Chapter 2

Markov Chains

Introduction

A Markov chain is a sequence of random variables $\{X_n;\ n = 0, 1, 2, \ldots\}$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking its values in a set E which could be arbitrary, but which will be for us either finite or countable, and which possesses the Markov property. Intuitively, a Markov chain has the property that, knowing the present state X_n, one can forget the past if one wants to predict the future. One way to construct a Markov chain is as follows. Let $\{Y_n,\ n \geq 1\}$ be mutually independent F–valued random variables, which are globally independent of X_0. Given a mapping $f : \mathbb{N} \times E \times F \to E$, we define $\{X_n,\ n \geq 1\}$ recursively by

$$X_n = f(n, X_{n-1}, Y_n).$$

In a way, this is the simplest model of non–mutually independent random variables.

The next two chapters will present many applications of Markov chains. Note that we shall restrict our presentation to *homogeneous* Markov chains (in the above recurrence relation, f does not depend upon n, and the Y_n's all have the same law), even though non–homogeneous chains are necessary in many applications. Even in those cases, understanding the long time behaviour of the homogeneous chains is crucial.

2.1 Definitions and elementary properties

We want to define and study Markov chains $\{X_n; n \in \mathbb{N}\}$ with values in a (finite or) countable state space E. We shall denote by $x, y, \dots$ generic points of E. We shall use the convention that whenever a condition involves a conditional probability $\mathbb{P}(A|B)$, that condition is assumed to be satisfied only when $\mathbb{P}(B) > 0$.

Definition 2.1.1. *The E–valued stochastic process $\{X_n; n \in \mathbb{N}\}$ is called a Markov chain whenever for all $n \in \mathbb{N}$, the conditional law of X_{n+1} given $X_0, X_1, \dots, X_n$ equals its conditional law given X_n, i.e $\forall x_0, \dots, x_{n+1} \in E$,*

$$\mathbb{P}(X_{n+1} = x_{n+1}|X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1}|X_n = x_n).$$

A simple criteria, which allows us in many cases to verify that a given process is a Markov chain, is given by the :

Lemma 2.1.2. *Let E and F be two countable sets, and let f be a mapping from $\mathbb{N} \times E \times F$ into E. Let $X_0, Y_1, Y_2, \dots$ be mutually independent r. v.'s, X_0 being E–valued, and the Y_n's being F–valued. Let $\{X_n,\ n \geq 1\}$ be the E–valued process defined by*

$$X_{n+1} = f(n, X_n, Y_{n+1}),\ n \in \mathbb{N}.$$

Then $\{X_n,\ n \in \mathbb{N}\}$ is a Markov chain.

Proof

$$\begin{aligned}
\mathbb{P}(X_{n+1} &= x_{n+1}|X_0 = x_0, \dots, X_n = x_n) \\
&= \frac{\mathbb{P}(X_0 = x_0, \dots, X_n = x_n, X_{n+1} = x_{n+1})}{\mathbb{P}(X_0 = x_0, \dots, X_n = x_n)} \\
&= \sum_{\{z; f(n,x_n,z) = x_{n+1}\}} \frac{\mathbb{P}(X_0 = x_0, \dots, X_n = x_n, Y_{n+1} = z)}{\mathbb{P}(X_0 = x_0, \dots, X_n = x_n)} \\
&= \sum_{\{z; f(n,x_n,z) = x_{n+1}\}} \mathbb{P}(Y_{n+1} = z) \\
&= \frac{\mathbb{P}(X_n = x_n, X_{n+1} = x_{n+1})}{\mathbb{P}(X_n = x_n)}
\end{aligned}$$

□

A Markov chain is the analogue of a deterministic sequence which is defined by a recurrence relation of the type :

$$x_{n+1} = f(n, x_n),$$

as opposed to a system "with memory", of the type :

$$x_{n+1} = f(n, x_n, x_{n-1}, \ldots, x_1, x_0).$$

Here the function $f(n, \cdot)$ is replaced by the "transition matrix" :

$$P_{xy} = \mathbb{P}(X_{n+1} = y | X_n = x).$$

From now on, this matrix $P = (P_{xy};\ x, y \in E)$ will be assumed to be independent of the time variable n. One then says that the Markov chain is **homogeneous**.

The chain constructed in Lemma 2.1.2 is homogeneous whenever f does not depend upon n, and the Y_n's all have the same law. We now state a variant of Lemma 2.1.2, whose proof is essentially identical, and which will be useful below.

Lemma 2.1.3. *Let E be a countable set, and f be a mapping from $E \times [0,1]$ into E, such that for all $x, y \in E$, the set $\{u \in [0,1];\ f(x,u) = y\}$ is a Borel subset of $[0,1]$. Let $X_0, Y_1, Y_2, \ldots$ be mutually independent r. v.'s, with X_0 taking its values in E, and the Y_n's being uniform on $[0,1]$, and let $\{X_n,\ n \geq 1\}$ be the E–valued random sequence defined by*

$$X_{n+1} = f(X_n, Y_{n+1}),\ n \in \mathbb{N}.$$

Then $\{X_n,\ n \in \mathbb{N}\}$ is a Markov chain.

The matrix P is called **Markovian** (or **stochastic**), in the sense that it has the property that $\forall x \in E$, the row vector $(P_{xy};\ y \in E)$ is a probability measure on E, or in other words :

$$P_{xy} \geq 0,\ \forall y \in E;\ \sum_{y \in E} P_{xy} = 1.$$

Remark 2.1.4. *P_{xy} is the entry in row x and column y of the matrix P. This notation may be surprising for the reader, but it is very convenient. It is more common to enumerate rows and columns, and hence to index them by $1, 2, \ldots$. We note moreover that our matrices are square matrices, with possibly an infinite number of rows and columns, in the case where $|E| = \infty$.*

As we will now see, the law of a Markov chain is entirely determined by the "initial law" $(\mu_x;\ x \in E)$, which is the law of X_0, and the transition matrix of the chain.

Definition 2.1.5. *Let μ be a probability on E, and P a Markovian matrix. An E–valued random sequence $\{X_n;\ n \in \mathbb{N}\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is called a (μ, P)–Markov chain (i.e. with initial law μ and transition matrix P) if :*

(i) $\mathbb{P}(X_0 = x) = \mu_x,\ \forall x \in E.$

(ii) $\mathbb{P}(X_{n+1} = y | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) = P_{xy},$
$\forall x_0, \ldots, x_{n-1}, x, y \in E.$

Proposition 2.1.6. *A necessary and sufficient condition for an E–valued random sequence $\{X_n,\ n \in \mathbb{N}\}$ to be a (μ, P)–Markov chain is that $\forall n \in \mathbb{N}$, the law of the r. v. $(X_0, X_1, \ldots, X_n)$ be given by*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = \mu_{x_0} P_{x_0 x_1} \times \cdots \times P_{x_{n-1} x_n}.$$

Proof Necessary Condition. If $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) > 0$, then

$$\begin{aligned}\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_n &= x_n | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) \\ &\times \cdots \times \mathbb{P}(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_0 = x_0),\end{aligned}$$

and the above identity follows from the definition. Otherwise, both sides of the identity in the statement are zero (consider the smallest index k such that $\mathbb{P}(X_0 = x_0, \ldots, X_k = x_k) = 0$).

Sufficient Condition. (i) The identity in the statement follows from the definition. Let us prove more than (ii).

$$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+p} &= x_{n+p} | X_0 = x_0, \ldots, X_n = x_n) \\ &= \frac{\mu_{x_0} P_{x_0 x_1} \times \cdots \times P_{x_{n+p-1} x_{n+p}}}{\mu_{x_0} P_{x_0 x_1} \times \cdots \times P_{x_{n-1} x_n}}\end{aligned}$$

(ii) now follows if we choose $p = 1$. □

We have in fact established :

Corollary 2.1.7. *If $\{X_n;\ n \in \mathbb{N}\}$ is a (μ, P)–Markov chain, then for all $n, p, x_0, \dots, x_{n+p}$*

$$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1}, \dots, X_{n+p} = x_{n+p} | X_0 = x_0, \dots, X_n = x_n)\\ = P_{x_n x_{n+1}} \times \cdots \times P_{x_{n+p-1} x_{n+p}}.\end{aligned}$$

A probability μ on E is considered to be a row vector, a mapping $g : E \to \mathbb{R}$ as a column vector, which justifies the notations

$$(\mu P)_y = \sum_{x \in E} \mu_x P_{xy}$$
$$(Pg)_x = \sum_{y \in E} P_{xy} g_y,$$

and the integral of a function g with respect to a measure μ is written (whenever the sum converges absolutely) as the product of a row vector on the left with a column vector on the right :

$$\mu g = \sum_{x \in E} \mu_x g_x.$$

Proposition 2.1.8. *Let $\{X_n, n \in \mathbb{N}\}$ ba a (μ, P)–Markov chain. Then*

(i) $\mathbb{P}(X_n = y | X_0 = x) = \mathbb{P}(X_{n+p} = y | X_p = x) = (P^n)_{xy}$

(ii) $\mathbb{P}(X_n = y) = (\mu P^n)_y$

(iii) $\mathbb{E}[g(X_n) | X_0 = x] = (P^n g)_x$

Proof

(i)

$$\begin{aligned}\mathbb{P}(X_n = y | X_0 = x) &= \sum_{x_1, \dots, x_{n-1}} \mathbb{P}(X_n = y, X_{n-1} = x_{n-1}, .., X_1 = x_1 | X_0 = x)\\ &= \sum_{x_1, \dots, x_{n-1}} \frac{\mu_x P_{x x_1} \times \cdots \times P_{x_{n-1} y}}{\mu_x}\\ &= \sum_{x_1, \dots, x_{n-1}} P_{x x_1} \times \cdots \times P_{x_{n-1} y}\\ &= (P^n)_{xy}\end{aligned}$$

(ii) We note that

$$\begin{aligned}\mathbb{P}(X_n = y) &= \sum_{x \in E} \mathbb{P}(X_n = y, X_0 = x) \\ &= \sum_{x \in E} \mathbb{P}(X_n = y | X_0 = x)\mu_x,\end{aligned}$$

and we use (i).

(iii) Again we use (i) starting from :

$$\mathbb{E}[g(X_n)|X_0 = x] = \sum_{y \in E} g_y \mathbb{P}(X_n = y | X_0 = x)$$

□

2.2 Examples

2.2.1 Random walk in $E = \mathbf{Z}^d$

Let $\{Y_n; n \in \mathbb{N}^*\}$ denote an i.i.d. $\mathbf{Z}^d$–valued random sequence, with the common law λ, and let X_0 be a $\mathbf{Z}^d$–valued r. v., independent of the Y_n's. Then the random sequence $\{X_n,\ n \geq 0\}$ defined by

$$X_{n+1} = X_n + Y_{n+1},\ \ n \in \mathbb{N}$$

is a (μ, P)–Markov chain, with μ=law of X_0, and $P_{xy} = \lambda_{y-x}$. The most classical case is that of the symmetric random walk starting from 0, i.e.

$$\mu = \delta_0, \quad \lambda_{\pm e_i} = \frac{1}{2d},$$

where $(e_1, \ldots, e_d)$ is an orthonormal basis of $\mathbb{R}^d$.

2.2.2 Bienaymé–Galton–Watson process

This is a branching process $\{Z_n; n \in \mathbb{N}\}$ where Z_n denotes the number of males in the n–th generation with a certain name, those individuals being all descendants of a common ancestor, the unique male in the generation 0

($Z_0 = 1$ p.s.). We assume that the i–th male from the n–th generation has ξ_i^n male children ($1 \le i \le Z_n$), in such a way that

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^n.$$

Our main assumption is that the r. v.'s $\{\xi_i^n, i = 1, 2, \ldots, n = 0, 1, 2, \ldots\}$ are i.i.d., so that in particular Z_n and $\{\xi_1^n, \ldots, \xi_p^n, \ldots\}$ are independent.

The random sequence $\{Z_n,\ n \in \mathbb{N}\}$ is a (μ, P) $\mathbb{N}$–valued Markov chain, with $\mu = \delta_1$ and

$$P_{xy} = (p^{*x})_y,$$

where p^{*x} denotes the x–th convolution power of the joint law p on $\mathbb{N}$ of the ξ_n^k's, i.e. the law of the sum of x i.i.d. r. v.'s, all having the law p.

2.2.3 A discrete time queue

We consider a queue at a counter. X_n denotes the number of customers who either are waiting or are being served at time n. Between time n and time $n+1$, Y_{n+1} new customers enter the queue, and whenever $X_n > 0$, Z_{n+1} customers leave the queue (with $Z_{n+1} = 0$ or 1). We assume that X_0, Y_1, Z_1, Y_2, $Z_2 \ldots$ are mutually independent, with $0 < \mathbb{P}(Y_n = 0) < 1$, and moreover $\mathbb{P}(Z_n = 1) = p = 1 - \mathbb{P}(Z_n = 0)$. We have

$$X_{n+1} = X_n + Y_{n+1} - \mathbf{1}_{\{X_n > 0\}} Z_{n+1}.$$

2.3 Strong Markov property

Let us first reformulate the Markov property. Let $\{X_n; n \in \mathbb{N}\}$ be an E–valued Markov chain, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Given a probability measure μ on E, we shall use the notation $\mathbb{P}_\mu$ to denote any probability on $(\Omega, \mathcal{F})$ such that under $\mathbb{P}_\mu$ the sequence $\{X_n,\ n \ge 0\}$ is a Markov chain with initial law μ, in other words μ is the law of X_0, that is

$$\mathbb{P}_\mu(X_0 = x) = \mu_x,\ x \in E.$$

Whenever $\mu = \delta_x$, we shall write $\mathbb{P}_x$ instead of $\mathbb{P}_{\delta_x}$. $\mathbb{P}_x$ can be interpreted as the conditional law of X, given that $X_0 = x$. For any $n \ge 0$, we define $\mathcal{F}_n$ as the sigma–algebra of those events which are "determined by $X_0, X_1, \ldots, X_n$", that is

$$\mathcal{F}_n = \left\{\{\omega; (X_0(\omega), \ldots, X_n(\omega)) \in B_n\},\ B_n \in \mathcal{P}(E^{n+1})\right\},$$

where we have used the following notation, which will be used again in this book : $\mathcal{P}(F)$ denotes the collection of all the subsets of F.

Theorem 2.3.1. *Let $\{X_n; n \geq 0\}$ be a (μ, P)–Markov chain. Then for any $n \in \mathbb{N}$, $x \in E$, conditionally upon $\{X_n = x\}$, $\{X_{n+p}; p \geq 0\}$ is a (δ_x, P) Markov chain, which is independent of $(X_0, \ldots, X_n)$. In other words, for all $A \in \mathcal{F}_n$ and any $m > 0$, $x_1, \ldots, x_m \in E$,*

$$\begin{aligned}&\mathbb{P}\left(A \cap \{X_{n+1} = x_1, \ldots, X_{n+m} = x_m\} | X_n = x\right)\\ &\quad = \mathbb{P}(A|X_n = x)\mathbb{P}_x(X_1 = x_1, \ldots, X_m = x_m)\end{aligned}$$

Proof It suffices to prove the result in the case where $A = \{X_0 = y_0, X_1 = y_1, \ldots, X_n = y_n\}$ (A is a finite or countable union of disjoint sets of that form, and the result in the general case will then follow from the σ–additivity of $\mathbb{P}$). It suffices to consider the case $y_n = x$, since otherwise both sides of the equality vanish. The left hand side of the identity in the statement equals

$$\frac{\mathbb{P}(X_0 = y_0, \ldots, X_n = x, X_{n+1} = x_1, \ldots, X_{n+m} = x_m)}{\mathbb{P}(X_n = x)},$$

which, applying Proposition 2.1.6 twice, is shown to equal

$$\frac{\mathbb{P}(A)}{\mathbb{P}(X_n = x)} \times P_{xx_1} \times P_{x_1x_2} \times \cdots \times P_{x_{m-1}x_m},$$

or in other words

$$\mathbb{P}(A|X_n = x)\mathbb{P}_x(X_1 = x_1, \ldots, X_m = x_m).$$

□

The preceding result says in particular that the past and the future of the chain are conditionally independent, given the position of the chain at the present time n.

We want now to extend the Markov property, replacing the fixed time n by a random time (but not any random time).

Definition 2.3.2. *A r. v. T taking values in the set $\mathbb{N} \cup \{+\infty\}$ is called a* **stopping time** *if $\forall n \in \mathbb{N}$,*

$$\{T = n\} \in \mathcal{F}_n.$$

In other words, the observation of $X_0, X_1, \ldots, X_n$, the trajectory of the chain up to time n, is enough to decide whether or not T equals n.

Example 2.3.3. *i) $\forall x \in E$, the first passage time S_x at state x :*

$$S_x = \begin{cases} \inf\{n \geq 0; X_n = x\} & \textit{if such an } n \textit{ exists}, \\ +\infty, & \textit{otherwise}; \end{cases}$$

is a stopping time, as well as the time of the first return to the state x :

$$T_x = \begin{cases} \inf\{n \geq 1; X_n = x\} & \textit{if such an } n \textit{ exists}, \\ +\infty, & \textit{otherwise.} \end{cases}$$

(With the convention that the infimum of the empty set is $+\infty$, it is sufficient to write : $T_x = \inf\{n \geq 1;\ X_n = x\}$.)

T_x is a stopping time, since

$$\{T_x = n\} = \{X_1 \neq x\} \cap \ldots \cap \{X_{n-1} \neq x\} \cap \{X_n = x\}.$$

ii) $\forall A \subset E$, the time of the first visit to the set A

$$T_A = \inf\{n \geq 1; X_n \in A\}$$

is a stopping time.
iii) On the other hand, the time of the last visit to A

$$L_A = \sup\{n \geq 1; X_n \in A\}$$

is not a stopping time, since we need to know the trajectory after time n, in order to decide whether or not $L_A = n$.

We shall denote by $\mathcal{F}_T$ the σ–algebra of events which are "determined by $X_0, X_1, \ldots, X_T$", which is defined as the σ–algebra of those events $B \in \mathcal{F}$ which are such that $\forall n \in \mathbb{N}$,

$$B \cap \{T = n\} \in \mathcal{F}_n.$$

Theorem 2.3.4. *(Strong Markov property) Let $\{X_n : n \geq 0\}$ be a (μ, P)–Markov chain, and T a stopping time. Conditionally upon $\{T < \infty\} \cap \{X_T = x\}$, $\{X_{T+n}; n \geq 0\}$ is a (δ_x, P)–Markov chain, which is independent of $\mathcal{F}_T$. In other words, for all $A \in \mathcal{F}_T$, and all $m > 0$, $x_1, \ldots, x_m \in E$,*

$$\begin{aligned} &\mathbb{P}(A \cap \{X_{T+1} = x_1, \ldots, X_{T+m} = x_m\} | X_T = x, T < \infty) \\ &= \mathbb{P}(A | X_T = x,\ T < \infty) \times \mathbb{P}_x(X_1 = x_1, \ldots, X_m = x_m) \end{aligned}$$

Proof It suffices to show that $\forall n \in \mathbb{N}$,

$$\begin{aligned} &\mathbb{P}(A \cap \{T = n\} \cap \{X_{T+1} = x_1, \ldots, X_{T+m} = x_m\} | X_T = x) \\ &= \mathbb{P}(A \cap \{T = n\} | X_T = x) \mathbb{P}_x(X_1 = x_1, \ldots, X_m = x_m), \end{aligned}$$

which follows from Theorem 2.3.1, and then to sum over all possible values of n. □

2.4 Recurrent and transient states

We define as above

$$T_x = \inf\{n \geq 1; X_n = x\}, \text{ and we state}$$

Definition 2.4.1. *$x \in E$ is said to be* **recurrent** *if $\mathbb{P}_x(T_x < \infty) = 1$, and* **transient** *otherwise (i.e. if $\mathbb{P}_x(T_x < \infty) < 1$).*

We define the number of returns to the state x :

$$N_x = \sum_{n \geq 1} \mathbf{1}_{\{X_n = x\}}$$

Proposition 2.4.2. *a) If x is recurrent,*

$$\mathbb{P}_x(N_x = +\infty) = 1.$$

b) If x is transient,

$$\mathbb{P}_x(N_x = k) = (1 - \Pi_x)\Pi_x^k,\ k \geq 0,$$

where $\Pi_x = \mathbb{P}_x(T_x < \infty)$ (in particular $N_x < \infty$, $\mathbb{P}_x$ a. s.)

Proof Let

$$\begin{aligned} T_x^2 &= \inf\{n > T_x, X_n = x\} \\ &= T_x + \inf\{n \geq 1, X_{T_x+n} = x\}. \end{aligned}$$

It is not hard to show that T_x^2 is a stopping time.

$$\begin{aligned} \mathbb{P}_x(T_x^2 < \infty) &= \mathbb{P}_x(T_x^2 < \infty | T_x < \infty)\mathbb{P}_x(T_x < \infty) \\ &= \sum_{n=1}^{\infty} \mathbb{P}_x(T_x^2 = T_x + n | T_x < \infty)\mathbb{P}_x(T_x < \infty). \end{aligned}$$

But from Theorem 2.3.4 we deduce that

$$\begin{aligned} \mathbb{P}_x(T_x^2 =& T_x + n | T_x < \infty) \\ &= \mathbb{P}_x(X_{T_x+1} \neq x, \ldots, X_{T_x+n-1} \neq x, X_{T_x+n} = x | T_x < \infty) \\ &= \mathbb{P}_x(X_1 \neq x, \ldots, X_{n-1} \neq x, X_n = x) \\ &= \mathbb{P}_x(T_x = n). \end{aligned}$$

Finally

$$\begin{aligned} \mathbb{P}_x(T_x^2 < \infty) &= (\mathbb{P}_x(T_x < \infty))^2, \text{ soit} \\ \mathbb{P}_x(N_x \geq 2) &= (\mathbb{P}_x(T_x < \infty))^2, \end{aligned}$$

and iterating the same argument, we deduce that

$$\mathbb{P}_x(N_x \geq k) = (\mathbb{P}_x(T_x < \infty))^k, \;\; k \in \mathbb{N}.$$

Both statements of the Proposition follow easily from this identity. □

Corollary 2.4.3. *x is recurrent if and only if*

$$\sum_{n=0}^{\infty} (P^n)_{xx} = +\infty$$

Proof

$$\begin{aligned} \mathbb{E}_x(N_x) &= \sum_{n\geq 1} \mathbb{P}_x(X_n = x) \\ &= \sum_{n \geq 1} (P^n)_{xx} \end{aligned}$$

It follows from the Proposition that this quantity is infinite whenever x is recurrent. On the other hand, if x is transient,

$$\begin{aligned}\mathbb{E}_x(N_x) &= \sum_{k=1}^{\infty} k(1-\Pi_x)\Pi_x^k \\ &= \frac{\Pi_x}{1-\Pi_x} < \infty\end{aligned}$$

□

Definition 2.4.4. *We say that the state y is* **accessible** *from x (denoted by $x \to y$) whenever there exists $n \geq 0$ such that $\mathbb{P}_x(X_n = y) > 0$. We say that x and y* **communicate** *(noted $\leftrightarrow$) whenever both $x \to y$ and $y \to x$.*

The relation $x \leftrightarrow y$ is an equivalence relation, and we can partition E into equivalence classes modulo the relation $\leftrightarrow$.

Note that $x \to y \Leftrightarrow \exists n \geq 0$ s. t. $(P^n)_{xy} > 0$, since $\mathbb{P}_x(X_n = y) = (P^n)_{xy}$ (Proposition 2.1.8(i)).

Theorem 2.4.5. *Let $C \subset E$ be an equivalence class for the relation $\leftrightarrow$. Then all states in C either are recurrent, or else they all are transient.*

PROOF Let $x, y \in C$. It suffices to show that x transient $\Rightarrow$ y transient (since then y recurrent $\Rightarrow$ x recurrent). Since $x \leftrightarrow y$, $\exists\ n, m > 0$ such that $(P^n)_{xy} > 0$ et $(P^m)_{yx} > 0$. But $\forall r \geq 0$,

$$(P^{n+r+m})_{xx} \geq (P^n)_{xy}(P^r)_{yy}(P^m)_{yx})$$

and

$$\sum_{r=0}^{\infty}(P^r)_{yy} \leq \frac{1}{(P^n)_{xy}(P^m)_{yx}} \sum_{n=0}^{\infty}(P^{n+r+m})_{xx} < \infty.$$

□

Definition 2.4.6. *A (μ, P)–Markov chain is said to be* **irreducible** *whenever E consists of a single equivalence class. It is said to be* **irreducible and recurrent** *if it is irreducible and all states are recurrent.*

Proposition 2.4.7. *Any irreducible Markov chain on a* **finite** *state space E is recurrent.*

PROOF Whenever E is finite, at least one state must be visited infinitely many times with positive probability, hence a.s. from Proposition 2.4.2, and that state (as well as all states) is (are) recurrent. □

2.5 The irreducible and recurrent case

In this section, we assume that the chain is both irreducible and recurrent. We start by studying the **excursions** of the chain between two successive returns to the state x :

$$\mathcal{E}_k = (X_{T_x^k}, X_{T_x^k+1}, \dots, X_{T_x^{k+1}}),\ k \geq 0.$$

These excursions are random sequences whose length is random and finite ≥ 2, composed of elements of $E \backslash \{x\}$, except for the first and the last one, which are equal to x. Denote by U the set of sequences

$$u = (x, x_1, \dots, x_n, x),$$

with $n \geq 1$, $x_\ell \neq x$, $1 \leq \ell \leq n$. U is countable, and it is the set of all possible excursions $\mathcal{E}_0, \mathcal{E}_1, \dots$. Hence these r. v.'s take their values in a countable set, and their probability law is characterized by the quantities

$$\mathbb{P}(\mathcal{E}_k = u),\ \ u \in U.$$

Proposition 2.5.1. *Under $\mathbb{P}_x$, the sequence $(\mathcal{E}_0, \mathcal{E}_1, \dots)$ of excursions is i.i.d., in other words there exists a probability $\{p_u,\ u \in U\}$ on U such that for all $k > 0$, $u_0, \dots, u_k \in U$,*

$$\mathbb{P}_x(\mathcal{E}_0 = u_0, \mathcal{E}_1 = u_1, \dots, \mathcal{E}_k = u_k) = \prod_{\ell=0}^{k} p_{u_\ell}.$$

PROOF This is a consequence of the strong Markov property. Indeed, $\{\mathcal{E}_0 = u_0\} \in \mathcal{F}_{T_x}$, and the event

$$\{\mathcal{E}_1 = u_1, \dots, \mathcal{E}_k = u_k\}$$

is of the form

$$\{X_{T_x+1} = x_1, \dots, X_{T_x+p} = x_p\},$$

for some $p > 0$, $x_1, \dots, x_p \in E$. Consequently

$$\begin{aligned}
&\mathbb{P}_x(\mathcal{E}_0 = u_0, \mathcal{E}_1 = u_1, \dots, \mathcal{E}_k = u_k) \\
&= \mathbb{P}_x(\{\mathcal{E}_0 = u_0\} \cap \{X_{T_x+1} = x_1, \dots, X_{T_x+p} = x_p\} | T_x < \infty) \\
&= \mathbb{P}_x(\mathcal{E}_0 = u_0)\mathbb{P}_x(X_1 = x_1, \dots, X_p = x_p) \\
&= \mathbb{P}_x(\mathcal{E}_0 = u_0)\mathbb{P}_x(\mathcal{E}_0 = u_1, \dots, \mathcal{E}_{k-1} = u_k) \\
&= \mathbb{P}_x(\mathcal{E}_0 = u_0)\mathbb{P}_x(\mathcal{E}_0 = u_1) \times \dots \times \mathbb{P}_x(\mathcal{E}_0 = u_k) \\
&= p_{u_0} p_{u_1} \times \cdots \times p_{u_k},
\end{aligned}$$

where $\{p_u, u \in U\}$ is the law of $\mathcal{E}_0$ under $\mathbb{P}_x$. □

A measure on the set E is a "row vector" $\{\gamma_x; x \in E\}$ such that $0 \le \gamma_x < \infty$, $\forall x$. Whenever the measure is finite, $\sum_{x \in E} \gamma_x < \infty$, we can normalize it, to make it a probability on E, $\left(\frac{\gamma_x}{\sum_z \gamma_z}, \ x \in E \right)$. A measure γ is said to be invariant (with respect to the transition matrix P) whenever

$$\gamma P = \gamma, \text{ i.e.}$$

$$\sum_{y \in E} \gamma_y P_{yx} = \gamma_x, \ x \in E.$$

A measure γ is said to be strictly positive if $\gamma_x > 0, \ \forall x \in E$.

A probability measure γ is invariant iff the chain (γ, P) has the property that γ is the law of X_n, $\forall n \in \mathbb{N}$, hence $\forall n$, $\{X_{n+m}; m \in \mathbb{N}\}$ is a (γ, P)–Markov chain.

Remark 2.5.2. *An invariant probability is a probability π which satifies $\pi P = \pi$, or equivalently $\forall x \in E$,*

$$\sum_{y \neq x} \pi_y P_{yx} = \pi_x (1 - P_{xx}),$$

that is

$$\mathbb{P}(X_n \neq x, X_{n+1} = x) = \mathbb{P}(X_n = x, X_{n+1} \neq x),$$

which means that at equilibrium, the mean number of departures from the state x between time n and time $n+1$ equals the mean number of arrivals at the state x between time n and time $n+1$. The relation which characterizes the invariant probability is very intuitive.

Theorem 2.5.3. *Let $\{X_n; n \in \mathbb{N}\}$ be a Markov chain with transition matrix P, which we assume to be irreducible and recurrent. Then there exists a strictly positive invariant measure γ, which is unique up to a multiplicative constant.*

Proof Existence Let γ_y^x denote the mean number of visits to the state y during the excursion $\mathcal{E}_0$ starting from x, that is

$$\begin{aligned}
\gamma_y^x &= \mathbb{E}_x \sum_{n=1}^{T_x} \mathbf{1}_{\{X_n=y\}} \\
&= \sum_{n=1}^{\infty} \mathbb{P}_x(X_n = y, n \le T_x) \\
&= \sum_{z\in E}\sum_{n=1}^{\infty} \mathbb{P}_x(\{X_{n-1} = z, n-1 < T_x\} \cap \{X_n = y\}) \\
&= \sum_{z\in E}\left(\sum_{n=2}^{\infty} \mathbb{P}_x(X_{n-1} = z, n-1 \le T_x)\right) P_{zy} \\
&= (\gamma^x P)_y.
\end{aligned}$$

Note that we have used recurrence to obtain the for to last equality. We now exploit the irreducibility of the chain. $\exists\ n, m$ such that $(P^n)_{xy} > 0$, $(P^m)_{yx} > 0$. Hence, since $\gamma_x^x = 1$,

$$\begin{aligned}
0 < (P^n)_{xy} = \gamma_x^x (P^n)_{xy} \le (\gamma^x P^n)_y = \gamma_y^x \\
\gamma_y^x (P^m)_{yx} \le (\gamma^x P^m)_x = \gamma_x^x = 1.
\end{aligned}$$

Consequently γ^x is a strictly positive measure, which satisfies $\gamma_x^x = 1$.
Uniqueness Let λ denote an invariant measure such that $\lambda_x = 1$. We shall first prove that $\lambda \ge \gamma^x$, then that $\lambda = \gamma^x$. Note that this part of the proof of the theorem exploits only irreducibility (and not recurrence).

$$
\begin{aligned}
\lambda_y &= P_{xy} + \sum_{z_1 \neq x} \lambda_{z_1} P_{z_1 y} \\
&= P_{xy} + \sum_{z_1 \neq x} P_{xz_1} P_{z_1 y} + \sum_{z_1, z_2 \neq x} \lambda_{z_2} P_{z_2 z_1} P_{z_1 y} \\
&\geq \sum_{n=0}^{\infty} \sum_{z_1, \ldots, z_n \neq x} P_{xz_n} P_{z_n z_{n-1}} \times \cdots \times P_{z_1 y} \\
&= \sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1} = y, T_x \geq n+1) \\
&= \gamma_y^x.
\end{aligned}
$$

Hence $\mu = \lambda - \gamma^x$ is also an invariant measure, and $\mu_x = 0$. Let $y \in E$, and n be such that $(P^n)_{yx} > 0$. Then

$$0 = \mu_x = \sum_{z \in E} \mu_z (P^n)_{zx} \geq \mu_y (P^n)_{yx}.$$

Hence $\mu_y = 0$, and this holds $\forall y \in E$. □

We have seen that a state x is recurrent whenever

$$\mathbb{P}_x(T_x < \infty) = 1.$$

Let $m_x = \mathbb{E}_x(T_x)$.

If this quantity is finite, then x is called **positive recurrent**, and otherwise it is called **null recurrent** .

Theorem 2.5.4. *Assume again that the chain is irreducible. A state x is positive recurrent iff all the states are positive recurrent, iff there exists an invariant probability π, with $\pi = (\pi_x = m_x^{-1},\ x \in E)$.*

Proof Note that

$$m_x = \sum_{y \in E} \gamma_y^x$$

Hence if x is positive recurrent, then the probability $\pi = (\pi_y = \dfrac{\gamma_y^x}{m_x},\ y \in E)$ is an invariant probability.

Conversely, if π is an invariant probability, from the irreducibility (see the end of the proof of existence in Theorem 2.5.3), π is strictly positive, hence if x is an arbitrary state, $\lambda = \left(\lambda_y = \frac{\pi_y}{\pi_x},\ y \in E\right)$ is an invariant measure which satisfies $\lambda_x = 1$. From the irreducibility and the proof of uniqueness in Theorem 2.5.3,

$$m_x = \sum_{y\in E} \gamma_y^x = \sum_{y\in E} \frac{\pi_y}{\pi_x} = \frac{1}{\pi_x} < \infty.$$

Hence x, as well as all the states, is positive recurrent. □

The following dichotomy follows from the two preceding Theorems : in the *irreducible and recurrent* case, the chain is *positive recurrent* whenever there exists an *invariant probability*, *null recurrent* if *one (hence all) invariant measure(s)* has *infinite* total mass ($\sum_i \pi_i = +\infty$). In particular, if $|E| < \infty$, there do not exist null recurrent states, rather, any recurrent state is positive recurrent.

Corollary 2.5.5. *Let $\{X_n\}$ be an irreducible Markov chain which is positive recurrent. To any $x \in E$, we associate $T_x = \inf\{n > 0,\ X_n = x\}$. Then for all $y \in E$,*

$$\mathbb{E}_y(T_x) < \infty.$$

Proof Note that

$$T_x \geq T_x \mathbf{1}_{\{T_y < T_x\}},$$

whence taking the expectation under $\mathbb{P}_x$,

$$m_x \geq \mathbb{E}_x(T_x|T_y < T_x)\mathbb{P}_x(T_y < T_x).$$

But it follows from the strong Markov property that $\mathbb{E}_x(T_x|T_y < T_x) > \mathbb{E}_y(T_x)$, and from the irreducibility that $\mathbb{P}_x(T_y < T_x) > 0$. The result has been established. □

Remark 2.5.6. ***The nonirreducible case***. *For simplicity, we consider here only the case $|E| < \infty$. There exists at least one recurrent class (which is positive recurrent), hence there exists a least one invariant probability. Any invariant probability charges only recurrent states. If there is only one* *recurrent class, then the chain possesses one and only one invariant probability. Otherwise, to each recurrent class we can associate a unique invariant probability whose support is that class, and all invariant measures are convex linear combinations of these, which are the extremal ones. Hence as soon as there are at least two different recurrent classes, there is an uncountable number of invariant probabilities.*

We restrict ourself again to the irreducible case. We can now establish the ergodic theorem, which is a generalization of the law of large numbers.

Theorem 2.5.7. *Suppose that the chain is irreducible and positive recurrent. Let $\pi = (\pi_x,\ x \in E)$ denote its unique invariant probability. If $f : E \to \mathbb{R}$ is bounded, then $\mathbb{P}$ a. s., as $n \to \infty$,*

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \to \sum_{x\in E} \pi_x f(x).$$

PROOF By assumption, there exists c such that $|f(x)| \le c$, $\forall x \in E$.

Let

$$N_x(n) = \sum_{1\le k\le n} \mathbf{1}_{\{X_k=x\}}$$

denote the number of returns to the state x before time n. We want to study the limit as $n \to \infty$ of

$$\frac{N_x(n)}{n}.$$

Let $S_x^0, S_x^1, \dots, S_x^k, \dots$ denote the lengths of the excursions $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_k, \dots$ starting from x. We have

$$S_x^0 + \cdots + S_x^{N_x(n)-1} \le n < S_x^0 + \cdots + S_x^{N_x(n)}.$$

Hence

$$\frac{S_x^0 + \cdots + S_x^{N_x(n)-1}}{N_x(n)} \le \frac{n}{N_x(n)} \le \frac{S_x^0 + \cdots + S_x^{N_x(n)}}{N_x(n)}$$

But since the r. v.'s $\mathcal{E}_k$ are i. i. d. (hence the same is true for the S_x^k's), as $n \to \infty$,

$$\frac{S_x^0 + \cdots + S_x^{N_x(n)}}{N_x(n)} \to \mathbb{E}_x(T_x) = m_x \ \ \mathbb{P}_x \text{ a. s.},$$

since $N_x(n) \to +\infty$ $\mathbb{P}_x$ a. s.. Again from the law of large numbers,

$$\frac{n}{N_x(n)} \to m_x \ \mathbb{P}_x \text{ a. s., that is}$$

$$\frac{N_x(n)}{n} \to \frac{1}{m_x} \ \mathbb{P}_x \text{ a. s.}$$

This convergence is also true $\mathbb{P}_\mu$ a. s., for any initial law μ , since the limit of $\frac{N_x(n)}{n}$ is the same for the chain $\{X_n; n \geq 0\}$ and for the chain $\{X_{T_x+n}; n \geq 0\}$.

Let now $F \subset E$. We define $\bar{f} = \sum_{x \in E} \pi_x f(x)$, $c = \sup_x |f(x)|$.

$$\begin{aligned}
\left| \frac{1}{n} \sum_{k=1}^{n} f(X_k) - \bar{f} \right| &= \left| \sum_{x \in E} \left(\frac{N_x(n)}{n} - \pi_x \right) f(x) \right| \\
&\leq c \sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| + c \sum_{x \notin F} \left(\frac{N_x(n)}{n} + \pi_x \right) \\
&= c \sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| + c \sum_{x \in F} \left(\pi_x - \frac{N_x(n)}{n} \right) + 2c \sum_{x \notin F} \pi_x \\
&\leq 2c \sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| + 2c \sum_{x \notin F} \pi_x
\end{aligned}$$

We choose a finite F such that $\sum_{x \notin F} \pi_x \leq \frac{\varepsilon}{4c}$, and then $N(\omega)$ such that $\forall n \geq N(\omega)$,

$$\sum_{x \in F} \left| \frac{N_x(n)}{n} - \pi_x \right| \leq \frac{\varepsilon}{4c},$$

which proves the result. □

We shall state a central limit theorem in the next section.

2.6 The aperiodic case

We have just shown that in the irreducible, positive recurrent case,

$$\frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k = y\}} \to \pi_y \text{ a. s.},$$

as $n \to \infty$. Taking the expectation under $\mathbb{P}_x$, we deduce that

$$\frac{1}{n}\sum_{k=1}^{n}(P^k)_{xy} \to \pi_y, \ \ \forall x, y \in E.$$

We see that the Cesaro means of the $(P^k)_{xy}$'s converge. This raises the natural question : is it true under the above assumptions that as $n \to \infty$,

$$(P^n)_{xy} \to \pi_y, \ \ \forall x, y \in E \quad ?$$

It is easily seen this is not the case

Consider a random walk on $E = \mathbf{Z}/N$, where N is an even integer (we identifiy 0 and N)

$$X_n = X_0 + Y_1 + \cdots + Y_n,$$

with the Y_n's i. i. d. with values in $\{-1, 1\}$, in other words

$$X_n = (X_0 + Y_1 + \cdots + Y_n) \bmod N.$$

This chain is irreducible, and positive recurrent since E is finite. But $(P^{2k+1})_{xx} = 0$, for all $x \in E$. In the particular case $N = 2$, we have $P^{2k} = I$ and $P^{2k+1} = P$.

In order for the desired convergence to be true, we need an additional assumption :

Definition 2.6.1. *A state $x \in E$ is said to be* **aperiodic** *if $\exists N$ such that*

$$(P^n)_{xx} > 0, \quad \text{for all } n \geq N.$$

Lemma 2.6.2. *If P is irreducible and there exists an aperiodic state x, then $\forall y, z \in E$, $\exists M$ such that $(P^n)_{yz} > 0$, $\forall n \geq M$. In particular, all states are aperiodic.*

Proof From the irreducibility, $\exists r, s \in \mathbb{N}$ such that $(P^r)_{yx} > 0$, $(P^s)_{xz} > 0$. Moreover

$$\left(P^{r+n+s}\right)_{yz} \geq (P^r)_{yx}(P^n)_{xx}(P^s)_{xz} > 0,$$

as soon as $n \geq N$. Hence we have the desired property with $M = N + r + s$. □

Remark 2.6.3. *Suppose we are in the irreducible, positive recurrent case. Let π be the invariant probability, so that $\pi_y > 0$, $\forall y \in E$. Hence the fact that there exists N such that $\forall n \geq N$, $(P^n)_{xy} > 0$ is a necessary condition for the convergence $(P^n)_{xy} \to \pi_y$ to hold. We shall now see that it is a sufficient condition.*

Theorem 2.6.4. *Suppose that P is irreducible, positive recurrent and aperiodic. Let π denote the unique invariant probability. If $\{X_n; n \in \mathbb{N}\}$ be a (μ, P)–Markov chain, $\forall y \in E$,*

$$\mathbb{P}(X_n = y) \to \pi_y, \;\; n \to \infty,$$

in other words

$$(\mu P^n)_y \to \pi_y,$$

for any initial law μ. In particular, $\forall x, y \in E$,

$$(P^n)_{xy} \to \pi_y.$$

Proof We shall use a coupling argument. Let $\{Y_n, n \in \mathbb{N}\}$ be a (π, P)–Markov chain, independent of $\{X_n; n \in \mathbb{N}\}$, and $x \in E$ be arbitrary. Let

$$T = \inf\{n \geq 0; \;\; X_n = Y_n = x\}$$

Step 1 We show that $\mathbb{P}(T < \infty) = 1$.

$\{W_n = (X_n, Y_n); n \in \mathbb{N}\}$ is an $E \times E$–valued Markov chain, with initial law λ (where $\lambda_{(x,u)} = \mu_x \pi_u$), and transition matrix $\tilde{P}_{(x,u)(y,v)} = P_{xy} P_{uv}$. Since P is aperiodic, $\forall x, u, y, v$, for all n large enough

$$(\tilde{P}^n)_{(x,u)(y,v)} = (P^n)_{xy} (P^n)_{uv} > 0.$$

Hence $\tilde{P}$ is irreducible. Moreover, $\tilde{P}$ possesses an invariant probability

$$\tilde{\pi}_{(x,u)} = \pi_x \pi_u.$$

Hence, from Theorem 2.5.4, $\tilde{P}$ is positive recurrent. T is the first passage time of the chain $\{W_n\}$ at the point (x, x) ; it is finite a. s.

Step 2 Define

$$Z_n = \begin{cases} X_n, & n \leq T; \\ Y_n, & n > T. \end{cases}$$

From the strong Markov property, both processes $\{X_{T+n};\ n \geq 0\}$ and $\{Y_{T+n}); n \geq 0\}$ are (δ_x, P)–Markov chains, independent of $(X_0, \ldots, X_T)$. Consequently, $\{Z_n, n \in \mathbb{N}\}$ is, as well as $\{X_n\}$, a (μ, P)–Markov chain.

Step 3 We now conclude. We have the three identities

$$\begin{aligned}
\mathbb{P}(Z_n = y) &= \mathbb{P}(X_n = y) \\
\mathbb{P}(Y_n = y) &= \pi_y \\
\mathbb{P}(Z_n = y) &= \mathbb{P}(X_n = y, n \le T) + \mathbb{P}(Y_n = y, n > T)
\end{aligned}$$

Hence

$$\begin{aligned}
|\mathbb{P}(X_n = y) - \pi_y| &= |\mathbb{P}(Z_n = y) - \mathbb{P}(Y_n = y)| \\
&\le \mathbb{P}(n < T) \\
&\to 0,
\end{aligned}$$

as $n \to \infty$. □

Remark 2.6.5. *One can define the period of a state $x \in E$ as the biggest common divisor of the integers n such that $(P^n)_{xx} > 0$. One can show with an argument very close to that of Lemma 2.6.2 that whenever P is irreducible, all states have the same period. A state is said to be aperiodic if its period is 1. The equivalence of the two definitions of aperiodicity is proved in exercise 2.10.6 below.*

We now make precise the speed of convergence in the preceding theorem, under an additional assumption, called Doeblin's condition :
$\exists\ n_0 \in \mathbb{N}$, $\beta > 0$ and a probability ν on E such that

$$(D) \quad (P^{n_0})_{xy} \ge \beta \nu_y, \ \forall x, y \in E.$$

Remark 2.6.6. *Condition (D) is equivalent to the condition*

$$\exists x \in E,\ n_0 \ge 1 \ \text{ such that } \inf_{y \in E} (P^{n_0})_{yx} > 0.$$

This implies that this state x is aperiodic. But it does not imply irreductibility (it is easy to construct a counterexample). We shall see in exercise 2.10.4 that this condition implies existence of a unique recurrence class, and of a unique invariant probability.

Lemma 2.6.7. *If P is irreducible and aperiodic, and E is finite, then condition (D) is satisfied.*

PROOF Choose $x \in E$. $\forall y \in E$, $\exists n_y$ such that $n \geq n_y \Rightarrow (P^n)_{yx} > 0$. Let $\bar{n} = \sup_{y \in E} n_y$, $\alpha = \inf_y (P^{\bar{n}})_{yx}$. Then $\alpha > 0$, and $\forall y \in E$,

$$(P^{\bar{n}})_{yx} \geq \alpha.$$

Hence condition (D) is satisfied with $n_0 = \bar{n}$, $\beta = \alpha$, $\nu = \delta_x$. $\square$

On the other hand, Doeblin's condition is rarely satisfied in the case $\text{card}E = +\infty$, since then typically $\forall n \in \mathbb{N}$, $y \in E$,

$$\inf_{x \in E} (P^n)_{xy} = 0.$$

Theorem 2.6.8. *Suppose that P is irreducible and satisfies Doeblin's condition (D). Then P is aperiodic, positive recurrent, and if π denotes its invariant probability,*

$$\sum_{y \in E} |(P^n)_{xy} - \pi_y| \leq 2\,(1-\beta)^{[n/n_0]}, \ \ \forall x \in E, \ \ n \in \mathbb{N},$$

where $[n/n_0]$ stands for the integer part of n/n_0.

Let us first introduce a tool which will be useful in the proof of this theorem.

Definition 2.6.9. *A coupling of two probabilities p and q on E is any pair (X, Y) of E–valued r. v.'s, such that p is the law of X and q is the law of Y.*

Lemma 2.6.10. *Let p and q denote two probabilities on E. We have the identity*

$$||p-q||_1 = 2 \inf_{(X,Y) \text{ coupling of } p,\,q} \mathbb{P}(X \neq Y).$$

PROOF First note that whenever (X, Y) is a coupling of p and q,

$$\begin{aligned}\mathbb{P}(X = Y) &= \sum_{x \in E} \mathbb{P}(X = Y = x) \\ &\leq \sum_{x \in E} p_x \wedge q_x,\end{aligned}$$

whence

$$\begin{aligned}\mathbb{P}(X \neq Y) &\geq 1 - \sum_{x \in E} p_x \wedge q_x \\ &= \sum_{x \in E} (p_x - q_x)^+,\end{aligned}$$

and

$$\begin{aligned}\|p - q\|_1 &= \sum_{x \in E} |p_x - q_x| \\ &\leq 2\mathbb{P}(X \neq Y).\end{aligned}$$

On the other hand, define $\alpha = \sum_{x \in E} p_x \wedge q_x$. If ξ, U, V and W are mutually independent r. v.'s, satisfying $\mathbb{P}(\xi = 1) = 1 - \mathbb{P}(\xi = 0) = \alpha$, the law of U is r defined by $r_x = \alpha^{-1} p_x \wedge q_x$, the law of V is $\bar{p}$ defined by $\bar{p}_x = (1-\alpha)^{-1}(p_x - q_x)^+$, and the law of W is $\bar{q}$ defined by $\bar{q}_x = (1-\alpha)^{-1}(q_x - p_x)^+$, then

$$\begin{aligned}X &= \xi U + (1 - \xi)V, \\ Y &= \xi U + (1 - \xi)W\end{aligned}$$

is a coupling (X, Y) of p and q, such that $2\mathbb{P}(X \neq Y) = \|p - q\|_1$. □

Proof of theorem 2.6.8 The chain being irreducible, Doeblin's condition (D) clearly implies that it is aperiodic.

Step 1 We first show that for any two probabilities μ and ν on E,

$$\|\mu P^n - \nu P^n\|_1 \leq 2(1 - \beta)^{[n/n_0]}. \tag{2.1}$$

To prove this, from Lemma 2.6.10, it suffices to construct a coupling (X_n, Y_n) of the probabilities μP^n and νP^n such that

$$\mathbb{P}(X_n \neq Y_n) \leq (1 - \beta)^{[n/n_0]}.$$

Suppose that $n = kn_0 + m$, with $m < n_0$. Given (X_0, Y_0) with the law $\mu \times \nu$ on $E \times E$, for $\ell = 0, 1, \ldots, k-1$, we define $(X_{(\ell+1)n_0}, Y_{(\ell+1)n_0})$ in terms of $(X_{\ell n_0}, Y_{\ell n_0})$ as follows. Given a sequence $\{\xi_\ell, U_\ell, V_\ell,\ \ell \geq 0\}$ of mutually indpendent r. v. 's, the ξ_ℓ's being Bernoulli with $\mathbb{P}(\xi_\ell = 1) = \beta = 1 - \mathbb{P}(\xi_\ell = 0)$, the law of the U_ℓ's being $\bar{m} = \beta^{-1} m$ and the V_ℓ's uniform on $[0, 1]$. Define

$$Q_{xy} = (1 - \beta)^{-1}((P^{n_0})_{xy} - m_y),$$

and $f : E \times [0,1] \to E$ such that for all $x, y \in E$, $\{u;\ f(x,u) = y\}$ is a Borel subset of $[0,1]$, and provided V is uniform on $[0,1]$, the law of $f(x,V)$ is $Q_{x\cdot}$, $x \in E$. We now let

$$\begin{aligned} X_{(\ell+1)n_0} &= \xi_\ell U_\ell + (1-\xi_\ell) f(X_{\ell n_0}, V_\ell) \\ Y_{(\ell+1)n_0} &= \xi_\ell U_\ell + (1-\xi_\ell) f(Y_{\ell n_0}, V_\ell). \end{aligned}$$

Note that we have really constructed a coupling $(X_{\ell n_0}, Y_{\ell n_0})$ of $\mu P^{\ell n_0}$ and $\nu P^{\ell n_0}$, for $\ell = 0, \dots, k$, which is such that

$$\mathbb{P}(X_{\ell n_0} \neq Y_{\ell n_0}) \le \mathbb{P}(\cap_{m=0}^{\ell} \xi_m = 0) = (1-\beta)^\ell.$$

It remains to construct a coupling (X_n, Y_n) of μP^n and νP^n, such that $\{X_n \neq Y_n\} \subset \{X_{kn_0} \neq Y_{kn_0}\}$, which is easy.

Step 2 We now show that for any probability μ on E, $\{\mu P^n,\ n \ge 0\}$ is a Cauchy sequence in the Banach space $\ell^1(E)$. If $\nu = \mu P^m$, it follows from (2.1) that

$$\|\mu P^{n+m} - \mu P^n\|_1 = \|\nu P^n - \mu P^n\|_1 \le 2c^{n-n_0},$$

where $c = (1-\beta)^{1/n_0}$. The result follows.

Step 3 It follows from the second step that the sequence of probabilities $\{\mu P^n,\ n \ge 0\}$ converges in $\ell^1(E)$, towards a probability π on E. But

$$\pi P = \lim_{n\to\infty} \mu P^{n+1} = \pi,$$

hence π is invariant, and the chain is positive recurrent. Consequently, from (2.1), for any probability μ on E,

$$\|\mu P^n - \pi\|_1 \le 2(1-\beta)^{[n/n_0]},$$

which establishes the claimed speed of convergence, together with aperiodicity. $\square$

We now state a central limit theorem for irreducible, positive recurrent and aperiodic Markov chains. Such a chain, if it also satisfies

$$\sum_{y\in E} |(P^n)_{xy} - \pi_y| \le M t^n, \quad x \in E,\ n \in \mathbb{N}$$

with $M \in \mathbb{R}$ and $0 < t < 1$, is said to be *uniformly ergodic*. We have just shown that Doeblin's condition implies uniform ergodicity. That property implies the central limit theorem.

Theorem 2.6.11. *Let $\{X_n; n \in \mathbb{N}\}$ be an E–valued Markov chain, with an irreducible transition matrix P, which is moreover uniformly ergodic and aperiodic. Let π denote the unique invariant probability of the chain, and $f : E \to \mathbb{R}$ be such that*

$$\sum_{x\in E} \pi_x f^2(x) < \infty \quad \text{et} \quad \sum_{x\in E} \pi_x f(x) = 0.$$

Then as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_1^n f(X_k) \text{ converges in law to } \sigma_f Z,$$

where $Z \simeq N(0,1)$ and

$$\begin{aligned}\sigma_f^2 &= \sum_{x\in E} \pi_x (Qf)_x^2 - \sum_x \pi_x (PQf)_x^2 \\ &= 2\sum_x \pi_x (Qf)_x f_x - \sum_x \pi_x f_x^2,\end{aligned}$$

with

$$(Qf)_x = \sum_{n=0}^{\infty} \mathbb{E}_x[f(X_n)], \;\; x \in E.$$

Note that the uniform ergodicity property implies that the series which defines the operator Q converges. The reader may consult [28], Corollary 5 and the references in that paper, for a proof, and other conditions under which the theorem holds. One of the other versions (without the uniform ergodicity, but with a stronger moment condition on f) is established in [13], Theorem 4.3.18.

2.7 Reversible Markov chain

Consider the irreducible, positive recurrent case. The formulation of the Markov property "past and future are conditionally independent given the present" tells us that whenever $\{X_n; n \in \mathbb{N}\}$ is a Markov chain, it follows that $\forall N$, $\{\hat{X}_n^N = X_{N-n}; 0 \le n \le N\}$ is also a Markov chain. In general, the time reversed chain is not homogeneous, except if $\{X_n\}$ is initialized with its invariant probability π.

Proposition 2.7.1. *Let $\{X_n; n \in \mathbb{N}\}$ be a (π, P)–Markov chain whose transition matrix P is supposed to be irreducible. Denote by π its invariant probability. Then the time–reversed chain $\{\hat{X}_n^N; 0 \le n \le N\}$ is a $(\pi, \hat{P})$–Markov chain, with*

$$\pi_y \hat{P}_{yx} = \pi_x P_{xy}, \ \forall\, x, y \in E$$

PROOF

$$\begin{aligned}
&\mathbb{P}(\hat{X}_{p+1} = x | \hat{X}_p = y) \\
&= \mathbb{P}(X_n = x | X_{n+1} = y) \\
&= \mathbb{P}(X_{n+1} = y | X_n = x) \times \frac{\mathbb{P}(X_n = x)}{\mathbb{P}(X_{n+1} = y)}.
\end{aligned}$$

□

We say that the chain $\{X_n; n \in \mathbb{N}\}$ is **reversible** if $\hat{P} = P$, which holds iff the following *detailed balance equation* is satisfied :

$$\pi_x P_{xy} = \pi_y P_{yx}, \ \forall\, x, y \in E,$$

where π denotes the invariant probability. It is easily checked that whenever a probability π satisfies this relation, then it is P–invariant. The converse need not be true.

Remark 2.7.2. *If π is the invariant probability of an irreducible (and hence also positive recurent) Markov chain, the chain need not be reversible. Suppose that card$E \ge 3$. Then there may exist $x \ne y$ such that $P_{xy} = 0 \ne P_{yx}$. Consequently $\pi_x P_{xy} = 0 \ne \pi_y P_{yx}$. The transitions from y to x of the original chain correspond to the transitions from x to y of the time reversed chain, hence $P_{yx} \ne 0 \Rightarrow \hat{P}_{xy} \ne 0$, whence $\hat{P} \ne P$.*

Remark 2.7.3. *Given the transition matrix P of an irreducible positive recurrent Markov chain, one might like to compute its invariant probability. This problem is not always solvable.*

Another problem, which will appear in the next chapter, is to determine an irreducible transition matrix P whose associated Markov chain admits a given probability π as its invariant probability.

The second problem is rather easy to solve. In fact there are always many solutions. The simplest way to solve this problem is to look for P such that *the associated chain is reversible with respect to π. In other words, it suffices to find an irreducible transition matrix P such that the quantity $\pi_x P_{xy}$ be symmetric in x, y.*

In order to solve the first problem, one can try to find π such that

$$\pi_x P_{xy} = \pi_y P_{yx}, \quad \forall x,\, y \in E,$$

which, unlike solving $\pi P = \pi$, implies no summation with respect to x. But that equation has a solution only if the chain is reversible with respect to its unique invariant probability measure, which need not be the case.

Suppose now that we are given a pair (P, π), and that we want to check whether or not π is the invariant probability of the chain with the irreducible transition matrix P. If the quantity $\pi_x P_{xy}$ is symmetric in x, y, then the answer is yes, and we have an additional property, namely the reversibility. If this is not the case, one needs to check whether or not $\pi P = \pi$. One way to do that verification is given by the next Proposition, whose elementary proof is left to the reader.

Proposition 2.7.4. *Let P be an irreducible transition matrix, and π a strictly positive probability on E. For each pair x, $y \in E$, we define*

$$\hat{P}_{xy} = \begin{cases} \frac{\pi_y}{\pi_x} P_{yx}, & \text{if } x \neq y, \\ P_{xx}, & \text{if } x = y. \end{cases}$$

π is the invariant probability of the chain having the transition matrix P, and $\hat{P}$ is the transition matrix of the time–reversed chain iff for all $x \in E$,

$$\sum_{y \in E} \hat{P}_{xy} = 1.$$

2.8 Speed of convergence to equilibrium

Suppose we are in the irreducible, positive recurrent and aperiodic case. We then know that for all $x, y \in E$, $(P^n)_{x,y} \to \pi_y$ as $n \to \infty$, where π denotes the unique invariant probability measure. More generally, we expect that for a large class of functions $f : E \to \mathbb{R}$, $(P^n f)_x \to \langle f, \pi \rangle$ as $n \to \infty$ for all $x \in E$, where here and below

$$\langle f, \pi \rangle = \sum_{x \in E} f(x) \pi_x.$$

In this section, we want to discuss at which speed the above convergence holds.

2.8.1 The reversible finite state case

Let us first consider the simplest case, i. e. assume that E is finite (we write $d = |E|$) and that the process is reversible. We first note that we can identify $L^2(\pi)$ with $\mathbb{R}^d$, equipped with the product

$$\langle f, g\rangle_\pi = \sum_{x\in E} f(x)g(x)\pi_x.$$

Next the reversibility of P is equivalent to the fact that P, as an element of $\mathcal{L}(L^2(\pi))$, is a selfadjoint operator, in the sense that

$$\begin{aligned}\langle Pf, g\rangle_\pi &= \sum_{x,y\in E} P_{x,y}f(y)g(x)\pi_x \\ &= \sum_{x,y\in E} P_{y,x}f(y)g(x)\pi_y \\ &= \langle f, Pg\rangle_\pi,\end{aligned}$$

where we have used the detailed balance equation for the second identity. We now check that the operator norm of P, as an element of $\mathcal{L}(L^2(\pi))$, is at most one. Indeed, if $\|\cdot\|_\pi$ denotes the usual norm in $L^2(\pi)$,

$$\begin{aligned}\|Pf\|_\pi^2 &= \sum_{x\in E} \left[(Pf)_x\right]^2 \pi_x \\ &= \sum_{x\in E} \left(\mathbb{E}[f(X_t)|X_0 = x]\right)^2 \pi_x \\ &\leq \mathbb{E}[f^2(X_t)|X_0 = x]\pi_x \\ &= \sum_{x\in E} f^2(x)\pi_x,\end{aligned}$$

where we have used Schwartz's (or equivalently Jensen's) inequality for the inequality, and the invariance of π for the last identity.

In order to be able to work in $\mathbb{R}^d$ equipped with the Euclidean norm, let us introduce the new $d \times d$ matrix

$$\tilde{P}_{x,y} := \sqrt{\frac{\pi_x}{\pi_y}} P_{x,y}.$$

In matrix notation, $\tilde{P} = \Pi^{1/2} P \Pi^{-1/2}$, where $\Pi_{x,y} = \delta_{x,y}\pi_x$ is a diagonal matrix. Moreover, if we denote by $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^d$, for any $f : E \to \mathbb{R}$, i. e. f is a collection of real numbers indexed by the d elements of the set E, in other words an element of $\mathbb{R}^d$, denoting $g = \Pi^{-1/2} f$, we have

$$\begin{aligned}
\|\tilde{P}f\|^2 &= \sum_{x\in E} (P\Pi^{-1/2} f)_x^2 \\
&= \|Pg\|_\pi^2 \\
&\le \|g\|_\pi^2 \\
&= \|f\|^2.
\end{aligned}$$

First note that f is an eigenvector of $\tilde{P}$ if and only if $g = \Pi^{-1/2} f$ is a right eigenvector of P, and $g' = \Pi^{1/2} f$ is a left eigenvector of P, associated with the same eigenvalue. We have that $\tilde{P}$ is a symmetric $d \times d$ matrix, whose norm is bounded by one. Hence from elementary results in linear algebra, $\tilde{P}$ admits the eigenvalues $-1 \le \lambda_d \le \lambda_{d-1} \le \lambda_2 \le \lambda_1 \le 1$. Let us establish the

Lemma 2.8.1. *We have* $\lambda_2 < \lambda_1 = 1$ *and* $-1 < \lambda_d$.

Proof If e denotes the vector whose x–th component equals $\sqrt{\pi_x}$, we have $(\tilde{P}e)_x = \sqrt{\pi_x} \sum_{y\in E} P_{x,y} = e_x$, and we have an eigenvector for the eigenvalue $\lambda_1 = 1$. This is also an eigenvalue of P, the associated right eigenvector being the vector $\Pi^{-1/2} e = (1, 1, \dots, 1)$ and the associated left eigenvector being the vector $\Pi^{1/2} e = \pi$.

The equality $\lambda_2 = \lambda_1$ would mean that the eigenspace associated to the eigenvalue 1 would be two–dimensional, in other words there would exist f linearly independent of e such that $\tilde{P} f = f$, which would imply that $f' = \Pi^{1/2} f$, considered as a row vector, would be such that $f'P = f'$. Now there would exist $\alpha \in \mathbb{R}$ such that $(f' + \alpha\pi)_x \ge 0$, for all $x \in E$. We would have a second invariant measure linearly independent of π, which contradicts irreducibility.

Finally if -1 were an eigenvalue of $\tilde{P}$, it would also be an eigenvalue of P, hence there would exist f such that $Pf = -f$, then we would have $P^{2n} f = f$, hence $f = \lim_{n\to\infty} P^{2n} f = \langle f, \pi\rangle$. But $g = \Pi^{1/2} f$ is an eigenvector of $\tilde{P}$ associated with the eigenvalue -1, hence it is orthogonal to e, in other words $\langle f, \pi\rangle = 0$, hence $f \equiv 0$, and -1 is not an eigenvalue. $\square$

Denote by $g_1, \dots, g_d$ the orthonormal basis of $L^2(\pi)$ made of right eigenvectors of P, corresponding respectively to the eigenvalues $1, \lambda_2, \dots, \lambda_d$. For any $f \in L^2(\pi)$ we have, since $g_1 = (1, \ldots, 1)$,

$$\begin{aligned}
f - \langle f, \pi \rangle &= \sum_{\ell=2}^{d} \langle f, g_\ell \rangle_\pi g_\ell \\
Pf - \langle f, \pi \rangle &= \sum_{\ell=2}^{d} \lambda_\ell \langle f, g_\ell \rangle_\pi g_\ell \\
P^n f - \langle f, \pi \rangle &= \sum_{\ell=2}^{d} \lambda_\ell^n \langle f, g_\ell \rangle_\pi g_\ell \\
\| P^n f - \langle f, \pi \rangle \|_\pi^2 &= \sum_{\ell=2}^{d} \lambda_\ell^{2n} \langle f, g_\ell \rangle_\pi^2 \\
&\leq \sup_{2 \leq \ell \leq d} \lambda_\ell^{2n} \| f - \langle f, \pi \rangle \|_\pi^2,
\end{aligned}$$

hence

Proposition 2.8.2.

$$\| P^n f - \langle f, \pi \rangle \|_\pi \leq\leq (1 - \beta)^n \| f - \langle f, \pi \rangle \|_\pi,$$

where $\beta := (1 - \lambda_2) \wedge (1 + \lambda_d)$ *is the spectral gap.*

2.8.2 The general case

More generally, the same is true with

$$\beta := 1 - \sup_{f \in L^2(\pi),\ \|f\|_\pi = 1} \| Pf - \langle f, \pi \rangle \|_\pi.$$

Indeed, with this β, considering only the case $f \neq 0$, since all inequalities below are clearly true for $f = 0$, we have

$$\begin{aligned}
\| Pf - \langle f, \pi \rangle \|_\pi &= \left\| P \left(\frac{f}{\|f\|_\pi} \right) - \langle \frac{f}{\|f\|_\pi}, \pi \rangle \right\|_\pi \times \|f\|_\pi \\
&\leq (1 - \beta) \|f\|_\pi.
\end{aligned}$$

Finally we check that Proposition 2.8.2 still holds in the general case, with β defined above. Note that

$$\begin{aligned}
\| P^{n+1} f - \langle f, \pi \rangle \|_\pi &= \| P[P^n f - \langle f, \pi \rangle] \|_\pi \\
&\leq (1 - \beta) \| P^n f - \langle f, \pi \rangle \|_\pi.
\end{aligned}$$

The result follows by induction.

In practice the problem is to estimate the spectral gap β precisely. We shall describe one such result in section 3.3 below. The notion of spectral gap will appear again below in section 7.10.

This content of this section was inspired by the treatment in [53]. For a more complete introduction to this topic, see [51].

2.9 Statistics of Markov chains

The aim of this section is to introduce the basic notions for the estimation of the parameters of a Markov chain.

We have seen that for all $n > 0$, the law of the random vector $(X_0, X_1, \dots, X_n)$ depends only on the initial law μ and on the transition matrix P. We want to see under which conditions one can estimate the pair (μ, P), given the observation of $(X_0, X_1, \dots, X_n)$, in such a way that the error tends to zero, as $n \to \infty$.

Let us first discuss the estimation of the invariant probability μ.

For any $x \in E$,

$$\hat{\mu}_x^n = \frac{1}{n+1} \sum_{\ell=0}^{n} \mathbf{1}_{\{X_\ell = x\}}$$

is a consistent estimator of μ_x, since the following is an immediate consequence of the ergodic theorem :

Proposition 2.9.1. *For any $x \in E$, $\hat{\mu}_x^n \to \mu_x$ a. s., as $n \to \infty$.*

Let us now discuss the estimation of the P_{xy}'s, $x, y \in E$. We choose the estimator

$$\hat{P}_{xy}^n = \frac{\displaystyle\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x, X_{\ell+1} = y\}}}{\displaystyle\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell = x\}}}$$

We have the

Proposition 2.9.2. *For any $x, y \in E$, $\hat{P}_{xy}^n \to P_{xy}$ a. s. as $n \to \infty$.*

PROOF We clearly have

$$\hat{P}^n_{xy} = \left(\frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell=x\}}\right)^{-1} \frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell=x, X_{\ell+1}=y\}}.$$

We know that

$$\frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell=x\}} \to \mu_x.$$

For $n \geq 0$, define $\widetilde{X}_n = (X_n, X_{n+1})$. It is not very hard to check that $\{\widetilde{X}_n, n \geq 0\}$ is an irreducible and positive recurrent $\widetilde{E} = \{(x,y) \in E \times E, P_{xy} > 0\}$–valued Markov chain, with transition matrix $\widetilde{P}_{(x,y)(u,v)} = \delta_{yu} P_{uv}$. It admits the invariant probability $\widetilde{\mu}_{(x,y)} = \mu_x P_{xy}$. The ergodic theorem applied to the chain $\{\widetilde{X}_n\}$ implies that a.s., as $n \to \infty$,

$$\frac{1}{n}\sum_{\ell=0}^{n-1} \mathbf{1}_{\{X_\ell=x, X_{\ell+1}=y\}} \to \mu_x P_{xy}$$

□

2.10 Exercises

Exercise 2.10.1. *Show that the $E = \{1,2,3\}$–valued Markov chain $\{X_n,\ n \in \mathbb{N}\}$ whose transition matrix is*

$$P = \begin{pmatrix} 1 & 0 & 0 \\ p & 1-p-q & q \\ 0 & 0 & 1 \end{pmatrix} \ (p, q > 0, p+q < 1)$$

starting from $X_0 = 2$, first changes its value at a random time $T \geq 1$ whose law is geometric. Show moreover that X_T is independent of T, and give the law of X_T. Finally show that $X_t = X_T$ if $t \geq T$.

Exercise 2.10.2. *Let $\{X_n;\, n \in \mathbb{N}\}$ be an $E = \{1, 2, 3, 4, 5\}$–valued Markov chain, with the transition matrix*

$$P = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

Find the equivalence classes, the transient and recurrent states, and all invariant measures of $\{X_n\}$.

Exercise 2.10.3. *Consider a Markov chain $\{X_n;\, n \in \mathbb{N}\}$ taking values in the finite state $E = \{1, 2, 3, 4, 5, 6\}$, with a transition matrix P whose off–diagonal entries are given by*

$$P = \begin{pmatrix} \cdot & 2/3 & 1/3 & 0 & 0 & 0 \\ 1/4 & \cdot & 0 & 0 & 1/5 & 2/5 \\ 0 & 0 & \cdot & 1/2 & 0 & 0 \\ 0 & 0 & 2/3 & \cdot & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdot & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & \cdot \end{pmatrix}$$

1. *Find the diagonal entries of the transition matrix P.*

2. *Show that E can be partitioned into three equivalence classes to be specified, of which one ($\mathcal{T}$) is transient and two ($\mathcal{R}_1$ and $\mathcal{R}_2$) are recurrent.*

3. *Find an invariant probability whose support is $\mathcal{R}_1$ and another one whose support is $\mathcal{R}_2$. Find all invariant probabilities.*

Exercise 2.10.4. *Let P be a Markovian matrix over a finite or countable set E, which satisfies Doeblin's condition (D) of section 2.6.*

1. *Suppose first that condition (D) is satisfied with $n_0 = 1$. Show that there exists at least one recurrent state, which is visited infinitely often by the chain, for any starting point. Deduce that the chain has a unique recurrent class. (Indication : first show that there exists $x \in E$, $\beta > 0$ such that the chain can be simulated as follows : at each time n, set $X_n = x$ with probability β, and with probability $1 - \beta$, do a certain Markovian transition).*

2. *Show that the result is still true in the general case of the assumption (D). (Indication : consider the sub–chain* $\{X_{kn_0},\ k = 0, 1, \dots\}$*).*

Exercise 2.10.5. *Show that whenever x is recurrent,* $\sum_{n\geq 0}(P^n)_{xy} = +\infty$ *iff* $x \leftrightarrow y$, $= 0$ *iff* $x \not\rightarrow y$.

Exercise 2.10.6 (Equivalence of the two definitions of aperiodicity)**.** *Let* $x \in E$*. Define* $N_x = \{n,\ (P^n)_{xx} > 0\}$.

1. *Show that whenever N_x contains two consecutive integers, the greatest common divisor of the elements of N_x is 1.*

2. *Show that if* $n, n+1 \in N_x$*, then* $\{n^2, n^2+1, n^2+2, \dots\} \subset N_x$.

3. *Show that if the GCD of the elements of N_x is 1, then there exists* $n \in \mathbb{N}$ *such that* $\{n, n+1\} \subset N_x$.

4. *Conclude that the two definitions of aperiodicity of a state x are equivalent.*

Exercise 2.10.7. *Consider an* $E = \{1, 2, 3, 4, 5, 6\}$*–valued Markov chain* $\{X_n;\ n \in \mathbb{N}\}$ *with the transition matrix P, whose off–diagonal entries are specified by*

$$P = \begin{pmatrix} \cdot & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & \cdot & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 7/8 & 0 \\ 1/4 & 1/4 & 0 & \cdot & 1/4 & 1/4 \\ 0 & 0 & 3/4 & 0 & \cdot & 0 \\ 0 & 1/5 & 0 & 1/5 & 1/5 & \cdot \end{pmatrix}$$

1. *Find the diagonal terms of the transition matrix P.*

2. *Find the equivalence classes of the chain.*

3. *Show that 4 and 6 are transient states, and that the other states can be grouped in two recurrent classes to be specified. In the sequel, we let* $\mathcal{T} = \{4, 6\}$*,* $\mathcal{C}$ *be the recurrent class containing 1, and* $\mathcal{C}'$ *the other recurrent class. For all* $x, y \in E$*, define* $\rho_x := \mathbb{P}_x(T < \infty)$*, where* $T := \inf\{n \geq 0;\ X_n \in \mathcal{C}\}$.

4. *Show that*

$$\rho_x = \begin{cases} 1, & \text{if } x \in \mathcal{C}; \\ 0, & \text{if } x \in \mathcal{C}'; \end{cases}$$

and that $0 < \rho_x < 1$, *if* $x \in \mathcal{T}$.

5. *Using the decomposition* $\{T < \infty\} = \{T = 0\} \cup \{T = 1\} \cup \{2 \le T < \infty\}$ *and conditioning in the computation of* $\mathbb{P}_x(2 \le T < \infty)$ *by the value of* X_1, *establish the formula*

$$\rho_x = \sum_{y \in E} P_{xy} \rho_y, \qquad \text{if } x \in \mathcal{T}.$$

6. *Compute* ρ_4 *and* ρ_6.

7. *Deduce (without any serious computation !) the values of* $\mathbb{P}_4(T_{\mathcal{C}'} < \infty)$ *and* $\mathbb{P}_6(T_{\mathcal{C}'} < \infty)$, *where* $T_{\mathcal{C}'} := \inf\{n \ge 0;\ X_n \in \mathcal{C}'\}$.

Exercise 2.10.8. *Consider an* $E = \{1, 2, 3, 4, 5, 6\}$*–valued Markov chain* $\{X_n;\ n \in \mathbb{N}\}$ *with the transition matrix* P, *whose off–diagonal entries are given by*

$$P = \begin{pmatrix} \cdot & 1/4 & 1/3 & 0 & 0 & 0 \\ 1/4 & \cdot & 0 & 1/4 & 1/3 & 0 \\ 1/2 & 0 & \cdot & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdot & 1/2 & 1/3 \\ 0 & 0 & 0 & 1/2 & \cdot & 1/2 \\ 0 & 0 & 0 & 1/3 & 1/4 & \cdot \end{pmatrix}.$$

1. *Find the diagonal entries of the transition matrix* P.

2. *Show that* E *is the union of two equivalence classes to be specified, one* $\mathcal{R}$ *being recurrent and the other* $\mathcal{T}$ *transient.*

3. *Define* $T := \inf\{n \ge 0;\ X_n \in \mathcal{R}\}$ *and* $h_x = \mathbb{E}_x(T)$, *for* $x \in E$. *Show that* $h_x = 0$ *for* $x \in \mathcal{R}$, *and that* $1 < h_x < \infty$ *pour* $x \in \mathcal{T}$.

4. *Show that for all* $x \in \mathcal{T}$,

$$h_x = 1 + \sum_{y \in E} P_{xy} h_y.$$

Deduce the values of h_x, $x \in \mathcal{T}$.

Exercise 2.10.9. *Given $0 < p < 1$, we consider an $E = \{1, 2, 3, 4\}$–valued Markov chain $\{X_n;\, n \in \mathbb{N}\}$ with the transition matrix P given by*

$$P = \begin{pmatrix} p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \\ p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \end{pmatrix}.$$

1. *Show that the chain $\{X_n\}$ is irreducible and recurrent .*
2. *Compute its unique invariant probability π.*
3. *Show that the chain is aperiodic. Deduce that P^n tends, as $n \to \infty$, towards the matrix*
$$\begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \end{pmatrix}.$$
4. *Compute P^2. Show that this transition matrix coincides with the above limit. Determine the law of X_2, as well as that of X_n, $n \geq 2$.*
5. *Define $T_4 = \inf\{n \geq 1,\, X_n = 4\}$. Compute $\mathbb{E}_4(T_4)$.*

Exercise 2.10.10. *Consider an $E = \{0, 1, 2, 3, 4\}$–valued Markov chain $\{X_n,\, n \in \mathbb{N}\}$ with the transition matrix :*

$$P = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ p & 0 & \frac{1-p}{2} & 0 & \frac{1-p}{2} \\ p & \frac{1-p}{2} & 0 & \frac{1-p}{2} & 0 \\ p & 0 & \frac{1-p}{2} & 0 & \frac{1-p}{2} \\ p & \frac{1-p}{2} & 0 & \frac{1-p}{2} & 0 \end{pmatrix},$$

where $0 < p < 1$. Let $T := \inf\{n \geq 1,\, X_n = 0\}$.

1. *Show that the chain $\{X_n\}$ is irreducible and recurrent. We shall denote its invariant probability by π.*
2. *Show that under $\mathbb{P}_0$, the law of T is a geometric law to be specified. Show that $\mathbb{E}_0(T) = \frac{p+1}{p}$.*

3. *Let*

$$N_n = \sum_{k=1}^{n} \mathbf{1}_{\{X_k=0\}}, \quad M_n = \sum_{k=1}^{n} \mathbf{1}_{\{X_k\neq 0\}}.$$

Compute the limits as $n \to \infty$ *of* $n^{-1}N_n$ *and* $n^{-1}M_n$.

4. *Give an intuitive argument to support the identity*

$$\pi_1 = \pi_2 = \pi_3 = \pi_4.$$

Deduce the probability π*, exploiting this identity .*

5. *Show the following general result. If there exists a one–to–one mapping* τ *from* E *into itself, such that*

$$P_{\tau x,\tau y} = P_{xy}, \quad \forall x, y \in E,$$

then the invariant probability π *has the following property :* $\pi_{\tau x} = \pi_x$, $x \in E$. *Deduce a rigourous argument for the result in the previous question.*

Exercise 2.10.11 (Random walk in $\mathbf{Z}$)**.** *Let*

$$X_n = X_0 + Y_1 + \cdots + Y_n,$$

where the X_n*'s take their values in* $\mathbf{Z}$*, the* Y_n*'s in* $\{-1, 1\}$*,* $X_0, Y_1, .., Y_n, ..$ *being a sequence of independent r. v.'s, and for all* n,

$$\mathbb{P}(Y_n = 1) = p = 1 - \mathbb{P}(Y_n = -1), \quad 0 < p < 1.$$

1. *Show that the chain* $\{X_n\}$ *is irreducible.*

2. *Show that whenever* $p \neq 1/2$*, the chain is transient (use the law of large numbers).*

3. *Consider the case* $p = 1/2$*. Show that the chain is recurrent (evaluate* $\sum_{n\geq 1}(P^n)_{00}$ *with the help of Stirling's formula* $n! \simeq \sqrt{2\pi n}(\frac{n}{e})^n$*). Show that the chain is null recurrent (look for an invariant measure). Determine the quantities*

$$\limsup_{n\to\infty} X_n \quad \textit{and} \quad \liminf_{n\to\infty} X_n.$$

Exercise 2.10.12 (Random walk in $\mathbf{Z}^d$). *We let*

$$X_n = X_0 + Y_1 + \cdots + Y_n,$$

where the X_n's take their values in $\mathbf{Z}^d$,the Y_n's being i. i. d., globally independent of X_0, and their law is specified by

$$\mathbb{P}(Y_n = \pm e_i) = (2d)^{-1}, \quad 1 \le i \le d,$$

where $\{e_1, \ldots, e_d\}$ is the canonical basis of $\mathbf{Z}^d$.

1. *Show that the common characteristic function of the Y_n's is given by*

$$\phi(t) = d^{-1} \sum_{j=1}^{d} \cos(t_j),$$

and that

$$(P^n)_{00} = (2\pi)^{-d} \int_{[-\pi,\pi]^d} \phi^n(t) dt.$$

2. *Deduce that for all $0 < r < 1$,*

$$\sum_{n \ge 0} r^n (P^n)_{00} = (2\pi)^{-d} \int_{[-\pi,\pi]^d} (1 - r\phi(t))^{-1} dt.$$

3. *Show that $\forall \alpha > 0$, the mapping*

$$(r,t) \to (1 - r\phi(t))^{-1}$$

is bounded on $]0,1] \times ([-\pi,\pi]^d \backslash C_\alpha)$, where $C_\alpha = \{t \in \mathbb{R}^d,\ \|t\| \le \alpha\}$, and that whenever $\|t\|$ is sufficiently small, $r \to (1-r\phi(t))^{-1}$ is positive and increasing.

4. *Deduce from the fact that $1 - \phi(t) \simeq \|t\|^2/2$, as $t \to 0$, that $\{X_n\}$ is an irreducible $\mathbf{Z}^d$–valued Markov chain, which is null recurrent if $d = 1, 2$, transient if $d \ge 3$.*

Exercise 2.10.13. *Consider again the $\mathbf{Z}$–valued random walk from Exercise 2.10.11 in the symmetric case ($p = 1/2$). The goal of this exercise is to establish the null recurrence of the walk by a method which is completely* *different from that of Exercise 2.10.11. Suppose for simplicity that* $X_0 = x \in \mathbf{Z}$.

For all $a, b \in \mathbf{Z}$ *with* $a < x < b$, *let*

$$\begin{aligned} T_{a,b} &= \inf\{n \geq 0;\ X_n \notin]a, b[\}, \\ T_a &= \inf\{n \geq 0;\ X_n = a\}, \\ T_b &= \inf\{n \geq 0;\ X_n = b\}. \end{aligned}$$

We note that

$$X_{n\wedge T_{a,b}} = x + \sum_{k=1}^{n} Y_k \mathbf{1}_{\{T_{a,b}>k-1\}}.$$

1. *Show that the r. v.'s* Y_k *and* $\mathbf{1}_{\{T_{a,b}>k-1\}}$ *are independent. Deduce that*

$$\mathbb{E} X_{n\wedge T_{a,b}} = x.$$

2. *Show that* $|X_{n\wedge T_{a,b}}| \leq \sup(|a|, |b|)$, $T_{a,b} < \infty$ *a. s., and*

$$\mathbb{E} X_{T_{a,b}} = x.$$

3. *Establish the identities*

$$\mathbb{P}(X_{T_{a,b}} = a) = \frac{b-x}{b-a}, \quad \mathbb{P}(X_{T_{a,b}} = b) = \frac{x-a}{b-a}.$$

4. *Show that* $\mathbb{P}(T_a < T_n) \to 1$, *as* $n \to \infty$.

5. *Show that* $T_a < \infty$ *a. s., and similarly that* $T_b < \infty$ *a. s. Deduce that the chain is recurrent.*

6. *In the sequel we consider without loss of generality the case* $x = 0$, *for the sake of notational simplicity. Show that for all* $n \geq 1$,

$$X^2_{n\wedge T_{a,b}} = \sum_{k=1}^{n} (1 - 2X_{k-1}Y_k) \mathbf{1}_{\{T_{a,b}>k-1\}}.$$

7. *Deduce that* $\mathbb{E}(X^2_{T_{a,b}}) = \mathbb{E}(T_{a,b}) = -ab$, *and that for all* $a \in \mathbf{Z}$, $\mathbb{E}(T_a) = +\infty$, *which shows that the chain is null recurrent.*

Exercise 2.10.14 (Reflected random walk). *The $\{Y_n\}$'s being defined as in Exercise 2.10.11, we define the $\mathbb{N}$–valued Markov chain $\{X_n\}$ by the recurrence formula*

$$X_{n+1} = X_n + \mathbf{1}_{\{X_n > 0\}} Y_{n+1} + \mathbf{1}_{\{X_n = 0\}}.$$

We assume that $X_0 \in \mathbb{N}$. We shall denote below by $\{X'_n\}$ the (unreflected) random walk from Exercise 2.10.11, with the same X_0 and the same $\{Y_n\}$'s. Below we shall use freely the results from Exercise 2.10.11.

1. *Show that the chain $\{X_n\}$ is irreducible, as an $\mathbb{N}$–valued chain. Give its transition matrix.*

2. *Show that a. s., $X_n \geq X'_n$, $\forall n$. Conclude that $\{X_n\}$ is transient in the case $p > 1/2$.*

3. *Let $T = \inf\{n \geq 0,\, X_n = 0\}$. Show that $X_n = X'_n$ whenever $T \geq n$. Conclude that the chain is recurrent in the case $p \leq 1/2$ (one can e. g. show that the state 1 is recurrent).*

4. *Show that the chain is null recurrent in the case $p = 1/2$, positive recurrent in the case $p < 1/2$ (one might check that in the first (resp. the second) case, the measure $(1/2, 1, 1, 1, \ldots)$ (resp. the probability μ defined by*

$$\mu_0 = \frac{1-2p}{2(1-p)}, \quad \mu_x = \frac{1-2p}{2} \frac{p^{x-1}}{(1-p)^{x+1}}, \; x \geq 1)$$

   is an invariant measure).

Exercise 2.10.15 (Birth and death Markov chain). *Let $\{X_n\}$ be an $E = \mathbb{N}$–valued Markov chain with the transition P given by*

$$P_{x,x-1} = q_x, \quad P_{x,x} = r_x, \quad P_{x,x+1} = p_x,$$

where for all $x \in \mathbb{N}$, $p_x + r_x + q_x = 1$, $q_0 = 0$, $q_x > 0$ if $x > 0$, and $p_x > 0$ for all $x \in \mathbb{N}$.

For $x \in \mathbb{N}$, we let $\tau_x = \inf\{n \geq 0,\, X_n = x\}$. Given three states a, x and b such that $a \leq x \leq b$, we define $u(x) = \mathbb{P}_x(\tau_a < \tau_b)$. Let $\{\gamma_x,\, x \in \mathbb{N}\}$ be defined by $\gamma_0 = 1$ and for $x > 0$, $\gamma_x = q_1 \times \cdots \times q_x / p_1 \times \cdots \times p_x$.

1. *Show that the chain is irreducible.*

2. *For $a < x < b$, establish a relation between $u(x) - u(x+1)$ and $u(x-1) - u(x)$. Compute $u(a) - u(b)$ in terms of the γ_x's, and deduce that for $a < x < b$,*
$$u(x) = \sum_{y=x}^{y=b-1} \gamma_y / \sum_{y=a}^{y=b-1} \gamma_y.$$
Consider the particular case where $p_x = q_x$ for all $x > 0$.

3. *Compute $\mathbb{P}_1(\tau_0 = \infty)$ and show that the chain is recurrent iff $\sum_0^\infty \gamma_y = +\infty$.*

4. *Find the invariant measures, and deduce that the chain is positive recurrent iff*
$$\sum_{x=1}^{\infty} \frac{p_0 p_1 \times \cdots \times p_{x-1}}{q_1 q_2 \times \cdots \times q_x} < \infty.$$

5. *Show that in the positive recurrent case, the chain is reversible. (Hint : one might first note that for $x > 0$ the relation $\pi_x = (\pi P)_x$ can be written*
$$\pi_x P_{x,x-1} + \pi_x P_{x,x+1} = \pi_{x-1} P_{x-1,x} + \pi_{x+1} P_{x+1,x}.$$
Then consider the case $x = 0$, and show by recurrence that
$$\pi_x P_{x,x+1} = \pi_{x+1} P_{x+1,x}, \quad \forall x \geq 0).$$

Exercise 2.10.16 (Queue). *We consider a discrete time queue, which evolves as follows : at each time $n \in \mathbb{N}$, one customer arrives with probability p, $(0 < p < 1)$ and no customer arrives with probability $1 - p$. During each unit time interval when at least one customer is present, one customer is served and leaves the queue with probability q, $0 < q < 1$, and nobody leaves the queue with probability $1 - q$ (a customer who arrives at time n leaves at the earliest at time $n + 1$). All the above events are mutually independent. Denote by X_n the number of customers in the queue at time n.*

1. *Show that $\{X_n, n \in \mathbb{N}\}$ is an irreducible $E = \mathbb{N}$–valued Markov chain. Determine its transition matrix P_{xy}, $x, y \in \mathbb{N}$.*

2. *Give a necessary and sufficient condition on p and q for the chain $\{X_n\}$ to possess an invariant probability. We assume below that this condition is satisfied. Specify the unique invariant probability $\{\pi_x,\, x \in \mathbb{N}\}$ of the chain $\{X_n\}$.*

3. *Compute $\mathbb{E}_\pi(X_n)$.*

4. *We add the information that the customers are being served according to the order in which they arrive. Denote by T the sojourn time in the queue of a customer who arrives at an arbitrary fixed time. Assuming that the queue is initialized with its invariant probability, what is the expectation of T ?*

Exercise 2.10.17 (Queue)**.** *We consider a queue at a counter. X_n denotes the number of customers in the queue at time n. Between times n and $n+1$, Y_{n+1} new customers join the queue, and provided $X_n > 0$, Z_{n+1} customers leave the queue. We assume that X_0, Y_1, Z_1, Y_2, $Z_2 \dots$ are mutually independent, the Y_n's having all the same law, such that $0 < \mathbb{P}(Y_n = 0) < 1$, and the Z_n's satisfying $\mathbb{P}(Z_n = 1) = p = 1 - \mathbb{P}(Z_n = 0)$.*

1. *Show that $(X_n; n \in \mathbb{N})$ is a Markov chain, and give its transition matrix.*

2. *Let φ denote the common characteristic function of the Y_n's, ρ that of the Z_n's, Ψ_n that of X_n. Compute Ψ_{n+1} in terms of Ψ_n, φ and ρ.*

3. *Show that there is a unique invariant probability iff $\mathbb{E}(Y_1) < p$, and determine its characteristic function.*

Exercise 2.10.18 (Queue)**.**

- **A** *Let X denote the random number of individuals in a given population, and $\phi(u) = \mathbb{E}[u^X]$, $0 \le u \le 1$ its generating function. Each individual is selected with probability q ($0 < q < 1$), independently from the others. Let Y denote the number of selected individuals in the initial population of X individuals. Show that the generating function ψ of Y (defined as $\psi(u) = \mathbb{E}[u^Y]$) is given by*
$$\psi(u) = \phi(1 - q + qu).$$

- **B** *We consider a service system (equipped with an infinite number of servers), and we denote by X_n ($n = 0, 1, 2, \dots$) the number of customers* *who are present in the system at time n. We assume that at time $n + 1/3$ each of the X_n customers leaves the system with probability $1 - p$, and stays with probability p (independently from the others, and of all the other events) (we denote by X'_n the number of remaining customers), and at time $n+2/3$ Y_{n+1} new customers join the queue. We assume that the random variables $X_0, Y_1, Y_2, \dots$ are mutually independent, and globally independent of the ends of services, and that the joint law of the Y_n's is the Poisson distribution with parameter $\lambda > 0$ (i. e. $\mathbb{P}(Y = k) = e^{-\lambda}\lambda^k/k!$, and $\mathbb{E}[u^{Y_n}] = \exp[\lambda(u-1)]$).*

1. *Show that $\{X_n;\ n \geq 0\}$ is an irreducible $E = \mathbb{N}$–valued Markov chain.*
2. *Compute $\mathbb{E}[u^{X_{n+1}}|X_n = x]$ in terms of u, p, λ and x.*
3. *Denote by $\phi_n(u) = \mathbb{E}[u^{X_n}]$ the generating function of X_n. Compute ϕ_{n+1} in terms of ϕ_n, and show that*

$$\phi_n(u) = \exp\left[\lambda(u-1)\sum_{0}^{n-1} p^k\right] \phi_0(1 - p^n + p^n u).$$

4. *Show that $\rho(u) = \lim_{n\to\infty} \phi_n(u)$ exists and does not depend on ϕ_0, and that ρ is the generating function of a Poisson distribution, whose parameter is to be specified in terms of λ and p.*
5. *Show that $\{X_n;\ n \geq 0\}$ is positive recurrent and specify its invariant probability.*

Exercise 2.10.19. *Let $X_0, A_0, D_0, A_1, D_1, \dots$ be $\mathbb{N}$–valued mutually independent random variables. The D_n's are Bernoulli random variables with parameter q, i. e. $\mathbb{P}(D_n = 1) = 1 - \mathbb{P}(D_n = 0) = q$, $0 < q < 1$. The A_n's all have the same law defined by $\mathbb{P}(A_n = k) = r_k$, $k \in \mathbb{N}$, where $0 \leq r_k < 1$, $0 < r_0 < 1$ and $\sum_{k=0}^{\infty} r_k = 1$. We assume that $p = \sum_k k r_k < \infty$.*

We consider the sequence of r. v. $\{X_n;\ n \in \mathbb{N}\}$ defined by

$$X_{n+1} = (X_n + A_n - D_n)^+,\ n \geq 0,$$

with the usual notation $x^+ = \sup(x, 0)$.

1. *Show that $\{X_n;\ n \in \mathbb{N}\}$ is an $E = \mathbb{N}$–valued Markov chain. Give its transition matrix P, and show that the chain is irreducible.*

   We assume from now on that $X_0 = 0$. Let $T = \inf\{n > 0;\ X_n = 0\}$. Define $S_n = \sum_{k=0}^{n-1}(A_k - D_k)$.

2. *Show that* $X_n \geq S_n$*, and that* $X_{n+1} = S_{n+1}$ *on the event* $\{T > n\}$*.*

3. *Show that* $S_n/n \ \to p - q$ *a. s., as* $n \to \infty$*.*

4. *Show that whenever* $p < q$*,* $T < \infty$ *a. s.*

5. *Assume that* $p > q$*. Show that* $\{X_n;\, n \in \mathbb{N}\}$ *visits 0 at most a finite number of times.*

6. *In the case* $p \neq q$*, specify in which case the chain is recurrent, and in which case it is transient.*

   Assume from now on that $\mathbb{P}(A_n = 1) = 1 - \mathbb{P}(A_n = 0) = p$*, where* $0 < p < 1$ *(p is again the expectation of* A_n*).*

7. *Specify the transition matrix* P *in this case.*

8. *Show that if* $p = q$*, the chain is null recurrent. (one might use the result of question 3. from Exercise 2.10.11 in order to show the recurrence, and then look for an invariant measure).*

9. *Assume that* $p < q$*. Show that the chain has a unique invariant probability* π *on* $\mathbb{N}$*, and that* $\pi_k = (1-a)a^k$*, with* $a = p(1-q)/q(1-p)$ *(one might first establish a recurrence relation for the sequence* $\Delta_k = \pi_k - \pi_{k+1}$*). Show that the chain is positive recurrent.*

Exercise 2.10.20 (Discrete Aloha)**.** *The aim of this exercise is to study the following communication protocol : users arrive at times* $\{1, 2, \ldots, n, \ldots\}$ *and they want to transmit a message through a channel, which has the capacity to transmit only one message at a time. Whenever several users try to transmit a message at the same time, no message is transmitted, each user knows this, and he makes a new attempt later. We look for a "distributed" retransmission policy, i. e. such that each user may decide when to try to retransmit, without knowing the intentions of the other users. The "discrete Aloha" protocol prescribes that each user whose message has been blocked at time n makes a new attempt at time* $n+1$ *with probability p. If he decides not to try at time* $n+1$*, he again makes an attempt at time* $n+2$ *with probability p, and so on until by chance he does try. Let* Y_n *denote the number of "new" messages (i. e. which have not been presented before) arriving at time n. We assume that the* $\{Y_n\}$*'s are i. i. d., with* $\mathbb{P}(Y_n = i) = a_i$*,* $i \in \mathbb{N}$*, and* $\mathbb{E}(Y_n) > 0$*. Let* X_n *denote the number of delayed messages which are waiting to be transmitted at time n.*

1. *Show that $\{X_n\}$ is a Markov chain, and give its transition matrix.*
2. *Show that $\{X_n\}$ is irreducible, but not positive recurrent.*

Exercise 2.10.21 (Programming)**.** *Consider again the queue from Exercise 2.10.16.*

1. *Simulate and plot a trajectory $\{X_n,\ n \geq 0\}$ from $n = 1$ to $n = 1000$, with $p = 1/2$ and successively $q = 3/5, 7/13, 15/29, 1/2$.*
2. *Since $\{X_n\}$ is irreducible, positive recurrent and aperiodic, $(P^n)_{yx} \to \pi_x$. Plot either the empirical histogram or the empirical distribution function of $(P^n)_{y\cdot}$, for $n = 100, 500, 1000$, and a sample size of 10^4. Show the histogram (resp. the distribution function) of π on the same picture. Treat the cases $p = 1/2$, $q = 3/5, 7/13$.*
3. *Graphically compare the quantities*

$$n^{-1} \sum_{k=1}^{n} \mathbf{1}_{\{X_k = x\}}, \quad x \in \mathbb{N}$$

and the histogram of π, for $n = 10^3, 10^4, 10^5$. Treat the cases $p = 1/2$, $q = 3/5, 7/13$. For each value of q, choose the interval of values of x from the previous results.

Exercise 2.10.22 (Ordering a database)**.** *Suppose that the memory of a computer contains n items $1, 2, \ldots, n$. This memory receives successive requests, each consisting of one of the items. The closer the item is to the head of the list, the faster the access is. Assume that the successive requests are i. i. d. r. v.'s. If the common law of those r. v.'s were known, the best choice would be to order the data in decreasing order of their associated probability of being requested. But this probability $(p_1, p_2, \ldots, p_n)$ is either unknown or slowly varying. Assume that $p_k > 0$, $\forall 1 \leq k \leq n$.*

We need to choose a method of replacement of the data after their request, in such a way that in the long run the time taken to get the requested data will be as small as possible.

We will compare two such methods. The first one consists of replacing systematically any item which has been requested at the head of the list. The second one consists of moving each item which has been requested one step ahead. In both cases, we have an irreducible Markov chain with values in the *set E of all permutations of the set $\{1, 2, \ldots, n\}$. Denote by Q the transition matrix of the first chain, and by π the associated invariant measure, by P the transition matrix of the second chain and μ the corresponding invariant measure. To the Markov chain with transition matrix Q, we associate the quantity*

$$J_Q \stackrel{def}{=} \sum_{k=1}^{n} \pi(\textit{position of } k) p_k,$$

where $\pi(\textit{position of } k)$ is the expectation under π of the position of the element k.

To the Markov chain with transition matrix P, we associate the quantity

$$J_P \stackrel{def}{=} \sum_{k=1}^{n} \mu(\textit{position of } k) p_k.$$

1. *Show that the chain with the transition matrix Q is not reversible.*

2. *Show that any irreducible and positive recurrent Markov chain which satisfies*

   (i) $P_{k\ell} > 0 \Leftrightarrow P_{\ell k} > 0$

   (ii) *For any excursion $k, k_1, k_2, \ldots, k_m, k$,*

   $$P_{kk_1} \prod_{i=2}^{m} P_{k_{i-1}k_i} P_{k_m k} = P_{kk_m} \prod_{i=m-1}^{1} P_{k_{i+1}k_i} P_{k_1 k}$$

   is reversible (this is known as the "Kolmogorov cycle condition").

3. *Show that P satisfies (i) and (ii).*

4. *Show that the second procedure is preferable, in the sense that $J_P < J_Q$.*