda > 0$ and vectors $h, \mu \in \mathbb{R}^N$ with positive coordinates such that

$$Ph = \lambda h, \quad P^*\mu = \lambda\mu, \quad \langle h, \mu\rangle = \langle \mathbb{1}, \mu\rangle = 1, \tag{5.13}$$

$$|\lambda^{-k}P^k f - \langle f, \mu\rangle h| \leq Ce^{-\gamma k}|f| \quad \text{for any } k \geq 0,\ f \in \mathbb{R}^N, \tag{5.14}$$

$$\|\lambda^{-k}(P^*)^k \nu - \langle h, \nu\rangle \mu\| \leq Ce^{-\gamma k}\|\nu\| \quad \text{for any } k \geq 0,\ \nu \in \mathbb{R}^N, \tag{5.15}$$

where C and γ are positive numbers not depending on k, f, and ν.

Hint: Consider the semigroup generated by the operator $Qf = \lambda^{-1}h^{-1}P(hf)$, where h and λ are constructed in Exercise 66.

Exercise 68 (Perron–Frobenius theorem)**.** Let P be a strongly irreducible matrix. Prove that the conclusion of Exercise 67 remains valid.

Exercise 69. Under the hypotheses of Exercise 68, prove that the positive eigenvalue λ constructed there is a simple root of the characteristic polynomial of P and all other eigenvalues are in the open circle of radius λ centre at zero.

Exercise 70. Let $D \subset \mathbb{R}^d$ be an open set and let $D \ni \alpha \mapsto P(\alpha)$ be a C^s-smooth matrix function with non-negative entries such that $P(\alpha)$ is strongly irreducible for any $\alpha \in D$. Prove that the eigenvalue $\lambda(\alpha)$ constructed in Exercise 69 is a C^s-function of $\alpha \in D$.

6 Doeblin coupling approach

Measurable sets, random variables, and integral

As before, we are dealing with the space $\Omega = \mathcal{A}^{\mathbb{Z}_+}$, where $\mathcal{A} = \{1, \dots, N\}$, and a probability measure $\mathbb{P} = \{\mathbb{P}_t, t \in \mathbb{Z}_+\}$. Setting $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$, we now define the concepts of a measurable set and a measurable function.

Definition 6.1. A function $\xi : \Omega \to \overline{\mathbb{R}}_+$ is said to be *cylindrical* if there is an integer $t \geq 1$ and a function $\tilde{\xi} : \Omega_t \to \overline{\mathbb{R}}_+$ such that $\xi(\omega) = \tilde{\xi}(\Pi_t \omega)$ for any $\omega \in \Omega$. A function $\xi : \Omega \to \overline{\mathbb{R}}_+$ is said to be $\mathbb{P}$-*measurable* if there is a sequence of cylindrical functions $\xi_t : \Omega \to \mathbb{R}_+$ such that, for $\mathbb{P}$-almost every $\omega \in \Omega$, we have

$$\xi_t(\omega) \leq \xi_{t+1}(\omega) \quad \text{for any } t \geq 1, \qquad \xi(\omega) = \lim_{t \to \infty} \xi_t(\omega). \tag{6.1}$$

Finally, a set $\Gamma \subset \Omega$ is said to be $\mathbb{P}$-*measurable* if its indicator function $\mathbb{1}_\Gamma$ is $\mathbb{P}$-measurable.

Exercise 71. Let $\Gamma \subset \Omega$. Prove that the following properties are equivalent:

- **(a)** Γ is $\mathbb{P}$-measurable;
- **(b)** there is an increasing sequence $\{\Gamma_t \subset \mathscr{C}_t, t \geq 1\}$ such that

$$\mathbb{P}\Bigl(\Gamma \vartriangle \bigcup_{t=1}^{\infty} \Gamma_t\Bigr) = 0\,; \tag{6.2}$$

Definition 6.2. Given a $\mathbb{P}$-measurable function $\xi : \Omega \to \overline{\mathbb{R}}_+$, we define its *integral* as

$$\mathbb{E}\,\xi \equiv \int_\Omega \xi \, d\mathbb{P} = \lim_{t \to \infty} \mathbb{E}\,\xi_t, \tag{6.3}$$

where $(\xi_t)_{t \geq 1}$ is the sequence entering (6.1). If this quantity if finite, then we say that ξ is $\mathbb{P}$-*integrable*.

Exercise* 72. Prove that the integral $\mathbb{E}\xi$ can be calculated by the formula

$$\int_\Omega \xi \, d\mathbb{P} = \sup_{\tilde{\xi} \le \xi} \int_\Omega \tilde{\xi} \, d\mathbb{P}, \tag{6.4}$$

where the supremum is taken over all cylindrical functions $\tilde{\xi} : \Omega \to \mathbb{R}_+$ such that $\tilde{\xi}(\omega) \le \xi(\omega)$ for any $\omega \in \Omega$. In particular, integral (6.3) is well defined in the sense it does not depend on the choice of $(\xi_t)_{t \ge 1}$.

A measurable function $\xi : \Omega \to \overline{\mathbb{R}}_+$ will often be called a *random variable*, and the integral $\int_\Omega \xi d\mathbb{P}$ is called the *mean value* of ξ and is denoted by $\mathbb{E}\,\xi$. If $\Gamma \subset \Omega$ is $\mathbb{P}$-measurable, then its probability $\mathbb{P}(\Gamma)$ is defined as the mean value $\mathbb{E}\mathbb{1}_\Gamma$.

Exercise 73. Prove that $\mathbb{E}\xi$ is a linear non-decreasing function of ξ and an affine function of $\mathbb{P}$.

Exercise 74. Let $\xi : \Omega \to \overline{\mathbb{R}}_+$ be a $\mathbb{P}$-integrable random variable. Prove that

$$\mathbb{P}\{\xi > a\} \le a^{-1} \mathbb{E}\,\xi \quad \text{for any } a > 0. \tag{6.5}$$

Exercise 75. Let the transition function of a Markov process be such that $p_{i1} > 0$ for any $i \in \mathcal{A}$ and let τ be defined by (4.2) with $A = \{1\}$. Prove that τ is $\mathbb{P}^\mu$-almost surely finite for any initial measure $\mu \in \mathcal{P}(\mathcal{A})$.

Exercise* 76. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process with state space $\mathcal{A}$ and let $A \subset \mathcal{A}$ be a subset possessing the following properties:

(a) for any $i \notin A$ there is an integer $t \ge 1$ such that $P_t(i, A) > 0$.

(b) for any $i \in A$ and $j \notin A$, we have $p_{ij} = 0$.

Prove that, for any initial measure $\mu \in \mathcal{P}(\mathcal{A})$, the stopping time τ_A defined by (4.2) is $\mathbb{P}^\mu$-a.s. finite and satisfies the inequality

$$\mathbb{E}^\mu e^{\alpha \tau_A} \le C < \infty \tag{6.6}$$

where α and C are positive numbers not depending on μ.

Uniqueness of stationary measure and exponential mixing

Let $(\xi_t, \mathbb{P}^i)$, $i \in \mathcal{A}$ be a Markov process with a transition matrix P. We define the extended phase space

$$\boldsymbol{\mathcal{A}} = \mathcal{A} \times \mathcal{A} = \{\boldsymbol{i} = (i_1, i_2) : i_1, i_2 \in \mathcal{A}\}$$

and consider a Markov process $(\boldsymbol{\xi}_t, \mathbb{P}^{\boldsymbol{i}})$, $\boldsymbol{i} \in \boldsymbol{\mathcal{A}}$ whose transition matrix $\boldsymbol{P} = (\boldsymbol{p}_{\boldsymbol{ij}})$ is given by $\boldsymbol{p}_{\boldsymbol{ij}} = p_{i_1 j_1} p_{i_2 j_2}$.

Exercise 77. Prove that $\boldsymbol{P}$ is indeed a stochastic matrix and that

$$\mathbb{P}_t^{\boldsymbol{i}} = \mathbb{P}_t^{i_1} \otimes \mathbb{P}_t^{i_2} \quad \text{for any } t \in \mathbb{Z}_+, \ \boldsymbol{i} = (i_1, i_2) \in \boldsymbol{\mathcal{A}}. \tag{6.7}$$

In view of (6.7), the Markov process $(\boldsymbol{\xi}_t, \mathbb{P}^{\boldsymbol{i}})$ is called a *pair of independent copies* of $(\xi_t, \mathbb{P}^i)$. In what follows, we shall write $\boldsymbol{\xi}_t = (\xi_t^1, \xi_t^2)$.

Exercise 78. Prove that, for any $f : \mathcal{A}^{t+1} \to \mathbb{R}$ and any $\boldsymbol{i} = (i_1, i_2) \in \boldsymbol{\mathcal{A}}$, we have

$$\mathbb{E}^{\boldsymbol{i}} f(\xi_0^1, \dots, \xi_t^1) = \mathbb{E}^{i_1} f(\xi_0, \dots, \xi_t), \quad \mathbb{E}^{\boldsymbol{i}} f(\xi_0^2, \dots, \xi_t^2) = \mathbb{E}^{i_2} f(\xi_0, \dots, \xi_t), \tag{6.8}$$

where $\mathbb{E}^{\boldsymbol{i}}$ stands for the expectation associated with $\mathbb{P}^{\boldsymbol{i}}$.

Let us introduce the random time $\tau = \min\{t \geq 0 : \xi_t^1 = \xi_t^2\}$. Thus, τ is the first instant when the two components of the process $\boldsymbol{\xi}_t$ coincide. This is equivalent to saying that $\boldsymbol{\xi}_t$ hits the diagonal of the product space $\boldsymbol{\mathcal{A}}$.

Exercise 79. Suppose that the transition matrix P is strongly recurrent. Prove that, for any $\boldsymbol{i} \in \boldsymbol{\mathcal{A}}$, the stopping time τ is $\mathbb{P}^{\boldsymbol{i}}$-almost surely finite, and there are positive numbers γ and C such that

$$\mathbb{E}^{\boldsymbol{i}} e^{\gamma\tau} \leq C \quad \text{for any } \boldsymbol{i} \in \boldsymbol{\mathcal{A}}. \tag{6.9}$$

Up to now, we were studying only the canonical process $\boldsymbol{\xi}_t$ defined on the space $\boldsymbol{\Omega} = \boldsymbol{\mathcal{A}}^{\mathbb{Z}_+}$. We now use it to construct an auxiliary process by the relation

$$\eta_t = \begin{cases} \xi_t^2 & \text{for } t \leq \tau, \\ \xi_t^1 & \text{for } t > \tau. \end{cases} \tag{6.10}$$

Thus, the process η_t coincides with ξ_t^2 before the components meet for the first time and with ξ_t^1 afterwards. The following result is the key point of Doeblin's coupling construction.

Exercise* 80. Let $(\boldsymbol{\xi}_t, \mathbb{P}^{\boldsymbol{i}})$ be a pair of independent copies of a Markov process $(\xi_t, \mathbb{P}^i)$ in $\mathcal{A}$. Prove that, for any function $f : \mathcal{A} \to \mathbb{R}$, $\boldsymbol{i} = (i_1, i_2) \in \boldsymbol{\mathcal{A}}$, and $t \in \mathbb{Z}_+$, we have

$$\mathbb{E}^{\boldsymbol{i}} f(\eta_t) = \mathbb{E}^{\boldsymbol{i}} f(\xi_t^2) = \mathbb{E}^{i_2} f(\xi_t), \tag{6.11}$$

where in the left-most term $f(\eta_t)$ should be regarded as a function of the full trajectory $\boldsymbol{\omega}_0, \dots, \boldsymbol{\omega}_t$. In particular, the law of η_t under $\mathbb{P}^{\boldsymbol{i}}$ coincides with that of ξ_t under $\mathbb{P}^{i_2}$.

Hint: Use the strong Markov property.

Exercise* 81. Let $(\xi_t, \mathbb{P}^i)$ be a Markov process with a strongly recurrent transition matrix. Prove that, for any $f : \mathcal{A} \to \mathbb{R}$, $i_1, i_2 \in \mathcal{A}$, and $t \in \mathbb{Z}_+$, we have

$$|\mathfrak{P}_t f(i_1) - \mathfrak{P}_t f(i_2)| \leq C e^{-\gamma t} |f|, \tag{6.12}$$

where γ is the number entering (6.9). Use this to give a different proof of the uniqueness of stationary measure and exponential mixing (see Exercise 60).

Hint: Use (6.11) *and* (6.9).

Index

References

[Chu67] K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, New York, 1967.

[Doe38] W. Doeblin, *Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états*, Rev. Math. Union Interbalkan **2** (1938), 77–105.

[Fel68] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. I.*, John Wiley & Sons, New York, 1968.

[Kol37] A. Kolmogoroff, *Zur Umkehrbarkeit der statistischen Naturgesetze*, Math. Ann. **113** (1937), no. 1, 766–772.

[Sek57] T. Seki, *An elementary proof of Brouwer's fixed point theorem*, Tohoku Math. J. (2) **9** (1957), 105–109.