AMARUN
www.amarun.org
Comisión de Pedagogía - Abass Sagna
Recurrence and transience of Markov chains.

UCE, verano 2014

1 Introduction

Our aim in this short lecture is to introduce the Markov chains and to study the Recurrence and transience properties of a Markov chain.

2 Definition of a Markov chain

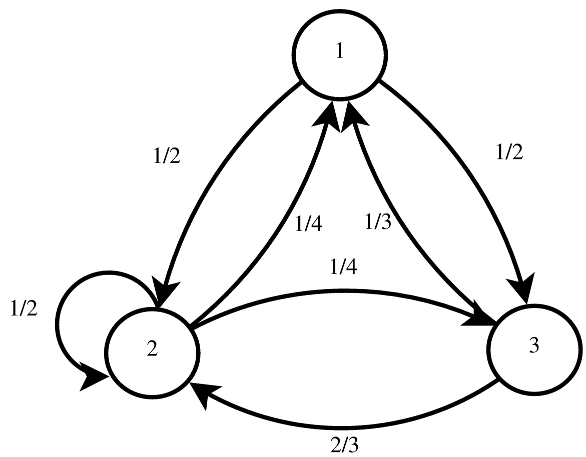
To introduce a Markov chain let us consider the following example. Let $(X_n)_{n\geq 0}$ be a stochastic process taking values in the set $E = \{1, 2, 3\}$. Suppose that at the initial time $n = 0$, X_0 is a random variable valued in $\{1, 2, 3\}$ with the following corresponding probabilities:

$$\mathbb{P}(X_0 = 1) = \frac{1}{2},\ \mathbb{P}(X_0 = 2) = \frac{1}{6},\ \mathbb{P}(X_0 = 3) = \frac{1}{3}.$$

We may then define the probability distribution μ of X_0 as:

$$\mu(1) = \mathbb{P}(X_0 = 1) = \frac{1}{2}, \quad \mu(2) = \mathbb{P}(X_0 = 2) = \frac{1}{6}, \quad \mu(3) = \mathbb{P}(X_0 = 3) = \frac{1}{3}.$$

For $n \geq 1$, the process (X_n) evolves according to the principle described by the following diagram:



The diagram is read as follows: if the process is at the state 1 at time n, then, at time $n+1$, it moves to the state 2 with a probability $\frac{1}{2}$ and to the state 3 with a probability $\frac{1}{2}$. If the process is at the state 2 at time n, then, at time $n+1$, it moves to the state 1 with a probability $\frac{1}{4}$, to the state 3 with a probability $\frac{1}{2}$ or stays at the state 2 with a probability $\frac{1}{4}$. Finally, if the process is at the state 3 at time n, at time $n+1$, it moves to the state 1 with a probability $\frac{1}{3}$ and to the state 2 with a probability $\frac{2}{3}$.

This diagram is determined by the knowledge of probability of transitions $\mathbb{P}(X_{n+1} = j | X_n = i)$, $i, j \in \{1, 2, 3\}$, and conversely. Let $(P(i,j))_{i,j \in E}$ be the matrix of transition probabilities where the rows correspond to the transition probabilities starting from the state i : $P(i,j) = \mathbb{P}(X_{n+1} = j | X_n = i)$, $j \in E = \{1, 2, 3\}$. We therefore have

$$\forall i \in E, \quad \sum_{j \in E} P(i,j) = 1.$$

The process $(X_n)_{n \geq 0}$ describe previously is a Markov chain. The probability distribution μ is called the initial probability distribution of the Markov chain and the set E is called the state space. The elements of E are called the states of the Markov chain.

DEFINITION 2.1. *A matrix $P = (P(i,j))_{i,j \in E}$ is a transition matrix if*

$$\forall i, j \in E, \quad P(i,j) \geq 0$$
$$\forall i \in E, \quad \sum_{j \in E} P(i,j) = 1.$$

EXAMPLE 2.1. The transition matrix related to the process defined previously is given by

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

REMARK 2.1. If P is a transition matrix then for every integer $n \geq 0$, P^n is a transition matrix.

DEFINITION 2.2. *Let E be a countable (finite or infinite) state space. Let μ be a probability on E and let P be a transition matrix on E. A process $(X_n)_{n \geq 0}$ is a (homogeneous) Markov chain on E, with initial probability distribution μ and transition matrix P if*

1. $\mathbb{P}(X_0 = i) = \mu(i)$, *for all $i \in E$,*

2. *For all $i_0, i_1, \ldots, i_{n+1} \in E$,*

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n).$$

EXAMPLE 2.2. Let the process $(X_n)_{n \geq 0}$ be defined by $X_0 = 0$ and $X_n = \sum_{i=1}^{n} Y_i$, where $(Y_i)_{i \geq 1}$ is a iid sequence of random variables defined as

$$Y_i = \begin{cases} +1 & \text{with probability } p \\ -1 & \text{with probability } 1-p. \end{cases}$$

Prove that $(X_n)_{n \geq 0}$ is a Markov chain with initial distribution $\mu = \delta_0$ and transition matrix $P = (P(i,j))_{i,j \in \mathbb{Z}}$, whith components are defined for every $i, j \in \mathbb{Z}$ as

$$P(i,j) = \mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} p & \text{if } j = i+1 \\ 1-p & \text{if } j = i-1 \\ 0 & \text{otherwise.} \end{cases}$$

Answer. We have $X_{n+1} = X_n + Y_{n+1}$. Then, for every $i_0, i_1, \dots, i_{n+1} \in \mathbb{Z}$, we have

$$\begin{aligned}\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) &= \mathbb{P}(Y_{n+1} = i_{n+1} - X_n | X_0 = i_0, \dots, X_n = i_n)\\ &= \mathbb{P}(Y_{n+1} = i_{n+1} - i_n | X_0 = i_0, \dots, X_n = i_n).\end{aligned}$$

Remind that Y_{n+1} is independent from $(X_0, \dots, X_n)$. Then,

$$\mathbb{P}(Y_{n+1} = i_{n+1} - i_n | X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(Y_{n+1} = i_{n+1} - i_n).$$

It remains to remark that $\mathbb{P}(Y_{n+1} = i_{n+1} - i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n)$. In fact,

$$\begin{aligned}\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) &= \mathbb{P}(Y_{n+1} = i_{n+1} - X_n | X_n = i_n)\\ &= \frac{\mathbb{P}(Y_{n+1} = i_{n+1} - i_n; X_n = i_n)}{\mathbb{P}(X_n = i_n)}\\ &= \frac{\mathbb{P}(Y_{n+1} = i_{n+1} - i_n)\mathbb{P}(X_n = i_n)}{\mathbb{P}(X_n = i_n)} \quad (Y_{n+1} \text{ and } X_n \text{ are independent})\\ &= \mathbb{P}(Y_{n+1} = i_{n+1} - i_n).\end{aligned}$$

Consider a Heads-Tails game by tossing a coin which has a probability p of getting a Heads and a probability $1 - p$ of getting a Tails. We gain 1\$ if Heads appears and we lose 1\$ when Tails appears. Let our initial stake be $X_0 = 0$ and let X_n be our wealth at the step n of the game. The process $(X_n)_{n \geq 0}$ may be defined as in the previous example. The fact that it is a Markov chain with initial probability distribution $\mu = \delta_0$ and transition matrix P is expected. In fact, our initial wealth $X_0 = 0$ ($\mu = \delta_0$), and, when X_n, our wealth at step n, is worth i, then X_{n+1} is worth $i + 1$ with probability p (when the result of the $n + 1$-th toss is Heads) and it is worth $i - 1$ with probability $1 - p$ (when the result of the $n + 1$-th toss is Tails). In other words, if our wealth at step n is worth i, at step $n + 1$, our wealth moves from i to $i + 1$ with probability p, and, from i to $i - 1$ with probability $1 - p$.

3 Some properties of Markov chains

The following result gives an other characterization of a Markov chain. It shows that the probability that a Markov chain follows a given trajectory is completely determined by its initial probability distribution and its transition matrix.

PROPOSITION 3.1. *Let $(X_n)_{n \geq 0}$ be a Markov chain with a state space E, an initial distribution μ and a transition matrix P. Then, for every $i_0, i_1, \dots, i_n \in E$,*

$$\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) = \mu(i_0)P(i_0, i_1) \dots P(i_{n-1}, i_n). \tag{3.1}$$

PROOF. We have

$$\begin{aligned}\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) &= \mathbb{P}(X_0 = i_0, \dots, X_{n-1} = i_{n-1})\mathbb{P}(X_n = i_n | X_0 = i_0, \dots, X_{n-1} = i_{n-1})\\ &= \mathbb{P}(X_0 = i_0, \dots, X_{n-1} = i_{n-1})\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1})\\ &= \mathbb{P}(X_0 = i_0, \dots, X_{n-1} = i_{n-1})P(i_{n-1}, i_n).\end{aligned}$$

The second equality follows from the definition of a Markov chain. Repeating the previous procedure with $\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1})$ leads to

$$\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = \mathbb{P}(X_0 = i_0, \ldots, X_{n-2} = i_{n-2})P(i_{n-2}, i_{n-1}).$$

We then may show by induction that

$$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_0 = i_0)P(i_0, i_1) \ldots P(i_{n-1}, i_n).$$

Hence the result. □

EXAMPLE 3.1. Considering the Example 2.1, we have:

1. $\mathbb{P}(X_0 = 2, X_1 = 1, X_3 = 2, X_4 = 3) = \mu(2)P(2,1)P(1,2)P(2,3) = \frac{1}{6} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{192}$.
2. $\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 2) = 0$, because $P(1,1) = 0$.

REMARKS AND NOTATIONS. Remark that the product of two transition matrix is a transition matrix and, if P is a transition matrix and μ is a probability, then μP is a probability. A probability distribution μ will be identified as a row vector so that μP will be a row vector defined for every $j \in E$ by $(\mu P)(j) = \sum_{i \in E} \mu(i)P(i,j)$.
A function $f : E \to \bar{\mathbb{R}}^+$ will be identified as a column vector and Pf will be the column vector defined for every $i \in E$ by $(Pf)(i) = \sum_{j \in E} P(i,j)f(j)$.
The N-th powers of the matrix P: P^2,..., P^N will be the usual matrix products:

$$P^2(i,k) = \sum_{j \in E} P(i,j)P(j,k), \ldots, \quad \text{with } P^0 = I \qquad (I)_{i,j} = \delta_{i,j} = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{otherwise.} \end{cases}$$

We denote by P_μ the probability with respect to the initial distribution μ: for every event A, $\mathbb{P}_\mu(A) = \mathbb{P}(A|X_0)$ with $X_0 \sim \mu$. If $\mu = \delta_i$ we simply denote $\mathbb{P}_i$ instead of $\mathbb{P}_{\delta_i}$: then, $\mathbb{P}_i(A) = \mathbb{P}(A|X_0 = i)$.

THEOREM 3.1. *Let $(X_n)_n \geq 0$ be a Markov chain with initial law μ and transition matrix P on a state space E. Then,*

1. $\forall n \geq 0, \forall j \in E,$
$$\mathbb{P}(X_n = j) = (\mu P^n)(j) \tag{3.2}$$
where $(\mu P^n)(j)$ is the j-th coordinate of the row vector μP^n.

2. $\forall k, n \geq 0, \forall i, j \in E,$
$$\begin{aligned} \mathbb{P}_i(X_n = j) &= \mathbb{P}(X_n = j \mid X_0 = i) \\ &= \mathbb{P}(X_{n+k} = j | X_k = i) \\ &= P^n(i,j). \end{aligned} \tag{3.3}$$

where $P^n(i,j)$ is the component (i,j) of the matrix P^n.

PROOF. 1. We have : $\forall n \geq 0, \forall j \in E$,

$$\begin{aligned}
\mathbb{P}(X_n = j) &= \sum_{i_0 \in E} \cdots \sum_{i_{n-1} \in E} \mathbb{P}(X_n = j, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\
&= \sum_{i_0 \in E} \cdots \sum_{i_{n-1} \in E} \mu(i_0) P(i_0, i_1) \cdots P(i_{n-1}, j) \\
&= \sum_{i_0 \in E} \cdots \sum_{i_{n-2} \in E} \mu(i_0) P(i_0, i_1) \cdots P(i_{n-3}, i_{n-2}) \underbrace{\sum_{i_{n-1} \in E} P(i_{n-2}, i_{n-1}) P(i_{n-1}, j)}_{=P^2(i_{n-2}, j)} \\
&= \sum_{i_0 \in E} \cdots \sum_{i_{n-2} \in E} \mu(i_0) P(i_0, i_1) \cdots P(i_{n-3}, i_{n-2}) P^2(i_{n-2}, j) \\
&\vdots \\
&= \sum_{i_0 \in E} \mu(i_0) P^n(i_0, j) \\
&= (\mu P^n)(j).
\end{aligned}$$

2. Starting from $X_0 = i$, $(X_n)_{n \geq 0}$ is a Markov chain with initial probability distribution $\mu = \delta_i$. We apply the item 1. above to $\mu = \delta_i$ to get the announced result. □

EXAMPLE 3.2. Let $(X_n)_{n \geq 0}$ be a Markov chain with transition matrix

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

Compute $P^n(1,1) = \mathbb{P}(X_n = 1 | X_0 = 1)$, means, the probability that, starting from the state 1 at time 0, the chain be at the state 1 at time n.

Answer. Let $p_{i,j}^{(n)} = P^n(i,j)$. On one hand, using the equality $P^n = P^{n-1}P$, we may express $p_{1,1}^{(n)}$ in terms of the $p_{1,j}^{(n-1)}$'s, $j = 1, 2$, as follows:

$$p_{1,1}^{(n)} = p\, p_{1,1}^{(n-1)} + (1-q)\, p_{1,2}^{(n-1)}. \tag{3.4}$$

On the other hand, since P^n is a transition matrix, we have:

$$p_{1,1}^{(n-1)} + p_{1,2}^{(n-1)} = 1. \tag{3.5}$$

Then, il follows from equations (3.4) and (3.5) that $p_{1,1}^{(n)} = (p+q-1)p_{1,1}^{(n-1)} + (1-q)$, with $p_{1,1}^{(0)} = 1$ (since $P^0 = I$). Then $p_{1,1}^{(n)}$ is an arithmetico-geometric sequence of the form $u_n = au_{n-1} + b$, which n-th term reads

$$u_n = \begin{cases} u_0 + nb & \text{if } a = 1 \\ a^n\left(u_0 - \frac{b}{1-a}\right) + \frac{b}{1-a} & \text{if } a \neq 1. \end{cases}$$

Hence,

$$p_{1,1}^{(n)} = \begin{cases} \left(p+q-1\right)^n \left(1 - \dfrac{1-q}{2-p-q}\right) + \dfrac{1-q}{2-p-q} & \text{if } p+q < 2 \\ 1 & \text{if } p+q = 2. \end{cases}$$

We see in particular that if $p = q = 0$ then $p_{1,1}^{(n)} = \frac{1+(-1)^n}{2}$. This is expected in fact.

4 States classification: recurrence and transience

4.1 The expected number of visits of a state

Let P be a transition matrix induced by a Markov chain $(X_n)_{n\geq 0}$. Let U be the potential operator associated to P, defined as

$$U = \sum_{k=0}^{\infty} P^k = I + P + P^2 + \cdots + P^n + \cdots$$

Let $j \in E$ and let $N_j = \sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=j\}}$ be the number of times that the chain visits the state j. We have

$$\begin{aligned} U(i,j) = \sum_{k=0}^{+\infty} P^k(i,j) &= \sum_{k=0}^{+\infty} P_i(X_k = j) \\ &= \sum_{k=0}^{+\infty} \mathbb{E}_i\big(\mathbf{1}_{\{X_k=j\}}\big) \\ &= \mathbb{E}_i\Big(\sum_{k=0}^{+\infty} \mathbf{1}_{\{X_k=j\}}\Big) \qquad \text{(using Fubini theorem)} \\ &= \mathbb{E}_i(N_j). \end{aligned}$$

So, we have the following result.

PROPOSITION 4.1. *Leaving at the state $i \in E$, the expected number of visits of the state j by the Markov chain, that is $\mathbb{E}_i(N_j)$, is given by*

$$\mathbb{E}_i(N_j) = U(i,j).$$

The following result gives a way to compute $U(i,j)$, $i,j \in E$. It is a solution to the so-called Dirichlet problem.

PROPOSITION 4.2. *$\forall j \in E$, $U(i,j) = \mathbb{E}_i(N_j)$ is the smallest nonnegative solution of the system of equations:*

$$u(i) = \begin{cases} 1 + (Pu)(i) & \text{if } i = j \\ (Pu)(i) & \text{if } i \neq j. \end{cases} \tag{4.1}$$

REMARK 4.1. The smallest solution means: for any other solution v of the system of equations (4.1), it holds $v(i) \geq u(i), \forall i \in E$. A nonnegative solution is a one satisfying : $u(i) \in [0, +\infty], \forall i \in E$.

EXAMPLE 4.1. Consider Example 2.1 with $E = \{1,2,3\}$ and

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

Compute $\mathbb{E}_1(N_2)$.

Answer. Let $u = (u(1), u(2), u(3))'$. Then Pu is the column vector

$$\Big(\frac{1}{2}u(2) + \frac{1}{2}u(3), \frac{1}{4}u(1) + \frac{1}{2}u(2) + \frac{1}{2}u(3), \frac{1}{3}u(1) + \frac{2}{3}u(2)\Big)',$$

so that the system of equations (4.1) reads

$$\begin{cases} u(1) &= \frac{1}{2}u(2) + \frac{1}{2}u(3) & (1) \\ u(2) &= \frac{1}{4}u(1) + \frac{1}{4}u(2) + \dfrac{1}{2}u(3) + 1 & (2) \\ u(3) &= \frac{1}{3}u(1) + \frac{2}{3}u(2) & (3) \end{cases}$$

Making the transformation $\big(2 \times (1) + (3)$ puis $(1) + 3 \times (3)\big)$ of the first and the second equalities above, we may write $u(1)$ and $u(3)$ in terms of $u(2)$:

$$\begin{cases} u(1) &= u(2) \\ u(2) &= \frac{1}{4}u(1) + \frac{1}{4}u(2) + \frac{1}{2}u(3) + 1 \\ u(3) &= u(2). \end{cases}$$

Now, putting back $u(1)$ and $u(3)$ in (2) lead to

$$\begin{cases} u(1) &= u(2) \\ u(2) &= u(2) + 1 \\ u(3) &= u(2). \end{cases}$$

This is possible only if $u(1) = u(2) = u(3) = +\infty$.

EXAMPLE 4.2. Let $(X_n)_{n\geq 0}$ be a Markov chain on $E = \{1, 2, 3\}$ with transition matrix

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

Compute $\mathbb{E}_i(N_1)$, for $i \in E$.

Answer. Let $u(i) = \mathbb{E}_i(N_1)$, $i \in E = \{1, 2, 3\}$. The system of equations reads:

$$\begin{cases} u(1) &= \frac{1}{2}u(2) + \frac{1}{2}u(3) + 1 \\ u(2) &= \frac{1}{4}u(1) + \frac{1}{4}u(2) + \frac{1}{2}u(3) \\ u(3) &= u(3) \end{cases}$$

which leads to

$$\begin{cases} u(1) &= \frac{6}{5} + u(3) \\ u(2) &= \frac{2}{5} + u(3). \end{cases}$$

As a consequence, any triplet $(\frac{6}{5} + u(3), \frac{2}{5} + u(3), u(3))$ is a solution to the Dirichlet equation. The smallest non negative solution $u = (u(1), u(2), u(3))$ is obtained by putting $u(3) = 0$, so that, $u = (\frac{6}{5}, \frac{2}{5}, 0)$. Finally, $\mathbb{E}_1(N_1) = \frac{6}{5}$, $\mathbb{E}_2(N_1) = \frac{2}{5}$ and $\mathbb{E}_3(N_1) = 0$.

Then starting for example from the state 2 at time 0, the chain makes a visit of the state 1 on average $6/5$ time. Once the chain reaches the state 3, it stays there for ever. We say that the state 3 is a *absorbent state*. We see at the same time that the chain makes a visit of the state 3 a infinite number of times and stays at the states 1 and 2 a finite number of times. The state 3 is said to be a *recurrent* state and the states 1 and 2 are said *transient*. This leads as, in a general setting, to the problem of classifying the states of a Markov chain, by saying which states are recurrent and which one are transient.

4.2 The first passage problem

Let $(X_n)_{n\geq 0}$ be a Markov chain on a state space E, with transition matrix P and let $i, j \in E$. One of the questions of interest is to determine the distribution of the first passage time of the chain at a given state. We will mainly tray to determine the probability that, leaving the state i, the chain makes a visit of the state j: $\mathbb{P}_i(\tau_j < +\infty)$, where τ_j is the hitting time of the state j, defined by

$$\tau_j = \inf\{n \geq 0, X_n = j\}.$$

The computation of the distribution of X_{τ_j} may also be of interest.

Let $f : E \longrightarrow \bar{\mathbb{R}}^+$, and let $\sigma_j = \inf\{n \geq 1, X_n = j\}$ be the first time the chain is back (the first time of return) at the state j. We notice that τ_j and σ_j are both stopping times and that:

- si $i = j$, alors $\mathbb{P}_i(\tau_j = 0) = 1$
- si $i \neq j$, alors $\mathbb{P}_i(\tau_j = \sigma_j) = 1$.

The following result shows how to compute explicitly $\mathbb{P}_i(\tau_j < +\infty)$.

THEOREM 4.1. *1. Let $u(i) = \mathbb{P}_i(\tau_j < +\infty)$, $i \in E$. Then, u is the smallest nonnegative solution of the following system of equations*

$$u(i) = \begin{cases} 1 & \textit{if } i = j \\ (Pu)(i) & \textit{if } i \neq j. \end{cases}$$

2. Let $v(i) = \mathbb{E}_i(\tau_j)$ be the expected time of returning at the state j, leaving the state i. Then v is the smallest nonnegative solution of the following system of equations:

$$v(i) = \begin{cases} 0 & \textit{if } i = j \\ 1 + (Pv)(i) & \textit{if } i \neq j. \end{cases}$$

EXAMPLE 4.3. Let $(X_n)_{n\geq 0}$ be a Markov chain with transition matrix P given for every $i = 0, \cdots, n-1$, by

$$P(i,j) = \begin{cases} p & \text{si } j = i+1 \\ q & \text{si } j = 0 \\ 0 & \text{sinon.} \end{cases}$$

with $0 < p, q < 1, p + q = 1$ and where we suppose that the state n is absorbent. Let

$$\tau = \inf\{k \geq 0, X_k = n\}.$$

Compute $\mathbb{E}_i(\tau)$, for $i = 0, \cdots, n$.

Answer. We know that the function $v(i) = \mathbb{E}_i(\tau)$ is the smallest nonnegative solution of:

$$v(i) = \begin{cases} 0 & \text{si } i = n \\ 1 + (Pv)(i) & \text{si } i \neq n \end{cases}$$

with

$$P = \begin{pmatrix} q & p & 0 & \cdots & \cdots & 0 \\ q & 0 & p & \ddots & & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ \vdots & \vdots & & & \ddots & 0 \\ q & 0 & \cdots & \cdots & \cdots & p \\ 0 & 0 & \cdots & & \cdots & 1 \end{pmatrix}.$$

This leads to the following system of equations:

$$\begin{cases} qv(0) + pv(1) &= v(0) - 1 \\ qv(0) + pv(2) &= v(1) - 1 \\ qv(0) + pv(3) &= v(2) - 1 \\ \vdots & \vdots \\ qv(0) + pv(n) &= v(n-1) - 1. \end{cases}$$

First remark that $v(n) = 0$. On the other hand, we may show, using a backward induction, that $\forall i = 0, \cdots, n-1$,

$$v(i) = 1 + p + \cdots + p^{n-i-1} + q(1 + p + \cdots + p^{n-i-1})v(0).$$

Then we deduce that

$$v(0) = \frac{1-p^n}{p^n(1-p)} \text{ and } v(i) = \frac{1-p^{n-i}}{p^n(1-p)}.$$

In this context, it is intuitive that when p goes to 1, the expected time of reaching the state n, leaving 0, is n. This is the case since $v(0)$ goes to n when p goes to 1. In fact,

$$v(0) = \frac{1-p^n}{p^n(1-p)} = \sum_{j=0}^{n-1} p^{j-n}$$

and the term on the right hand side of the above equation tends towards n when p tends to 1. We show likewise that, in accordance to the intuition, that $v(0)$ goes to $+\infty$ when p goes to 0 and that $v(i)$ goes to $n - i$ when p goes to 1 (and that for every $i = 0, \ldots, n-1$, $v(i)$ goes to $+\infty$ when p goes to 0).

EXERCISE 4.1. Consider Exemple 2.1 where $P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$ and $E = \{1, 2, 3\}$ and determine for every $i \in E$, $\mathbb{P}_i(\tau_2 < +\infty)$.

4.3 States classification

Let $(X_n)_{n\geq 0}$ be a Markov chain with transition probability P and potential operator U. Keep in mind that

$$U = \sum_{k=0}^{\infty} P^k \quad \text{and} \quad U(i,j) = \mathbb{E}_i(N_j),$$

with

$$N_j = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = j\}}.$$

Now, let $\tau_i = \inf\{n \geq 0, X_n = i\}$ and $\sigma_i = \inf\{n \geq 1, X_n = i\}$ be the first hitting time of the state i and the first time of return at the state i, respectively. Let $(\sigma_i^n)_{n\geq 1}$ be the sequence of the successive times of returns at the state i, defined by :

$$\sigma_i^1 = \sigma_i \quad \text{and} \quad \begin{cases} \sigma_i^n = \inf\{k > \sigma_i^{n-1}, X_k = i\} & \text{if } \sigma_i^{n-1} < +\infty \\ \sigma_i^n = +\infty & \text{otherwise.} \end{cases}$$

We have the following result which derives from the Markov property of a Markov chain.

PROPOSITION 4.3. *We have,* $\forall i \in E, \forall n \geq 1$,

$$\mathbb{P}_i(\sigma_i^n < +\infty) = \big(\mathbb{P}_i(\sigma_i < +\infty)\big)^n. \tag{4.2}$$

Remark that

$$N_i = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = i\}} = \mathbf{1}_{\{X_0 = i\}} + \sum_{n=1}^{\infty} \mathbf{1}_{\{\sigma_i^n < +\infty\}}.$$

As a consequence (owing to the previous remark, to Proposition 4.3 and applying Fubini's theorem) we get

$$\begin{aligned} \mathbb{E}_i(N_i) &= \mathbb{P}_i(X_0 = i) + \sum_{n=1}^{\infty} \mathbb{P}_i(\sigma_i^n < +\infty) \\ &= 1 + \sum_{n=1}^{\infty} (\mathbb{P}_i(\sigma_i < +\infty))^n. \end{aligned}$$

Then

$$\mathbb{E}_i(N_i) = U(i,i) = \begin{cases} = +\infty & \text{if } \mathbb{P}_i(\sigma_i < +\infty) = 1 \\ < +\infty & \text{if } \mathbb{P}_i(\sigma_i < +\infty) = a < 1. \end{cases}$$

DEFINITION 4.1. *Let* $(X_n)_{n\geq 0}$ *be a Markov chain with state space* E. *A state* i *is recurrent if*

$$\mathbb{P}_i(\sigma_i < +\infty) = 1$$

and it is transient if

$$\mathbb{P}_i(\sigma_i < +\infty) < 1.$$

The Markov chain is recurrent (transient) if every state is recurrent (transient).

In fact, we have the following result.

THEOREM 4.2. *Let* $i \in E$ *be a state of a Markov chain.*

1. i *is recurrent if and only if*

$$\mathbb{P}_i(N_i = \infty) = 1 \iff \mathbb{E}_i(N_i) = +\infty.$$

2. i *is transient if and only if*

$$\mathbb{P}_i(N_i = \infty) = 0 \iff \mathbb{E}_i(N_i) < +\infty.$$

EXERCISE 4.2. Consider the example where $E = \{0, \cdots, n\}$ and for every $i = 0, \cdots, n-1$,

$$P(i,j) = \begin{cases} p & \text{if } j = i+1 \\ q & \text{if } j = 0 \\ 0 & \text{otherwise} \end{cases}$$

and n is absorbent state. Classify the states of the chain.

EXERCISE 4.3. Consider 4.2 where $(X_n)_{n\geq 0}$ is a Markov chain with transition matrix

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

Compute $\mathbb{E}_i(N_i)$ for every $i \in E = \{1, 2, 3\}$. Deduce a classification of the states of the Markov chain.

EXAMPLE 4.4. *(Symmetric random work on $\mathbb{Z}$)* Consider a mobile moving randomly on $\mathbb{Z}$ following a Markov chain with transition matrix P, which components are defined as

$$P(i,j) = \begin{cases} p & \text{if } j = i+1, \\ 1-p & \text{if } j = i-1, \\ 0 & \text{otherwises.} \end{cases}$$

with $0 < p < 1$. Let $(Z_n)_{n\geq 1}$ be a sequence of iid random variables such that $\mathbb{P}(Z_n = 1) = p$, $\mathbb{P}(Z_n = -1) = 1-p$ et let $X_0 = 0$ and $X_n = Z_1 + \ldots + Z_n$. We have seen that $(X_n)_{n\geq 0}$ is a Markov chain with transition matrix P. Now, let us show that the Markov chain $(X_n)_{n\geq 0}$ is transient if $p \neq 1/2$ and that it is recurrent if $p = 1/2$.

1. Suppose that $p \neq 1/2$. Owing to the law of large numbers we have $\lim\limits_{n\to+\infty} \frac{1}{n} X_n = 2p - 1 \neq 0$. Then,

$$\lim_{n\to+\infty} X_n = \begin{cases} +\infty & \text{if } p > 1/2 \\ -\infty & \text{if } p < 1/2, \end{cases}$$

so that in both cases the $(X_n)_{n\geq 0}$ will be transient.

2. Suppose now that $p = 1/2$ and set $Y_i = \frac{1}{2}(Z_i + 1)$. We know that $T_n = \frac{1}{2}(X_n + n) = \sum_{i=1}^n Y_i$ is binomial random variable with parameters n and $p = 1/2$. So

$$P^{(n)}(0,0) = \mathbb{P}(X_n = 0) = \mathbb{P}\Big(T_n = \frac{n}{2}\Big) = \binom{\frac{n}{2}}{n} 2^{-n}.$$

If $n = 2k$ is an even number then $P^{(2k)}(0,0) = \binom{k}{2k} 4^{-k} = \frac{(2k)!}{(k!)^2}$. On the other hand, using the Stirling formula we get

$$k! \sim \sqrt{2\pi k}\left(\frac{k}{e}\right)^k \quad \text{and} \quad (2k)! \sim \sqrt{4\pi k}\left(\frac{2k}{e}\right)^{2k}.$$

This means that $P^{(2k)}(0,0) \sim (\pi k)^{-1/2}$ as k goes to $+\infty$, so that $U(0,0) = \sum_{n\geq 1} P^n(0,0) = +\infty$ and the state 0 is recurrent. As a consequence the Markov chain is recurrent since every state is accessible from state 0: for every $i \in \mathbb{Z}$, there exists an integer n such that $\mathbb{P}(X_n = i | X_0 = 0) > 0$.

EXAMPLE 4.5. *(Symmetric random walk in $\mathbb{Z}^d$)* Define the probability μ by

$$\mu(x) = \begin{cases} (2d)^{-1} & \text{if } x \in \mathcal{N}_d(0) \\ 0 & \text{otherwise} \end{cases}$$

where $\mathcal{N}_d(0) = \{z \in \mathbb{Z}^d, z = (0, \ldots, 0, \pm 1, 0, \ldots 0)\}$, means, the $2d$ belonging to the neighborhood of $0 \in \mathbb{Z}^d$. For example, $\mathcal{N}_2(0) = \{(0,1), (0,-1), (1,0), (-1,0)\}$. Let $Y_1, \ldots, Y_n$ be a sequence of iid and $\mathbb{Z}^d$-valued random variables with distributions μ. Set $X_0 = 0$ and $X_n = Y_1 + \ldots + Y_n$. We have shown that $(X_n)_{n \geq 0}$ is a Markov chain with initial distribution δ_0 and transition matrix P, which components (i,j) are defined by $P(i,j) = \mu(j-i)$.

1. Let m be the counting measure on $\mathbb{Z}^d$ and $f : \mathbb{Z}^d \to \mathbb{R}$ a integrable function w.r.t. m: $\sum_{z \in \mathbb{Z}^d} |f(z)| < +\infty$. Define the Fourier transform of f by

$$\hat{f}(\theta) = \sum_{z \in \mathbb{Z}^d} f(z) e^{i(\theta, z)}, \quad \theta \in \mathbb{R}^d,$$

where (a,b) stands for the dot product between a and b.

   (a) Show that if $f, g \in L^1(m)$ and the convolution product

$$h(z) = \sum_{y \in \mathbb{Z}^d} f(y) g(z-y),$$

   then $\hat{h}(\theta) = \hat{f}(\theta)\hat{g}(\theta)$, $\forall \theta \in \mathbb{R}^d$.

   (b) Show that if $f \in L^1(m)$, then $\hat{f}$ is bounded and

$$f(z) = \frac{1}{(2\pi)^d} \int_{]-\pi,\pi]^d} \hat{f}(\theta) e^{-i(\theta,z)} d\theta.$$

   (c) Set $\theta = (\theta_1, \ldots, \theta_d)$ Show that

$$\hat{\mu}(\theta) = \frac{1}{d} \sum_{k=1}^{d} \cos(\theta_k).$$

   Deduce that if $\theta \in]-\pi, \pi]^d$, $\theta \neq 0$, then $\hat{\mu}(\theta) < 1$ and that, as $\theta \to 0$,

$$1 - \hat{\mu}(\theta) \sim \frac{|\theta|^2}{2d}.$$

2. Set

$$U_\lambda(x,y) = \sum_{n=0}^{+\infty} \lambda^n P^n(x,y), \quad |\lambda| < 1.$$

   (a) Show that $\lim_{\lambda \to 1} U_\lambda(x,y) = U(x,y)$, where U is the potential matrix associated to P.

   (b) Set $u_\lambda = U_\lambda(0,y)$. Show that $u_\lambda \in L^1(m)$ and that

$$\hat{\mu}_\lambda(\theta) = \frac{1}{1 - \lambda\hat{\mu}(\theta)}.$$

   (c) Compute $U(0,0)$ and show that the Markov chain is recurrent if and only if $d \leq 2$.

References

[1] Baldi, P., Mazliak, L., Priouret, P. *Martingales et chaînes de Markov*, théorie élémentaire et exercices corrigés. Hermann, éditeurs des sciences et des arts.