Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques
Universitat de Barcelona

FINITE MARKOV CHAINS

Lidia Pinilla Peralta

Director:	**David Márquez-Carreras**
Realitzat a:	**Departament de Probabilitat, Lògica i Estadística. UB**

Barcelona, 30 de juny de 2015

Abstract

The purpose of this final project is to study and understand finite Markov chains. First of all, the topic is introduced by definitions and relevant properties. Subsequently, how to classify the chains will be the point in question. Finally, it is analyzed the behaviour of the chain in the long-time. Furthermore, different examples will be exposed throughout the project, to show how Markov chains work and make them easier to comprehend.

Acknowledgments

I would like to express my gratitude to my tutor David Márquez because he has helped me every week since I started this project on February. His advise and assistance have been essential and with his professional guidance I could accomplish this study and I could learned about it as well. My special thanks are extended to my family and friends for their support and consideration during all this period.

Contents

1 Introduction **1**

2 Markov chains in finite time **2**

2.1 Time-homogeneous Markov chains and examples 2

2.2 Chapman-Kolmogorov equation and n-step transition probabilities 10

3 Class structure **13**

3.1 Communicated states. Closed sets. Irreducible chains 13

3.2 Periodic states. Cyclic classes. First classification of the states of a Markov chain 15

3.3 Hitting times and absorption probabilities 18

4 Recurrent and transient states **22**

4.1 Definitions 22

4.2 Rules 25

4.3 Examples 27

5 Invariant distributions **30**

5.1 Invariant distributions or stationary distributions 30

5.2 Ergodicity 34

5.3 Examples 38

6 Limit of a time-homogeneous Markov chain **40**

6.1 Behaviour in mean 40

6.2 Mean recurrence time and invariant distributions 43

6.3 Limiting probability distribution and stationary distribution 45

6.4 Reversibility. Metropolis-Hastings algorithm 47

7 Conclusions **50**

A Probability review **51**

A.1 General outcomes 51

A.2 Conditional probability 52

A.3 Independence 53

A.4 Discrete Random Variables 54

A.5 Expected value 54

B Problems and applications **55**

1 Introduction

Since I began to study the grade of Mathematics in the University of Barcelona, probability and statistics have been the field of this science I enjoyed the most because I think that in this area it is possible to learn theoretically and their application is clear as well. For that reason, I decided that the topic of my final project was about it. During my studies, I have learned the main concepts about probability and with this study I will be able to consolidate and increase my knowledge in this discipline.

With regard to the specific topic, I did know what was a Markov chain thanks to the subject of 'Modelització', where they were briefly introduced to us. So that, I decided to choose this topic to learn it in more detail. It seemed intriguing to me as well, because I also thought that my mathematical skills will be improved by doing it. In particular, this project is focused on finite Markov chains, whose values are discrete. These types of processes have another relevant characteristic as well: the absence of memory, which means that the past is not relevant for the future, what is important is the present.

The project is divided into five parts:

- The first part consists of introduce the concept of Markov chain and explain some properties of them. In fact, the notions defined here will be indispensable for the subsequent parts.
- In the second part, the description of the characteristics of a Markov chain goes deeper. Moreover, it allows to start to classify the chains.
- The aim of the third part is to complete the classification of the Markov chains started in the preceding part. Furthermore, the evolution that the chain has over the time will take up an important place in this part.
- The fourth part goes beyond the previous ones, by starting to study the long-time properties of Markov chains. With statements and its demonstrations, the understanding will be better.
- In the fifth and final part, the explanation of the behaviour of Markov chains in the long-time will conclude as well as this study.

Throughout all the project, the presence of examples will be an important point in order to comprehend how Markov chains work.

Additionally, two annexes are included:

On one hand, the first annex provides the main outcomes studied in the subject of 'Probabilitats'. These results are important during all the project and their use in it will be continuous.

On the other hand, the second annex collects various exercises made in parallel as the theoretical project. This supplement allows to complete the topic.

2 Markov chains in finite time

We start introducing the definition of a Markov chain and we study the first properties of them. In particular, we focus on a specific type of Markov chains, which have an important property as we will see. After that, we determine what is a stopping time, a valuable notion that will be useful in the following chapters. We work with these new concepts and we give examples to clarify them. Subsequently, the concepts of Chapman-Kolmogorov equation and n-step transition probabilities will help us to study the evolution of the chain over the states. Finally, we show two applications of these ideas.

Throughout all this chapter we consider a probability space $(\Omega, \mathcal{A}, P)$ and a collection of discrete random variables $X_n : \Omega \to I$, where I is a countable set called set of states or state space. Therefore, every $i \in I$ defines a state.

2.1 Time-homogeneous Markov chains and examples

In this section we define what is a Markov chain, in particular, we work with what we call time-homogeneous Markov chain. Furthermore, we show some characteristics of these chains and various examples in order to understand them better.

Firstly, we introduce a necessary concept to the definition of a Markov chain.

Definition 2.1.1.: A stochastic process is a collection of random variables that are included into a random set $\{X_t : t \in T\}$; for every t, we have: $X_t : \Omega \to \mathbb{R}$. In other words, it represents the evolution of a system of random values over time.

We will study the case where T is discrete, that is: $T = \mathbb{Z}^+ = \{0, 1, 2, \dots\}$. Therefore, we will have the process: $\{X_n : n \geq 0\}$.

So now we can consider the following notion.

Definition 2.1.2. A matrix $\Pi = (p_{i,j} : i, j \in I)$ is a stochastic matrix or a transition matrix if

(a) $p_{i,j} \in [0, 1]$.

(b) $\sum_{j \in I} p_{i,j} = 1$, $\forall i \in I$.

In addiction, $p_{i,j}$, $\forall i, j \in I$, are called transition probabilities.

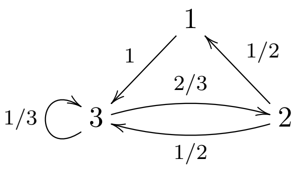
Observation 2.1.3. A stochastic process can be represented by transition matrices as well as by state diagrams.

Examples:

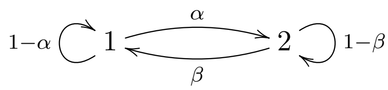
(i) Consider a stochastic process with transition matrix

$$\begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \\ 0 & 2/3 & 1/3 \end{pmatrix}$$

Then the state diagram is



(ii) Consider a stochastic process with state diagram



Then the transition matrix is

$$\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

In the following result we describe what is necessary to have a Markov chain.

Definition 2.1.4. A stochastic process $\{X_n : n \geq 0\}$, whose values are in a set of states I, is a Markov Chain with initial distribution $\gamma = \{\gamma_i : i \in I\}$ and transition matrix $\Pi = (p_{i,j} : i, j \in I)$ if

(a) $X_0 \overset{\mathcal{D}}{=} \gamma$, which means that for $i \in I$ we have $P(X_0 = i) = \gamma_i$.

(b) $P(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j \mid X_n = i)$, $\forall i_0, \ldots, i_{n-1}, i, j \in I$, $\forall n \geq 0$.

Observation 2.1.5. The last equality is called Markov property and it points out that given the present state, the future and the past states are independent.

As we have said before, we will focus on a particular type of Markov chains.

Definition 2.1.6. A Markov Chain defined as in **Definition 2.1.4.** is called a time-homogeneous Markov chain if

$$P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i) = p_{i,j}, \forall n \geq 0, \forall i, j \in I.$$

Moreover, we will write it as $HMC(\gamma, \Pi)$.

Observation 2.1.7. The definition above shows that the transition probability $p_{i,j}$ is independent of n.

Theorem 2.1.8. A stochastic process $\{X_n : n \geq 0\}$, whose values are in the set of states I, is a time-homogeneous Markov Chain (γ, Π) if, and only if,

$$P(X_0 = i_0, \ldots, X_n = i) = \gamma_{i_0} p_{i_0,i_1} \ldots p_{i_{n-1},i}, \qquad \forall i_0, \ldots, i_{n-1}, i \in I, \forall n \geq 0,$$

where $p_{i_0,i_1}, \ldots, p_{i_{n-1},i}$ are the transition probabilities associated with Π.

Proof.

$\boxed{\Rightarrow}$ Using the general product rule, we have

$$\begin{aligned}
&P\left(X_0 = i_0, \ldots, X_n = i\right) = \\
&= P\left(X_0 = i_0\right) P\left(X_1 = i_1 \mid X_0 = i_0\right) \times \ldots \times P\left(X_n = i \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\right) \\
&= \gamma_{i_0} p_{i_0,i_1} \ldots p_{i_{n-1},i}.
\end{aligned}$$

$\boxed{\Leftarrow}$ The first condition of the Markov property is satisfied because

$P\left(X_0 = i_0\right) = \sum_{i_1 \in I} \gamma_{i_0} p_{i_0,i_1} = \gamma_{i_0} \sum_{i_1 \in I} p_{i_0,i_1} = \gamma_{i_0}$. Consequently, $X_0 \overset{\mathcal{D}}{=} \gamma$.

Now, we study the second condition of the Markov property. On one hand, we have

$$\begin{aligned}
P\left(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i\right) &= \frac{P\left(X_0 = i_0, \ldots, X_n = i, X_{n+1} = j\right)}{P\left(X_0 = i_0, \ldots, X_n = i\right)} \\
&= \frac{\gamma_{i_0} p_{i_0,i_1} \cdots p_{i,j}}{\gamma_{i_0} p_{i_0,i_1} \cdots p_{i_{n-1},i}} = p_{i,j}.
\end{aligned}$$

On the other hand, we get

$$\begin{aligned}
P\left(X_{n+1} = j \mid X_n = i\right) &= \frac{P\left(X_n = i, X_{n+1} = j\right)}{P\left(X_n = i\right)} \\
&= \frac{\sum\limits_{i_0,\ldots,i_{n-1} \in I} P\left(X_0 = i_0, \ldots, X_n = i, X_{n+1} = j\right)}{\sum\limits_{i_0,\ldots,i_{n-1} \in I} P\left(X_0 = i_0, \ldots, X_n = i\right)} \\
&= \frac{\sum\limits_{i_0,\ldots,i_{n-1} \in I} \gamma_{i_0} p_{i_0,i_1} \cdots p_{i_{n-1},i} p_{i,j}}{\sum\limits_{i_0,\ldots,i_{n-1} \in I} \gamma_{i_0} p_{i_0,i_1} \cdots p_{i_{n-1},i}} = p_{i,j}.
\end{aligned}$$

Therefore, the second condition of the Markov property is satisfied. In addiction, as $p_{i,j}$ does not depend on n, we also obtain that this is a time-homogeneous Markov chain.

□

Proposition 2.1.9. Let $\{X_n : n \geq 0\}$ be a $HMC\left(\gamma, \Pi\right)$. Then, $\{X_{n+m} : m \geq 0\}$ is a $HMC\left(\mathcal{L}\left(X_m\right), \Pi\right)$, where $\mathcal{L}\left(X_m\right)$ is the law of the discrete random variable X_m.

Proof. The result follows by **Theorem** 2.1.8.

$$\begin{aligned}
&P\left(X_{0+m} = j_0, \ldots, X_{n+m} = j_n\right) = \\
&= \sum_{i_0,\ldots,i_{m-1} \in I} P\left(X_0 = i_0, \ldots, X_{m-1} = i_{m-1}, X_m = j_0, \ldots, X_{n+m} = j_n\right) \\
&= \sum_{i_0,\ldots,i_{m-1} \in I} \gamma_{i_0} p_{i_0,i_1} \ldots p_{i_{m-1},j_0} p_{j_0,j_1} \ldots p_{j_{n-1},j_n} = p_{j_0,j_1} \ldots p_{j_{n-1},j_n} P\left(X_m = j_0\right).
\end{aligned}$$

□

Now, we introduce the concept of stopping time, which will be important in successive chapters.

Definition 2.1.10 A random variable $T : \Omega \longrightarrow \{1, 2, \ldots\} \cup \{\infty\}$ such that the events $\{T = n\}$ only depend on $X_0, \ldots, X_n$, for $n \geq 0$, is called stopping time.

We also study an important characteristic related to time-homogeneous Markov chains and stopping times, which is called strong Markov property.

Theorem 2.1.11. Let $\{X_n : n \geq 0\}$ be a $HMC\,(\gamma, \Pi)$ and assume that T is a stopping time. Then, $\{X_{T+n} : n \geq 0\}$ is a $HMC\,(\delta_i, \Pi)$ conditionally on $T < \infty$ and $X_T = i$. Moreover, it is independent of $X_0, X_1, \ldots, X_T$.

Proof. Consider an event B determined by $X_0, \ldots, X_T$; so that, $B \cap \{T = m\}$ only depends on $X_0, X_1, \ldots, X_m$. Then, using the Markov property we have

$$
\begin{aligned}
&P\left(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_T = i\}\right) \\
&= P\left(\{X_m = j_0, X_{m+1} = j_1, \ldots, X_{m+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_m = i\}\right) \\
&= P\left(X_m = j_0, X_{m+1} = j_1, \ldots, X_{m+n} = j_n \mid B \cap \{T = m\} \cap \{X_m = i\}\right) \\
&\quad \times P\left(B \cap \{T = m\} \cap \{X_m = i\}\right) \\
&= P\left(X_m = j_0, X_{m+1} = j_1, \ldots, X_{m+n} = j_n \mid X_m = i\right) P\left(B \cap \{T = m\} \cap \{X_T = i\}\right) \\
&= P\left(X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n \mid X_0 = i\right) P\left(B \cap \{T = m\} \cap \{X_T = i\}\right).
\end{aligned}
$$

Taking the sum from $m = 0$ until infinity on both sides of the previous equality, we get

$$
\begin{aligned}
&P\left(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \cap \{T < \infty\} \cap \{X_T = i\}\right) \\
&= P_i\left(X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n\right) P\left(B \cap \{T < \infty\} \cap \{X_T = i\}\right).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\frac{P\left(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \cap \{T < \infty\} \cap \{X_T = i\}\right)}{P\left(\{T < \infty\} \cap \{X_T = i\}\right)} \\
&= \frac{P_i\left(X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n\right) P\left(B \cap \{T < \infty\} \cap \{X_T = i\}\right)}{P\left(\{T < \infty\} \cap \{X_T = i\}\right)}.
\end{aligned}
$$

Finally we obtain

$$
\begin{aligned}
&P\left(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \mid T < \infty, X_T = i\right) \\
&= P_i\left(X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n\right) P\left(B \mid T < \infty, X_T = i\right).
\end{aligned}
$$

□

Examples:

(i) Random walk on $\mathbb{Z}$.

Suppose a particle moving along a straight line in unit steps. Each step is one unit to the right with probability p or one unit to the left with probability $q = 1 - p$. The states are the possible positions.

Let $X_0 = 0$ be the initial distribution, which is the constant 0; it means that X_0 is the initial position. We consider $X_n = X_0 + \xi_1 + \ldots + \xi_n$, where $P(\xi_n = 1) = p$

and $P(\xi_n = -1) = q = 1 - p$, with $p \in (0,1)$ and $\{\xi_n : n \geq 1\}$ are independent and identically distributed random variables; that is to say that ξ_n is the movement in the nth stage.

Now, we verify that $\{X_n : n \geq 0\}$ is a Markov chain by proving that it satisfies the Markov property. Using that $\{\xi_n : n \geq 1\}$ are independent and identically distributed, for any $i_0, \ldots, i_{n-1}, i, j \in I$, we have

$$\begin{aligned}
&P(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = \\
&= \frac{P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j)}{P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)} \\
&= \frac{P(X_0 = i_0, \xi_1 = i_1 - i_0, \xi_2 = i_2 - i_1, \ldots, \xi_n = i - i_{n-1}, \xi_{n+1} = j - i)}{P(X_0 = i_0, \xi_1 = i_1 - i_0, \ldots, \xi_n = i - i_{n-1})} \\
&= \frac{P(X_0 = i_0, \xi_1 = i_1 - i_0, \xi_2 = i_2 - i_1, \ldots, \xi_n = i - i_{n-1}) P(\xi_{n+1} = j - i)}{P(X_0 = i_0, \xi_1 = i_1 - i_0, \ldots, \xi_n = i - i_{n-1})} \\
&= P(\xi_{n+1} = j - i) := p_{i,j}.
\end{aligned}$$

On the other hand, we get

$$\begin{aligned}
P(X_{n+1} = j \mid X_n = i) &= \frac{P(X_n + \xi_{n+1} = j, X_n = i)}{P(X_n = i)} \\
&= \frac{P(\xi_{n+1} = j - i, X_n = i)}{P(X_n = i)} \\
&= \frac{P(\xi_{n+1} = j - i) P(X_n = i)}{P(X_n = i)} \\
&= P(\xi_{n+1} = j - i) := p_{i,j}.
\end{aligned}$$

So that, $\{X_n : n \geq 0\}$ satisfies the Markov property. Furthermore, as $p_{i,j}$ does not depend on n, we obtain that this is a time-homogeneous Markov chain.

Finally, we write the transition matrix of this process:

$$\Pi = \begin{pmatrix}
\ddots & & & & & & \iddots \\
& 0 & p & 0 & 0 & 0 & \\
& q & 0 & p & 0 & 0 & \\
\cdots & 0 & q & 0 & p & 0 & \cdots \\
& 0 & 0 & q & 0 & p & \\
& 0 & 0 & 0 & q & 0 & \\
\iddots & & & & & & \ddots
\end{pmatrix}$$

The previous concepts of this example prove the next result.

Proposition 2.1.12. Let $\{Y_n : n \geq 1\}$ be a sequence of random variables, which are independent and identically distributed, whose values are in I. Then, $\left\{X_n = \sum_{i=1}^{n} Y_i : n \geq 0\right\}$ is a time-homogeneous Markov chain.

Generally, we have the following result.

Proposition 2.1.13. Let $\{Z_n : n \geq 1\}$ be a sequence of random variables, which are independent and identically distributed, such that $Z_n : \Omega \to I$. Consider $f : I \times I \to I$ and a random variable X_0, independent of $\{Z_n : n \geq 1\}$ and whose values are in I. Then, if we define $X_{n+1} = f(X_n, Z_{n+1})$, with $n \geq 0$, then $\{X_n : n \geq 0\}$ is a time-homogeneous Markov chain.

Proof. Firstly, we study the Markov property. Using that $\{Z_n : n \geq 1\}$ are independent and identically distributed, we get

$$
\begin{aligned}
& P(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) \\
&= \frac{P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j)}{P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)} \\
&= \frac{P(X_0 = i_0, f(X_0, Z_1) = i_1, \ldots, f(X_{n-1}, Z_n) = i, f(X_n, Z_{n+1}) = j)}{P(X_0 = i_0, f(X_0, Z_1) = i_1, \ldots, f(X_{n-1}, Z_n) = i)} \\
&= \frac{P(X_0 = i_0, f(i_0, Z_1) = i_1, \ldots, f(i_{n-1}, Z_n) = i, f(i, Z_{n+1}) = j)}{P(X_0 = i_0, f(i_0, Z_1) = i_1, \ldots, f(i_{n-1}, Z_n) = i)} \\
&= \frac{P(X_0 = i_0, f(i_0, Z_1) = i_1, \ldots, f(i_{n-1}, Z_n) = i) \, P(f(i, Z_{n+1}) = j)}{P(X_0 = i_0, f(i_0, Z_1) = i_1, \ldots, f(i_{n-1}, Z_n) = i)} \\
&= P(f(i, Z_{n+1}) = j) = p_{i,j}.
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
& P(X_{n+1} = j \mid X_n = i) = \frac{P(X_{n+1} = j, X_n = i)}{P(X_n = i)} = \frac{P(f(X_n, Z_{n+1}) = j, X_n = i)}{P(X_n = i)} \\
&= \frac{P(f(i, Z_{n+1}) = j, X_n = i)}{P(X_n = i)} = \frac{P(f(i, Z_{n+1}) = j) \, P(X_n = i)}{P(X_n = i)} \\
&= P(f(i, Z_{n+1}) = j) = p_{i,j}.
\end{aligned}
$$

Consequently, $\{X_n : n \geq 0\}$ verifies the Markov property. Moreover, as $p_{i,j}$ does not depend on n, we obtain that this is a time-homogeneous Markov chain.

□

(ii) Random walk on $\mathbb{Z}$ with absorbing barriers (also known as 'The Gambler's Ruin Problem').

Suppose two gamblers, A and B, who play at tossing a coin (so that, the possible outcomes are heads or tails) and they have a euros and b euros, respectively. The game finishes when one of the players get ruined.

We also consider that on each successive gamble, the player A wins 1 euro if the result is a head or loses 1 euro if the result is a tail, with probabilities p and $q = 1 - p$, respectively.

So that, the stochastic process in this case is: $\{X_n : n \geq 0\}$ and it points out the evolution of the capital that has the player A. Moreover, we write $X_0 = a$ and $X_{n+1} = X_n + \xi_{n+1}$, such that $X_n : \Omega \to \{0, 1, 2, \ldots, a + b\}$. Then, the initial distribution is

$\gamma = \delta_{\{a\}}$[1] and the transition matrix is

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

(iii) Random walk on $\mathbb{Z}$ with reflecting barriers.

The idea is the same as the previous example but now, when one of the players gets ruined, the other one gives him one euro, so that the gamble does not end. Consequently, the transition matrix in this case is

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

(iv) Model for the inventory management.

A certain product is put up for sale to satisfy the requests of the consumers. Let Z_{n+1} be the number of demanded unities of the product between the moments n and $n+1$. Let X_0 be the initial value of the stock and $\{Z_n : n \geq 1\}$ independent and identically distributed random variables and independents of X_0 as well. We also know that the stock is updated after every stage. Therefore, let X_n be the value of the stock in the nth stage.

We consider the values s and S, which are the critical point and the maximum point, respectively, such that $0 < s < S$. In that way, if $X_n \leq s$, then we place the stock at the stage S, otherwise (that is to say that $X_n > s$) we keep the stock were it was.

Moreover, we assume that $X_0 \leq S$ and we write: $X_n = \{S, S-1, S-2, \ldots\}$, with $n \geq 0$. In addiction, considering that we can obtain negative values if the request is not handled, but it will be handled immediately after the replacement of the stock; then X_n can be expressed by the recurrence

$$X_{n+1} = \begin{cases} X_n - Z_{n+1}, & \text{if} \quad s < X_n \leq S, \\ S - Z_{n+1}, & \text{if} \quad X_n \leq s. \end{cases}$$

Using a similar argument that the one used in **Example (i)**, it can be proved that this model is represented by a time-homogeneous Markov chain.

[1] δ is called Dirac delta function and it is defined as

$$\delta_{\{a\}} = \begin{cases} 0, & \text{if } x \neq a, \\ 1, & \text{if } x = a. \end{cases}$$

(v) Ehrenfest model for statistical mechanics. A simple version.

This model allows us to represent the diffusion by a porous material, a membrane for example. Furthermore, it can determine the heat exchange between two systems with different temperatures.

Suppose that we have two boxes, labeled A and B, that contain a total of N particles. Let X_n be the number of particles in A at the moment n. Just after the moment n we take one particle, no matter in which box is located, and we change it to the other box at the moment $n+1$. So that, if we write $X_n = i$, with $i \in \{0, 1, \ldots, N\}$, then

$$X_{n+1} = \begin{cases} i-1, & \text{if the particle was in box } A, \\ i+1, & \text{if the particle was in box } B. \end{cases}$$

Now, we study the initial distribution of this process: if $X_0 = n_0$, then $\gamma = \delta_{\{n_0\}}$. Moreover, considering that if a particle is located in the state i, then the transition is only possible to the states $i-1$ and $i+1$, we have that the transition probabilities in this case are

$$p_{i,j} = \begin{cases} P\left(X_{n+1} = i+1 \mid X_n = i\right) = \frac{N-i}{N}, & \text{if } j = i+1, \\ P\left(X_{n+1} = i-1 \mid X_n = i\right) = \frac{i}{N}, & \text{if } j = i-1, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, the transition matrix that we obtain is

$$\Pi = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{N} & 0 & \frac{N-2}{N} & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{N} & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{3}{N} & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & \frac{N-2}{N} & 0 & \frac{2}{N} & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{N-1}{N} & 0 & \frac{1}{N} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}$$

Observation 2.1.14. This matrix and the one obtained in **Example (iii)** (Random walk on $\mathbb{Z}$ with reflecting barriers) are similar; the difference between both is that in this case the transition probability $p_{i,j}$ is not independent of i.

Finally, we verify that this is a time-homogeneous Markov chain. We write the model as: $X_{n+1} = X_n + Z_{n+1}$, with $n \geq 0$, $Z_n \in \{+1, -1\}$ and $P\left(Z_{n+1} = 1 \mid X_n = i\right) = \frac{N-i}{N}$, $P\left(Z_{n+1} = -1 \mid X_n = i\right) = \frac{i}{N}$. Then, it can be expressed recursively as:

$$X_{n+1} = f\left(Z_{n+1}, g\left(X_0, Z_1, \ldots, Z_n\right)\right),$$

that is to say that X_n is function of $X_0, Z_1, \ldots, Z_n$. Furthermore, if we define $X_n := g\left(X_0, Z_1, \ldots, Z_n\right)$, we obtain that Z_{n+1} and X_n are independent and then we can apply the **Proposition** 2.1.13., which implies that $\{X_n : n \geq 0\}$ is a time-homogeneous Markov chain.

2.2 Chapman-Kolmogorov equation and n-step transition probabilities

In this section, we study what is the probability that a Markov chain can be located in a certain state after n steps. In addiction, we work on the Chapman-Kolmogorov equation, which allow us to express this probability. Finally, we show some examples to assimilate this concept.

Definition 2.2.1. Let $\{X_n : n \geq 0\}$ be a time-homogeneous Markov chain. We define

$$p_{i,j}^{(m)} = P\left(X_{n+m} = j \mid X_n = i\right),$$

with $n, m \geq 0$ and $i, j \in I$, as the probability of, staying in the state i, we arrive to the state j in m steps; that concept is known as n-step transition probabilities. Moreover, the m-step transition matrix is

$$\Pi_m = \left(p_{i,j}^{(m)} : i, j \in I\right).$$

Observation 2.2.2. If $m = 1$, then $p_{i,j}^{(1)} = P\left(X_{n+1} = j \mid X_n = i\right) = p_{i,j}$, which is the same case that we have already studied.

Proposition 2.2.3. Let $\{X_n : n \geq 0\}$ be a time-homogeneous Markov chain and Π_m the m-step transition matrix. Then, $\Pi_m = \Pi^m = \Pi \times \overset{m)}{\cdots} \times \Pi$.

Proof. We should prove that $p_{i,j}^{(m)} = \sum_{k \in I} p_{i,k}^{(m-1)} p_{k,j}$, for $m \geq 2$; that is to say that $\Pi_m = \Pi_{m-1}\Pi$. Therefore, we have

$$\begin{aligned}
p_{i,j}^{(m)} &= P\left(X_{n+m} = j \mid X_n = i\right) = \frac{P\left(X_{n+m} = j, X_n = i\right)}{P\left(X_n = i\right)} \\
&= \frac{\sum_{k \in I} P\left(X_{n+m} = j, X_{n+m-1} = k, X_n = i\right)}{P\left(X_n = i\right)} \\
&= \sum_{k \in I} P\left(X_{n+m} = j \mid X_{n+m-1} = k, X_n = i\right) P\left(X_{n+m-1} = k \mid X_n = i\right) \\
&= \sum_{k \in I} p_{i,k}^{(m-1)} p_{k,j}.
\end{aligned}$$

□

Corollary 2.2.4. With the hypothesis of the previous proposition, we can obtain

- Π^m is a transition matrix.
- Considering that $\Pi^{l+k} = \Pi^l \Pi^k$, with $l, k \geq 0$, then $p_{i,j}^{(l+k)} = \sum_{h \in I} p_{i,h}^{(l)} p_{h,j}^{(k)}$, which is known as Chapman-Kolmogorov equation.
- If the Chapman-Kolmogorov equation is repeated by iteration, we get

$$p_{i,j}^{(m)} = \sum_{i_1 \in I} p_{i,i_1} p_{i_1,j}^{(m-1)} = \sum_{i_1, \ldots, i_{m-1} \in I} p_{i,i_1} p_{i_1,i_2} \cdots p_{i_{m-1},j}.$$

The following result shows what is the law or distribution of X_n:

Proposition 2.2.5. Let $\{X_n : n \geq 0\}$ be a $HMC(\gamma, \Pi)$. Then, $\forall k \in I$, we have

$$P(X_n = k) = \gamma_k^{(n)},$$

where $\gamma^{(n)} = \gamma \Pi^n$ is the law of X_n.

Proof. $P(X_n = k) = \sum_{h \in I} P(X_n = k \mid X_0 = h) P(X_0 = h) = \sum_{h \in I} \gamma_h p_{h,k}^{(n)}$,

where $\gamma_h = P(X_0 = h)$.

$\square$

Corollary 2.2.6. With the hypothesis of the previous proposition, we get

- $P(X_{l+k} = i) = \sum_{j \in I} P(X_{l+k} = i \mid X_l = j) P(X_l = j) = \sum_{j \in I} \gamma_j^{(l)} p_{j,i}^{(k)}$. In consequence, $\gamma^{(l+k)} = \gamma^{(l)} \Pi^k$.
- $\gamma_j^{(n)} = P(X_n = j) = \sum_{h \in I} \gamma_h p_{h,j}^{(n)} = \sum_{h \in I} \sum_{i_1, \ldots, i_{n-1} \in I} \gamma_h p_{h,i_1} p_{i_1,i_2} \ldots p_{i_{n-1},j}$.

The proposition below studies the distributions of the process in finite dimension.

Proposition 2.2.7. The initial distribution γ and the transition matrix Π establish the law of the random vector $(X_{n_1}, \ldots, X_{n_k})$, with $n_1, \ldots, n_k \geq 0$.

Proof. To show this result, we use the general product rule.

Firstly, we study the case $(X_0, \ldots, X_k)$. So that, we have

$$\begin{aligned} & P(X_0 = i_0, \ldots, X_{k-1} = i_{k-1}, X_k = i_k) = \\ & = P(X_0 = i_0) P(X_1 = i_1 \mid X_0 = i_0) \ldots P(X_k = i_k \mid X_0 = i_0, \ldots, X_{k-1} = i_{k-1}) \\ & = \gamma_{i_0} p_{i_0,i_1} p_{i_1,i_2} \ldots p_{i_{k-1},i_k}. \end{aligned}$$

Now, in general, the law of the vector $(X_{n_1}, \ldots, X_{n_k})$ with $n_1 < n_2 < \ldots < n_k$ is

$$\begin{aligned} & P\left(X_{n_1} = i_1, \ldots, X_{n_{k-1}} = i_{k-1}, X_{n_k} = i_k\right) = \\ & = P(X_{n_1} = i_1) P(X_{n_2} = i_2 \mid X_{n_1} = i_1) \ldots P\left(X_{n_k} = i_k \mid X_{n_1} = i_1, \ldots, X_{n_{k-1}} = i_{k-1}\right) \\ & = \sum_{h \in I} \gamma_h p_{h,i_1}^{(n_1)} p_{i_1,i_2}^{(n_2 - n_1)} \ldots p_{i_{k-1},i_k}^{(n_k - n_{k-1})}. \end{aligned}$$

$\square$

To conclude this section we study what is the probability $p_{i,i}^{(n)}$ in the following important two examples:

(i) The most general two-state chain, with $\alpha, \beta > 0$, has transition matrix of the form

$$\Pi = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

And it is represented by the following state diagram

$$\begin{tikzcd} 1 \arrow[loop left, "1-\alpha"] \arrow[rr, bend left, "\alpha"] & & 2 \arrow[ll, bend left, "\beta"] \arrow[loop right, "1-\beta"] \end{tikzcd}$$

To get the value of $p_{1,1}^{(n)}$ we use the relation $\Pi^{n+1} = \Pi^n \Pi$. Then

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

On the other hand, we know that $p_{1,1}^{(n)} + p_{1,2}^{(n)} = 1$. So that, we get

$$p_{1,1}^{(n+1)} = p_{1,1}^{(n)} (1-\alpha-\beta) + \beta.$$

Therefore, we have the following recurrence relation for $p_{1,1}^{(n)}$

$$\begin{cases} p_{1,1}^{(n+1)} &= p_{1,1}^{(n)} (1-\alpha-\beta) + \beta, \\ p_{1,1}^{(0)} &= 1. \end{cases}$$

Solving the previous recurrence relation by induction or using the geometric series, we obtain that it has a unique solution, which is

$$p_{1,1}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} (1-\alpha-\beta)^n .$$

Observation 2.2.8. The case $\alpha + \beta = 0$ is not possible because we have supposed that $\alpha, \beta > 0$.

(ii) Virus mutation.

Suppose a virus can exist in N different strains and in each generation either stays the same, or with probability α mutates to another strain, which is chosen at random. What is the probability that the strain in the nth generation is the same as the original?

We can model this process as a Markov chain with N states: $\{1, \dots, N\}$ and a transition matrix Π given by the next transition probabilities

$$p_{i,j} = \begin{cases} 1-\alpha, & \text{if } i = j, \\ \frac{\alpha}{N-1}, & \text{if } i \neq j. \end{cases}$$

Then, we need to compute $p_{1,1}^{(n)}$ and to get that we can assume that there are only two states: the initial state (state 1) and all the remaining states, which can be considered as one state (state 2). Therefore, we have a two-state time-homogeneous Markov chain that can be represented by the following state diagram

$$\begin{tikzcd} 1 \arrow[loop left, "1-\alpha"] \arrow[r, bend left, "\alpha"] & 2 \arrow[l, bend left, "\beta"] \arrow[loop right, "1-\beta"] \end{tikzcd}$$

where $\beta = \dfrac{\alpha}{N-1}$.

Finally, using the result obtained in the previous example, we get

$$p_{1,1}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} (1-\alpha-\beta)^n = \frac{1}{N} + \frac{N-1}{N} \left(1 - \frac{\alpha N}{N-1}\right)^n .$$

Observation 2.2.9. The last two examples we have already studied are not difficult because we have considered time-homogeneous Markov chains with two states and the matrices obtained are simple. However, when there are more states, we can get matrices whose powers can be difficult to compute.

3 Class structure

In this chapter, we begin to classify Markov chains. Firstly, we define concepts as states that communicate and absorbing, essential and passing through states. Moreover, we study when a Markov chain is or not irreducible, a notion related to the number of equivalence classes that has the chain. We also give some examples to clarity these ideas.

Subsequently, we study the periodicity that have the states, a notion that allows us to define what are the cyclic subclasses. We illustrate it all by some examples as well. Finally, we study the concepts of absorption probabilities and hitting times, which are related to the evolution of the chain over the time. Furthermore, we prove important results that involve the previous concepts and we show their application with examples.

Throughout all this chapter we consider a probability space $(\Omega, \mathcal{A}, P)$ and a collection of discrete random variables $X_n : \Omega \to I$, where I is a countable set and $\{X_n : n \geq 0\}$ is a $HMC\,(\gamma, \Pi)$ with $\Pi = (p_{i,j}; i, j \in I)$.

3.1 Communicated states. Closed sets. Irreducible chains

In this section we work with Markov chains by dividing it into small parts; so that, it is simpler to study them and they together explain the global behavior of the chain.

Definition 3.1.1. A state $j \in I$ is called accessible from a state $i \in I$ (written as $i \longrightarrow j$) if there exists $n \geq 0$ such that

$$P\left(X_n = j | X_0 = i\right) = p_{i,j}^{(n)} > 0.$$

Definition 3.1.2. A state $i \in I$ communicates with state $j \in I$ (written as $i \longleftrightarrow j$) if both $i \longrightarrow j$ and $j \longrightarrow i$.

Proposition 3.1.3. Given two different states $i, j \in I$, the following conditions are equivalent

(a) $i \longrightarrow j$.

(b) There exist states $i_1, i_2, \ldots, i_{n-1} \in I$ such that

$$p_{i,i_1} p_{i_1,i_2} \ldots p_{i_{n-1},j} > 0.$$

Proof. Considering that $p_{i,j}^{(n)} = \sum\limits_{i_1,\ldots,i_{n-1} \in I} p_{i,i_1} p_{i_1,i_2} \ldots p_{i_{n-1},j}$, we obtain directly the equivalence above. □

Observation 3.1.4.

- The communicate property is an equivalence relation in the set of states I.
- It can also be possible to make a partition of the set of states I in equivalence classes.

Definition 3.1.5. C is a closed class if, given $i \in C$ and $i \longrightarrow j$, then $j \in C$. It means that it is impossible to leave this class C. Moreover, we have that $\sum\limits_{j \in C} p_{i,j} = 1$, $\forall i \in C$.

Definition 3.1.6. A state $i \in I$ is called absorbing if $\{i\}$ is a closed class; in other words, if $p_{i,i} = 1$.

Definition 3.1.7. A Markov chain is irreducible if there is only one equivalence class; that is to say that all the states are communicated.

Definition 3.1.8. A state $i \in I$ is called inessential or passing through state if there exist $m \geq 1$ and $j \in I$ such that $p_{i,j}^{(m)} > 0$ but $p_{j,i}^{(n)} = 0, \forall n \geq 1$. In other words, it is possible to leave the state i but it is impossible to return to it. In addiction, the complementary of an inessential set of states is called essential set of states; it is composed of absorbing states and states where the leaving is possible and the returning as well.

Examples:

(i) Random walk on $\mathbb{Z}$.

As all the states communicate with one another, they are all essential and this chain has only one class, which is $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$; so that, the chain is irreducible.

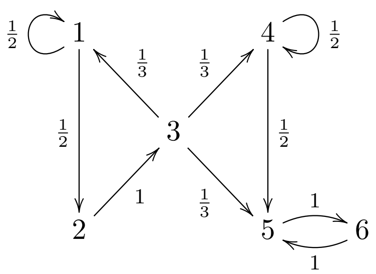
(ii) Random walk on $\mathbb{Z}$ with absorbing barriers.

On one hand, as 0 and N are absorbing states, we have that these states are essential. On the other hand, the states $1, \ldots, N-1$ are passing through states because it is possible to leave them, that is to say that you go to the absorbing states, but the retuning is not possible. Therefore, $\{0\}$, $\{1, \ldots, N-1\}$ and $\{N\}$ are the classes in this case, which means that the chain is not irreducible. Moreover, $\{0\}$ is a closed class because once you are there, it is impossible to leave it; the same argument can be applied to class $\{N\}$.

(iii) Consider the Markov chain with transition matrix

$$\Pi = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Therefore, the state diagram in this case is



As state 1 leads to states 2 and 3, and then the return to 1 is possible, that means that 1 is an essential state; the same happens with states 2 and 3. In addiction, the states 5 and 6 communicate with one another, so they are also essential. On the other hand, state 4 is inessential because it is possible to leave it, but it is impossible to return to it. Therefore, $\{1, 2, 3\}$, $\{4\}$ and $\{5, 6\}$ are the classes in this case, which implies that the chain is not irreducible. Furthermore, $\{5, 6\}$ is a closed class because once you are in this class, it is impossible to leave it.

3.2 Periodic states. Cyclic classes. First classification of the states of a Markov chain

We illustrate this section with a previous example. Consider a random walk on $\mathbb{Z}$, with $p \in (0, 1)$. We can decompose $\mathbb{Z}$ into two disjoint subsets C_1 and C_2, which are the sets of even and odd numbers, respectively. So, in one step it is necessary to go from C_1 to C_2. Moreover, in two steps we return again to the original set. Having this idea in mind, here we focus on this periodic behavior.

Definition 3.2.1. Let $i \in I$ be an essential state. The period of the state $i \in I$ is defined as: $d_i = gcd\left\{n \geq 1 : p_{i,i}^{(n)} > 0\right\}$. Moreover, if $d_i = 1$, then the state $i \in I$ is called aperiodic.

Observation 3.2.2. Chapman-Kolmogorov equation shows that if $p_{i,i}^{(n)} > 0$, then $p_{i,i}^{(kn)} > 0, \forall k \geq 0$. This is due to: $p_{i,i}^{(kn)} \geq p_{i,i}^{(n)} \overset{k)}{\cdots} p_{i,i}^{(n)} > 0$.

In the following result we show that the period is an equivalence class property, with respect to the relation $i \longrightarrow j$.

Proposition 3.2.3. Let us assume that the two states $i, j \in I$ are communicated. Then, $d_i = d_j$.

Proof. Consider $p_{i,j}^{(n)} > 0$ and $p_{j,i}^{(m)} > 0$, with $n, m \geq 1$. Then, we have that $p_{i,i}^{(n+hk+m)} \geq p_{i,j}^{(n)} \left(p_{j,j}^{(k)}\right)^h p_{j,i}^{(m)}$, $\forall k \geq 1$. This is because it is possible to go from the state i until the same state i in $n + hk + m$ steps by the following way

$$X_0 = i, X_n = j, X_{n+k} = j, \ldots, X_{n+hk} = j, X_{n+hk+m} = i.$$

Therefore, $\forall k \geq 1$ such that $p_{j,j}^{(k)} > 0$, we have that $p_{i,i}^{(n+hk+m)} > 0$, for any $h \geq 1$. So that, $d_i | n + hk + m, \forall h \geq 1$.

We have that $d_i | n + m$ because $p_{i,j}^{(n)} p_{j,i}^{(m)} > 0$. Then, as we have seen previously that $d_i | n + hk + m$, $\forall h \geq 1$, it is necessary that we have that $d_i | hk$, $\forall h \geq 1$. In particular, $d_i | k$, $\forall k \geq 1$ such that $p_{j,j}^{(k)} > 0$ and, consequently, we get that $d_i | d_j$. Using a similar argument we can obtain that $d_j | d_i$. In conclusion, $d_i = d_j$. □

Proposition 3.2.4. Let C be an equivalence class by the relation $\longleftrightarrow$ with period d. If $i, j \in C$, then: $p_{i,j}^{(r)} > 0$, $p_{i,j}^{(s)} > 0$, which implies that $r - s = \dot{d}$.

Proof. We consider $t \geq 1$ and $p_{j,i}^{(t)} > 0$, which exists because $j \longrightarrow i$. Then

- $\exists r > 0$ such that: $p_{i,i}^{(r+t)} \geq p_{i,j}^{(r)} p_{j,i}^{(t)} > 0 \Rightarrow r + t = \dot{d}$.
- $\exists s > 0$ such that: $p_{i,i}^{(s+t)} \geq p_{i,j}^{(s)} p_{j,i}^{(t)} > 0 \Rightarrow s + t = \dot{d}$.

Therefore, $r - s = \dot{d}$.

□

Now, let us assume that $p_{i,j}^{(r)} > 0$, with $r = ad + b$ and $0 \leq b \leq d - 1$. So, if $p_{i,j}^{(s)} > 0$, then $s = cd + b$, for any $c \geq 0$ (we have that $r - s = \dot{d}$ as well). In other words, each pair of states $i, j \in C$ can determine an element $\bar{b} \in \mathbb{Z}/d$ such that all the ways from i to j with positive probability have length b module d.

Thus, if $i \in C$, we can define

$$C_0 = \left\{ j \in C : p_{i,j}^{(n)} > 0 \Rightarrow n \equiv 0 \,(mod\ d) \right\}$$

$$C_1 = \left\{ j \in C : p_{i,j}^{(n)} > 0 \Rightarrow n \equiv 1 \,(mod\ d) \right\}$$

$$\vdots$$

$$C_{d-1} = \left\{ j \in C : p_{i,j}^{(n)} > 0 \Rightarrow n \equiv d - 1 \,(mod\ d) \right\}$$

And we have that $C = \bigcup_{j=0}^{d-1} C_j$.

Definition 3.2.5. The sets $C_0, C_1, \ldots, C_{d-1}$ are called cyclic subclasses.

Proposition 3.2.6. Consider $j \in C_b$, with $p_{j,k} > 0$. Then, $k \in C_{b+1}$, with $C_d = C_0$. It means that, beginning in $i \in C_0$, the process moves from the states of the subclass C_b to the states of the subclass C_{b+1} in one step.

Proof. We consider $m \geq 1$ such that $p_{i,j}^{(n)} > 0$. So, we have that

$$p_{i,k}^{(n+1)} \geq p_{i,j}^{(n)} p_{j,k}^{(1)} > 0.$$

Moreover, as $n \equiv b \,(mod\ d)$, then, $n + 1 \equiv b + 1 \,(mod\ d)$, so that: $k \in C_{b+1}$.

□

Examples:

(i) Random walk on $\mathbb{Z}$.

Assuming that $X_0 = 0$, previously we have seen that this Markov chain is irreducible because it has only one class, which is $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$, with period $d = 2$ because, for example, starting in state 0, the returning to this state is only possible in stages $2, 4, \ldots$. In addiction, the cyclic subclasses are

$$C_0 = \left\{ j \in \mathbb{Z} : p_{0,j}^{(n)} > 0 \Rightarrow n = \dot{2} \right\} = \{0, \pm 2, \pm 4, \ldots\}$$

$$C_1 = \left\{ j \in \mathbb{Z} : p_{0,j}^{(n)} > 0 \Rightarrow n = \dot{2} + 1 \right\} = \{\pm 1, \pm 3, \pm 5, \ldots\}.$$

(ii) Random walk on $\mathbb{Z}$ with absorbing barriers.

At the beginning of this section we have shown that this Markov chain is not irreducible because it has three classes: $\{0\}$, $\{1, \ldots, N-1\}$ and $\{N\}$. On one hand, the absorbing states in this example are 0 and N; therefore, they are aperiodic because, beginning in state 0, in each stage is possible to come back to it; the same argument is valid for state N. On the other hand, $\{1, \ldots, N-1\}$ is a class with period $d = 2$ because, for example, starting in state 1, the returning to this state is only possible in stages $2, 4, \ldots$.

Moreover, the cyclic subclasses are

$$C_0 = \left\{ j \in \{1, \ldots, N-1\} : p_{1,j}^{(n)} > 0 \Rightarrow n = \dot{2} \right\} = \{1, 3, 5, \ldots\}$$

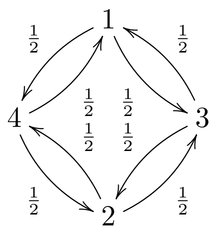
$$C_1 = \left\{ j \in \{1, \ldots, N-1\} : p_{1,j}^{(n)} > 0 \Rightarrow n = \dot{2} + 1 \right\} = \{2, 4, 6, \ldots\}.$$

Observation 3.2.7. Previously we have defined the period of a certain essential state, and using that notion we have also defined the cyclic subclasses. But, in the example above, we study both concepts when the states involved are inessential as well, in order to understand better their definitions, provided that both of them make sense; that is to say that $p_{i,i}^{(n)} > 0$, for $i \in I$.

(iii) Consider the HMC with state space $I = \{1, 2, 3, 4\}$, whose transition matrix is

$$\Pi = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix}$$

Therefore, the state diagram in this case is



This Markov chain is irreducible because it has only one class, which is $\{1, 2, 3, 4\}$. Thus, there is only one equivalence class, whose period is the period of any element. Moreover, all the states have period $d = 2$ because, for instance, starting in state 1, the returning to this state is only possible in stages $2, 4, 6, \ldots$.

Furthermore, the cyclic subclasses are

$$C_0 = \left\{ j \in \{1, 2, 3, 4\} : p_{1,j}^{(n)} > 0 \Rightarrow n = \dot{2} \right\} = \{1, 2\}$$

$$C_1 = \left\{ j \in \{1, 2, 3, 4\} : p_{1,j}^{(n)} > 0 \Rightarrow n = \dot{2} + 1 \right\} = \{3, 4\}.$$

Observation 3.2.8. The description of the cyclic subclasses depends on the state that we take as a reference. For instance, in the previous example, if we take the state $i = 3$ as a reference, instead of the state $i = 1$, we obtain that the cyclic subclasses are

$$C_0 = \left\{ j \in \{1, 2, 3, 4\} : p_{3,j}^{(n)} > 0 \Rightarrow n = \dot{2} \right\} = \{3, 4\}$$

$$C_1 = \left\{ j \in \{1, 2, 3, 4\} : p_{3,j}^{(n)} > 0 \Rightarrow n = \dot{2} + 1 \right\} = \{1, 2\}.$$

3.3 Hitting times and absorption probabilities

In this section we study the moment when the Markov chain enters a subset of the state space. In addiction, we analyze the time required to arrive to this subset.

Definition 3.3.1. Let $\{X_n : n \geq 0\}$ be a HMC with a finite number of states I and with transition matrix Π. The hitting time of a subset A of I is the random variable $H_A : \Omega \longrightarrow \mathbb{N} \cup \{\infty\}$ given by

$$H_A(\omega) = \inf \{n \geq 0 : X_n(\omega) \in A\},$$

where we assent that the infimum of the empty set is ∞. Furthermore, the probability starting from the state i that X_n gets into A is then

$$\lambda_i^A = P(H_A < \infty | X_0 = i).$$

When A is a closed set, λ_i^A is called absorption probability.

Observation 3.3.2.

- If A is a closed set and $i \in A$, then $\lambda_i^A = 1$.
- If A and B are closed sets, $A \cap B = \emptyset$ and $i \in B$, then $\lambda_i^A = 0$.

Definition 3.3.3. The mean time that $\{X_n : n \geq 0\}$ needs to reach A is called mean absorption time and it is given by

$$m_i^A = E_i(H_A) = \sum_{n=0}^{\infty} nP(H_A = n | X_0 = i) + \infty P(H_A = \infty | X_0 = i).$$

Observation 3.3.4. In the previous definition, E_i means that $p_i = P(\cdot | X_0 = i)$.

From now on, we study how to compute the value of the absorption probabilities and the mean absorption times in different cases.

Theorem 3.3.5. Let A be a closed class and $\left\{\lambda_i^A : i \in I\right\}$ the absorption probabilities. Then, $\left\{\lambda_i^A : i \in I\right\}$ are the solution to the system of linear equations

$$\left.\begin{array}{ll} \lambda_i^A = 1, & \text{if } i \in A, \\ \lambda_i^A = \sum\limits_{j \in I} p_{i,j} \lambda_j^A, & \text{if } i \notin A. \end{array}\right\}$$

Proof. On one hand, if $X_0 = i \in A$, then $H_A = 0$, so that $\lambda_i^A = 1$.

On the other hand, if $X_0 = i \notin A$, then $H_A \geq 1$. Using the Markov property, we have

$$P(H_A < \infty | X_0 = i, X_1 = j) = P(H_A < \infty | X_1 = j) = \lambda_j^A.$$

Therefore

$$\begin{aligned} \lambda_i^A &= P(H_A < \infty | X_0 = i) = \sum_{j \in I} \frac{P(H_A < \infty, X_0 = i, X_1 = j)}{P(X_0 = i)} \\ &= \sum_{j \in I} P(H_A < \infty | X_0 = i, X_1 = j) \frac{P(X_0 = i, X_1 = j)}{P(X_0 = i)} = \sum_{j \in I} p_{i,j} \lambda_j^A. \end{aligned}$$

□

Observation 3.3.6. We have already proved that the absorption probabilities are the solution to the previous system of linear equations, but this solution can not be unique. What we have shown is that it is the minimal non-negative solution; it means that if $\{X_i : i \in I\}$ is another solution with $X_i \geq 0$, $\forall i$, then $X_i \geq \lambda_i^A$, $\forall i$.

Observation 3.3.7. For the states $i \in I$, we can write the absorption probabilities as

$$\lambda_i^A = \sum_{j \in A} p_{i,j} + \sum_{j \notin A} p_{i,j} \lambda_j^A,$$

where $\sum\limits_{j \in A} p_{i,j}$ is the probability to reach A starting in the state i in one step.

Example:

In this example we work on a similar version of 'The Gambler's Ruin Problem', which we have already studied in the last chapter.

Consider the Markov chain, with $q = 1 - p \in (0, 1)$, with transition matrix

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

Therefore, the state diagram in this case is

$$\begin{tikzcd} 0 \arrow[loop left, "1"] & 1 \arrow[l, "q"'] \arrow[r, bend left, "p"] & 2 \arrow[l, bend left, "q"] \arrow[r, bend left, "p"] & 3 \arrow[l, bend left, "q"] & \cdots \end{tikzcd}$$

Suppose that you go to a casino with i euros and you gamble 1 euro in each play; so that, you can win 1 euro with probability p and you can lose 1 euro with probability $q = 1 - p$.

If we want to study the probability of getting ruined, we are talking about the absorption probabilities of the state 0, because $\{0\}$ is the only absorption class that has the chain. So that, we write $\lambda_i = \lambda_i^{\{0\}}$, for $i = 0, 1, \ldots$. Then, using the previous theorem, for $i \geq 1$ we have

$$\lambda_i = \begin{cases} 1, & \text{if } i = 0, \\ p\lambda_{i+1} + q\lambda_{i-1}, & \text{if } i \geq 1. \end{cases}$$

Solving the previous recurrence, we get

- If $p \neq q$, then $\lambda_i = A + B \left(\frac{q}{p}\right)^i$, with $A + B = 1$.

  On one hand, if $p < q$, then $\lambda_i = 1$, $\forall i \geq 0$. On the other hand, if $p > q$, then $\lambda_i = \left(\frac{q}{p}\right)^i$, $\forall i \geq 0$.

- If $p = q$, then $\lambda_i = 1$, $\forall i \geq 0$.

Theorem 3.3.8. Let A be a closed class and $\left\{m_i^A : i \in I\right\}$ the mean absorption times. Then, $\left\{m_i^A : i \in I\right\}$ are the solution to the system of linear equations

$$\left.\begin{array}{ll} m_i^A = 0, & \text{if } i \in A, \\ m_i^A = 1 + \sum\limits_{j \notin A} p_{i,j} m_j^A, & \text{if } i \notin A. \end{array}\right\}$$

Proof. On one hand, if $X_0 = i \in A$, then $H_A = 0$; so that, $m_i^A = 0$.

On the other hand, if $X_0 = i \notin A$, then $H_A \geq 1$. Using the Markov property, for $n \geq 1$ we have

$$\begin{aligned} P\left(H_A = n, X_1 = j | X_0 = i\right) &= \frac{P\left(H_A = n, X_1 = j, X_0 = i\right)}{P\left(X_0 = i\right)} \\ &= P\left(H_A = n | X_0 = i, X_1 = j\right) p_{i,j} \\ &= P\left(H_A = n | X_1 = j\right) p_{i,j}. \end{aligned}$$

The case where $P\left(H_A = \infty \mid X_0 = i\right) > 0$ is immediate because we get that $m_i^A = \infty$, if $i \notin A$, and then the equality $m_i^A = 1 + \sum\limits_{j \notin A} p_{i,j} m_j^A$, if $i \notin A$, is satisfied. Therefore, we focus on the case where $P\left(H_A = \infty | X_0 = i\right) = 0$, and we have

$$\begin{aligned} m_i^A &= E_i\left(H_A\right) = \sum_{n=1}^{\infty} n P\left(H_A = n | X_0 = i\right) + \infty P\left(H_A = \infty | X_0 = i\right) \\ &= \sum_{n=1}^{\infty} \left[n P\left(H_A = n | X_0 = i\right) - P\left(H_A = n | X_0 = i\right) + P\left(H_A = n | X_0 = i\right)\right] \\ &= 1 + \sum_{n=1}^{\infty} (n-1) P\left(H_A = n | X_0 = i\right) \\ &= 1 + \sum_{j \in I} \sum_{n=1}^{\infty} (n-1) P\left(H_A = n | X_1 = j\right) p_{i,j} = 1 + \sum_{j \in I} m_j^A p_{i,j} \\ &= 1 + \sum_{j \notin A} p_{i,j} m_j^A. \end{aligned}$$

□

Observation 3.3.9. We have already proved that the mean absorption times are the solution to the previous system of linear equations, but this solution can not be unique. What we have shown is that it is the minimal non-negative solution; it means that if $\{X_i : i \in I\}$ is another solution with $X_i \geq 0$, $\forall i$, then $X_i \geq m_i^A$, $\forall i$.

Example:

Consider the chain with transition matrix

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Therefore, the state diagram in this case is

$$\begin{tikzcd} 1 \arrow[loop left, "1"] & 2 \arrow[l, "1/2"'] \arrow[r, bend left, "1/2"] & 3 \arrow[l, bend left, "1/2"] \arrow[r, "1/2"'] & 4 \arrow[loop right, "1"] \end{tikzcd}$$

- Starting from state 2, what is the probability of absorption in state 4?

  We compute all the absorption probabilities in state 4:

  We write $\lambda_i = \lambda_i^{\{4\}} = P\left(X_n \text{ finishes in } 4 \mid X_0 = i\right)$, for $i = 1, 2, 3, 4$. Then

$$\left.\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &= \tfrac{1}{2}\lambda_1 + \tfrac{1}{2}\lambda_3 \\ \lambda_3 &= \tfrac{1}{2}\lambda_2 + \tfrac{1}{2}\lambda_4 \\ \lambda_4 &= 1 \end{aligned}\right\} \longrightarrow \left.\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &= \tfrac{1}{3} \\ \lambda_3 &= \tfrac{2}{3} \\ \lambda_4 &= 1 \end{aligned}\right\}$$

  Therefore, starting from state 2, the probability of absorption in state 4 is $\lambda_2 = \frac{1}{3}$.

- Starting from state 2, how many stages are necessary until the chain is absorbed in state 1 or in state 4?

  We compute all the mean absorption times in states 1 or 4:

  We write $m_i = m_i^{\{1,4\}}$, for $i = 1, 2, 3, 4$. Then

$$\left.\begin{aligned} m_1 &= 0 \\ m_2 &= 1 + \tfrac{1}{2}m_1 + \tfrac{1}{2}m_3 \\ m_3 &= 1 + \tfrac{1}{2}m_2 + \tfrac{1}{2}m_4 \\ m_4 &= 0 \end{aligned}\right\} \longrightarrow \left.\begin{aligned} m_1 &= 0 \\ m_2 &= 2 \\ m_3 &= 2 \\ m_4 &= 0 \end{aligned}\right\}$$

  Therefore, starting from state 2, until the chain is absorbed in state 1 or in state 4 are necessary $m_2 = 2$ stages.

4 Recurrent and transient states

In this chapter, initially we study the concepts of recurrent and transient states. Afterward, we work on the notion of the first instant that the chain is located in some state and also on the time of the rth step. Subsequently, the number of times that the chain is located in a certain state is an important idea as well; in fact, this concept and the stopping times help us to define what is a return probability. After that, we show rules and criteria to handle with this notion and, finally, we provide some examples for a complete understanding of them.

Throughout all this chapter we consider a probability space $(\Omega, \mathcal{A}, P)$ and a collection of discrete random variables $X_n : \Omega \to I$, where I is a countable set and $\{X_n : n \geq 0\}$ is a $HMC\,(\gamma, \Pi)$ with $\Pi = (p_{i,j}; i, j \in I)$.

4.1 Definitions

In this section, we study whether, starting in some state, it is possible to return to it or not. In addiction, we work on the time that the chain is located in a certain state.

Definition 4.1.1. The state $i \in I$ is recurrent or persistent if

$$P_i\left(X_n = i \text{ for infinite } n\right) = 1.$$

The state $i \in I$ is transient if

$$P_i\left(X_n = i \text{ for infinite } n\right) = 0.$$

That is to say that you always come back to a recurrent state, but you can never return to a transient state.

Observation 4.1.2. An absorbing state is recurrent.

Definition 4.1.3. The first instant that the chain is located in state $i \in I$ is defined as

$$T_i\left(\omega\right) = \inf\left\{X_n\left(\omega\right) = i : n \geq 1\right\}, \text{ for } \omega \in \Omega,$$

where we assent that the infimum of the empty set is ∞. Moreover, the rth instant that the chain is located in the state $i \in I$ can be defined by using the following recurrence

$$\begin{cases} T_i^0\left(\omega\right) = 0, \\ T_i^1\left(\omega\right) = T_i\left(\omega\right), \\ T_i^{(r+1)}\left(\omega\right) = \inf\left\{X_n\left(\omega\right) = i : n \geq T_i^{(r)}\left(\omega\right) + 1\right\}. \end{cases}$$

Furthermore, we can define the time of the rth excursion as

$$S_i^{(r)} = \begin{cases} T_i^{(r)} - T_i^{(r-1)} = \inf\left\{X_{T_i^{(r-1)}+n} = i : n \geq 1\right\}, & \text{if } T_i^{(r-1)} < \infty, \\ 0, & \text{otherwise.} \end{cases}$$

Taking into account the previous concepts and the definition of stopping times made in **chapter 2**, we get the following result.

Observation 4.1.4. For $k \geq 0$, the random variables $T_i^{(k)}$ and $S_i^{(k)}$ are stopping times because $\left\{T_i^{(k)} = n\right\}$ means that $X_n(\omega) = i$ and, before this, there are $r-1$ i's. So that, there is only dependence on $X_0, \ldots, X_n$.

Now, we show an interesting lemma related to the rth instant that the chain is located in a certain state and to the the time of the rth excursion as well. In its proof, we use stopping times and the strong Markov property, both ideas that have been seen previously.

Lemma 4.1.5. *$S_i^{(r)}$ is independent of $\left\{X_m : m \leq T_i^{(r-1)}\right\}$ conditional on $T_i^{(r-1)} < \infty$, for $r = 2, 3, \ldots$. Furthermore, $P\left(S_i^{(r)} = n \mid T_i^{(r-1)} < \infty\right) = P_i(T_i = n)$.*

Proof. Applying the strong Markov property to the stopping time $T = T_i^{(r-1)}$ and assuming that if $T < \infty$ then $X_T = i$, we have that $\{X_{T+n} : n \geq 0\}$ is a $HMC(\delta_i, \Pi)$ conditional on $T_i^{(r-1)} < \infty$, and it is also independent of $X_0, \ldots, X_T$. Moreover

$$S_i^{(r)} = \inf\{X_{T+n} = i : n \geq 1\}$$

is the first time that the chain $\{X_{T+n} : n \geq 0\}$ is located in the state i; so that

$$P\left(S_i^{(r)} = n \mid T_i^{(r-1)} < \infty\right) = P(T_i = n \mid X_0 = i) = P_i(T_i = n).$$

□

Definition 4.1.6. The number of times that the chain is located in the state i is defined as

$$N_i(\omega) = \#\{X_n(\omega) = i : n \geq 1\} = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = i\}}.$$

Observation 4.1.7. Considering the last definition, we have

- $$E_j(N_i) = E_j\left(\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=i\}}\right) = \sum_{n=1}^{\infty} E_j\left(\mathbb{1}_{\{X_n=i\}}\right) = \sum_{n=1}^{\infty} P(X_n = i \mid X_0 = j)$$
  $$= \sum_{n=1}^{\infty} p_{j,i}^{(n)}.$$
- In particular, if $i = j$, we get that $E_i(N_i) = \sum_{n=1}^{\infty} p_{i,i}^{(n)}$.

Now, we study the probability that the chain returns to a certain state, a concept related to the first instant that the chain is located in the state considered.

Definition 4.1.8. The return probability is defined as

$$\rho_{i,j} = P(T_j < \infty \mid X_0 = i),$$

$$\rho_{i,i} = P(T_i < \infty \mid X_0 = i).$$

We finish this section with an important result that involves the previous definition and the notion of number of times that the chain is located in a certain state.

Lemma 4.1.9. For $r \geq 1$, we have that

$$P\left(N_j \geq r \mid X_0 = i\right) = \rho_{i,j}\rho_{j,j}^{r-1}.$$

Proof. Note that if $X_0 = i$, then $\{N_j \geq r\} = \left\{T_j^{(r)} < \infty\right\}$.

We prove it by induction. Firstly, we study the case $r = 1$:

$$P\left(N_j \geq 1 \mid X_0 = i\right) = P\left(T_j < \infty \mid X_0 = i\right) = \rho_{i,j}.$$

Now we need to prove that if the statement holds for r, then it also holds for $r+1$. Using the induction hypothesis, which is

$$P\left(N_j \geq r \mid X_0 = i\right) = \rho_{i,j}\rho_{j,j}^{r-1},$$

as well as the **lemma 4.1.5.**, we obtain

$$\begin{aligned}
&P\left(N_j \geq r+1 \mid X_0 = i\right) = P\left(T_j^{(r+1)} < \infty \mid X_0 = i\right) \\
&= P\left(T_j^{(r)} < \infty, S_j^{(r+1)} < \infty \mid X_0 = i\right) \\
&= \frac{P\left(T_j^{(r)} < \infty, S_j^{(r+1)} < \infty, X_0 = i\right)}{P\left(X_0 = i\right)} \\
&= P\left(S_j^{(r+1)} < \infty \mid T_j^{(r)} < \infty, X_0 = i\right) P\left(T_j^{(r)} < \infty \mid X_0 = i\right) \\
&= \left[\sum_{n=0}^{\infty} P\left(S_j^{(r+1)} = n \mid T_j^{(r)} < \infty\right)\right] P\left(N_j \geq r \mid X_0 = i\right) \\
&= \left[\sum_{n=0}^{\infty} P_j\left(T_j = n\right)\right] \rho_{i,j}\rho_{j,j}^{r-1} \\
&= \left[\sum_{n=0}^{\infty} P\left(T_j = n \mid X_0 = j\right)\right] \rho_{i,j}\rho_{j,j}^{r-1} \\
&= P\left(T_j < \infty \mid X_0 = j\right) \rho_{i,j}\rho_{j,j}^{r-1} = \rho_{i,j}\rho_{j,j}^{r}.
\end{aligned}$$

□

Observation 4.1.10. In particular, we have

- $P\left(N_i \geq 0 \mid X_0 = i\right) = 1$.
- $P\left(N_i \geq r \mid X_0 = i\right) = \rho_{i,i}^{r}$, for $r \geq 1$.

4.2 Rules

In this section we study criteria and rules related to recurrent states and transient states. Furthermore, we work on the decomposition of a chain by using the first time that it is located in a certain state.

Theorem 4.2.1.

(a) If $\rho_{i,i} = P(T_i < \infty \mid X_0 = i) = 1$, then the state i is recurrent. In addiction, $\sum_{n=0}^{\infty} p_{i,i}^{(n)} = \infty$ and $E_i(N_i) = \infty$ as well.

(b) If $\rho_{i,i} = P(T_i < \infty \mid X_0 = i) < 1$, then the state i is transient. Moreover, $\sum_{n=0}^{\infty} p_{i,i}^{(n)} < \infty$ and $E_i(N_i) = \dfrac{\rho_{i,i}}{1-\rho_{i,i}}$.

Proof. On one hand, we suppose that $\rho_{i,i} = 1$. Using the property of probability **8.(b)**, we get

$$\begin{aligned} P(N_i = \infty \mid X_0 = i) &= \lim_{r\to\infty} P(N_i > r \mid X_0 = i) = \lim_{r\to\infty} P(N_i \geq r+1 \mid X_0 = i) \\ &= \lim_{r\to\infty} \rho_{i,i}^{r+1} = \lim_{r\to\infty} 1 = 1. \end{aligned}$$

Therefore, i is recurrent. Furthermore, we have

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} = E_i(N_i) = \infty.$$

On the other hand, consider $\rho_{i,i} < 1$. Using that $E(X) = \sum_{l=1}^{\infty} P(X \geq l)$, we obtain

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} = E_i(N_i) = \sum_{l=1}^{\infty} P(N_i \geq l \mid X_0 = i) = \sum_{l=1}^{\infty} \rho_{i,i}^{l} = \sum_{l=0}^{\infty} \rho_{i,i}^{l+1} = \frac{\rho_{i,i}}{1-\rho_{i,i}} < \infty.$$

In consequence, $P(N_i = \infty \mid X_0 = i) = 0$ and i is transient.

$\square$

Observation 4.2.2. Using the definitions of recurrent state and transient state, we note that the implication to the left in the previous theorem is also satisfied; that is to say that

- $\rho_{i,i} = P(T_i < \infty \mid X_0 = i) = 1$ if, and only if, the state i is recurrent. In addiction, $\sum_{n=0}^{\infty} p_{i,i}^{(n)} = \infty$ and $E_i(N_i) = \infty$ as well.
- $\rho_{i,i} = P(T_i < \infty \mid X_0 = i) < 1$ if, and only if, the state i is transient. Moreover, $\sum_{n=0}^{\infty} p_{i,i}^{(n)} < \infty$ and $E_i(N_i) = \dfrac{\rho_{i,i}}{1-\rho_{i,i}}$.

The next result show that being recurrent or transient is a class property.

Theorem 4.2.3. Let C be a class. Then, all the states that are in C are recurrent or transient.

Proof. Consider $i, j \in C$ and suppose that i is transient; as the states i and j belong to the same class, then there exist $n, m \geq 0$ such that $\rho_{i,j}^{(n)} > 0$ and $\rho_{j,i}^{(m)} > 0$. Therefore, for all $r \geq 0$, we have

$$\rho_{i,i}^{(n+r+m)} \geq \rho_{i,j}^{(n)} \rho_{j,j}^{(r)} \rho_{j,i}^{(m)} \Rightarrow \rho_{j,j}^{(r)} \leq \frac{\rho_{i,i}^{(n+r+m)}}{\rho_{i,j}^{(n)} \rho_{j,i}^{(m)}} \Rightarrow \sum_{r=0}^{\infty} \rho_{j,j}^{(r)} \leq \frac{1}{\rho_{i,j}^{(n)} \rho_{j,i}^{(m)}} \sum_{r=0}^{\infty} \rho_{i,i}^{(n+r+m)} < \infty.$$

Thus, j is transient.

Similarly, considering $i, j \in C$ and supposing that i is recurrent, we can obtain

$$\sum_{r=0}^{\infty} \rho_{j,j}^{(r)} = \infty.$$

So that, j is recurrent. □

Theorem 4.2.4. If C is a recurrent class, then C is a closed class.

Proof. Assume that C is not a closed class. In other words, there exist $i \in C$, $j \notin C$ and $m \geq 1$ such that $\rho_{i,j}^{(m)} > 0$.

We have as well that $P_i\left(\{X_m = j\} \cap \{X_n = i \text{ for infinite } n\}\right) = 0$; it means that if you leave the class C, then it is impossible to return to it. Therefore

$$P_i\left(X_n = i \text{ for infinite } n\right) < 1.$$

So that, i is not recurrent, that is to say that C is not recurrent, which stands in contradiction to the previous assumption. □

Theorem 4.2.5. If C is a finite closed class, then C is recurrent.

Proof. Suppose that C is a finite closed class. Consider that the Markov chain $\{X_n : n \geq 0\}$ starts in C, for instance, in state $j \in C$. Then, there exists $i \in C$ such that

$$P_j\left(X_n = i \text{ for infinite } n\right) > 0.$$

On the other hand, the Markov property implies that

$$P_j\left(X_n = i \text{ for infinite } n\right) = P_j\left(X_n = i \text{ for some } n\right) P_i\left(X_n = i \text{ for infinite } n\right).$$

As $P_i\left(X_n = i \text{ for infinite } n\right) \neq 0$, then i is not transient; that is to say that i is recurrent. Thus, C is recurrent. □

Observation 4.2.6.

- All finite Markov chains have at least one recurrent state.
- In all finite Markov chains, starting in a state i it is possible, at least, to go to a recurrent state j.
- If there exist two states $i, j \in I$ such that $j \longrightarrow i$, but not $i \longrightarrow j$, then j is transient.
- If I is finite, then j is transient if, and only if, there exists a state i such that $j \longrightarrow i$ but not $i \longrightarrow j$.

4.3 Examples

In this section we show some interesting examples in order to understand better the previous concepts of this chapter.

(i) Random walk on $\mathbb{Z}$.

We have already shown that there is only one class in this case, but we do not know if this chain is make up of recurrent states or transient states. We had

$$\begin{cases} S_0 = X_0 = 0, \\ S_n = \sum\limits_{i=1}^{n} X_i : \Omega \longrightarrow \{2k - n : k = 0, 1, \ldots, n\}, \text{ for } n \geq 1. \end{cases}$$

So that, we get

$$P(S_n = 2k - n) = \binom{n}{k} p^k (1-p)^{n-k}.$$

If we define $h = 2k - n$, then we obtain

$$P(S_n = h) = \binom{n}{\frac{n+h}{2}} p^{\frac{n+h}{2}} (1-p)^{\frac{n-h}{2}}.$$

Now, we study the summability of the series $p_{0,0}^{(n)}$:

$$\begin{cases} p_{0,0}^{(n)} = 0, & \text{if n is odd,} \\ p_{0,0}^{(n)} = P(S_n = 0) = \binom{n}{\frac{n}{2}} p^{\frac{n}{2}} (1-p)^{\frac{n}{2}}, & \text{if n is even.} \end{cases}$$

Therefore, considering only the even numbers, we have

$$\sum_{n=1}^{\infty} p_{0,0}^{(n)} = \sum_{n=1}^{\infty} \binom{2n}{n} p^n (1-p)^n.$$

Using the Stirling formula, which is $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ as n approaches infinity, and writing $q = 1 - p$, we obtain

$$\binom{2n}{n} p^n (1-p)^n = \binom{2n}{n} p^n q^n \approx \frac{(4pq)^n}{\sqrt{\pi n}}.$$

In particular, we study the case where $p = q = \frac{1}{2}$:

$$p_{0,0}^{(n)} = \binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \approx \frac{1}{\sqrt{\pi n}}.$$

So that, we get

$$\sum_{n=1}^{\infty} p_{0,0}^{(n)} \approx \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}} = +\infty,$$

this is because we have obtained a p-series with $p = \frac{1}{2} < 1$.

Therefore, using **Observation 4.2.2.**, the state 0 is recurrent. Thus, all the states are recurrent since there is only one class; in other words, the chain is recurrent.

(ii) Random walk on $\mathbb{Z}^2$.

In this case the state space is $I = \{(i,j) \in \mathbb{Z}^2\}$ and the transition probabilities are

$$p_{(i,j),(i',j')} = \begin{cases} p_1, & \text{if } i' = i+1, j' = j, \\ p_2, & \text{if } i' = i-1, j' = j, \\ p_3, & \text{if } i' = i, j' = j+1, \\ p_4, & \text{if } i' = i, j' = j-1, \end{cases} \qquad \text{with } p_1 + p_2 + p_3 + p_4 = 1.$$

From now on, we consider the particular case where $p_i = \frac{1}{4}$, for $i = 1, 2, 3, 4$; and we compute $p^{(2n)}_{(i,j),(i,j)}$.

Starting in state (i,j), to return to this state (i,j) in $2n$ stages it is necessary that, for some $k \in \mathbb{N}$, we have made k movements to the east, k to the west, $n-k$ to the north and $n-k$ to the south. So that, we express these $2n$ repetitions of an experience with 4 possible outcomes, which are east, west, north and south, with probability $\frac{1}{4}$, by using the multinomial distribution. Therefore, we get

$$p^{(2n)}_{(i,j),(i,j)} = \frac{1}{4^{2n}} \sum_{k=0}^{n} \frac{(2n)!}{k!k!\,(n-k)!\,(n-k)!} = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^2.$$

We observe that $p^{(2n)}_{(i,j),(i,j)} = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^2 = \left(p^{(n)}_{0,0}\right)^2$, where $p^{(n)}_{0,0}$ is the return probability to state 0 in n stages in a random walk on $\mathbb{Z}$.

Using again the Stirling formula, we get

$$p^{(2n)}_{(i,j),(i,j)} = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^2 \approx \frac{1}{\pi n}.$$

So that, we have

$$\sum_{n=1}^{\infty} p^{(2n)}_{(i,j),(i,j)} \approx \sum_{n=1}^{\infty} \frac{1}{\pi n} = +\infty,$$

this is because we have obtained an harmonic series.

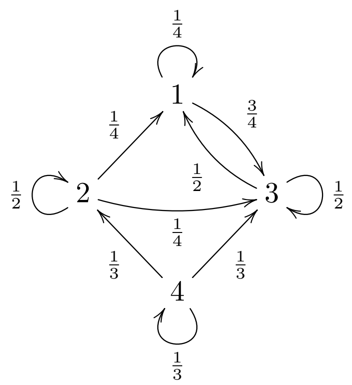
Therefore, using **Observation 4.2.2.**, the state $(0,0)$ is recurrent. In consequence, all the states are recurrent as there is only one class; that is to say that the chain is recurrent.

Observation 4.3.1. The case where there is not symmetry, that is $p \neq q$, has more difficulties and the result is that the chain is transient.

(iii) Consider the Markov chain with transition matrix

$$\Pi = \begin{pmatrix} \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Therefore, the state diagram in this case is



In this example there are three equivalent classes, which are

$C_1 = \{1, 3\}$, $C_2 = \{2\}$ and $C_3 = \{4\}$.

The state 2 is transient because, starting in this state, you can go to state 1 but, once you are there, you can never return to state 2; therefore, C_2 is a transient class. Similarly, the state 4 is transient because you can leave this state to go to state 3 but, then, you can not come back to state 4; thus, C_3 is a transient class. Finally, C_1 is a recurrent class because it is a closed class.

Now we study, starting in state 2, what is the probability of return to this state in 12 stages. If we begin in state 2, we can only move to states $1, 2$ or 3. If we go to states 1 or 3, it means that we are going to the closed class C_1; so that, it will not be possible to return to state 2. In consequence, the only possibility left is going repeatedly to state 2 during the 12 stages. Therefore, we obtain

$$p_{2,2}^{(12)} = (p_{2,2})^{12} = \frac{1}{2^{12}}.$$

5 Invariant distributions

In this chapter, firstly we introduce the concept of invariant distributions, also known as stationary distributions. After that, we study the existence of this type of distributions; in particular, we show the relation between this idea and the limit of $p_{i,j}^{(n)}$ as n approaches infinity. Subsequently, we work on the uniqueness of invariant distributions; specifically, we see that this notion is related to the concept of ergodicity. Finally, we study the concept of regularity and we give some examples to complete the explanation of the previous concepts.

Throughout all this chapter we consider a probability space $(\Omega, \mathcal{A}, P)$ and a collection of discrete random variables $X_n : \Omega \to I$, where I is a countable set and $\{X_n : n \geq 0\}$ is a $HMC\,(\gamma, \Pi)$ with $\Pi = (p_{i,j} : i, j \in I)$.

5.1 Invariant distributions or stationary distributions

The aim of this section is to study the long-time properties that has a HMC; in other words, we work on the value of $\lim\limits_{n\to\infty} p_{i,j}^{(n)}$. Moreover, we show that this concept is related to what we call invariant distributions.

Definition 5.1.1. Let Π be a transition matrix. A probability distribution μ over I is called invariant distribution or stationary distribution for Π if

$$\mu^t = \mu^t \Pi,$$

where $\mu_j = \sum\limits_{i\in I} \mu_i p_{i,j}$, $\forall j \in I$, and $\mu^t = \big(\mu_1, \mu_2, \ldots, \mu_{card(I)}\big)$.

In addiction, the following expression is also satisfied

$$\sum_{j\in I} \mu_j = 1,$$

because μ is a probability distribution over I.

Observation 5.1.2. The previous definition shows that $\mu^t = \mu^t \Pi^n$, $\forall n \geq 1$; so that, $\{X_n : n \geq 0\}$ is a $HMC\,(\mu, \Pi)$, whose law is

$$\mu^{(n)}\,(k) = \mu_k^{(n)} = P\,(X_n = k) = \sum_{i\in I} \mu_i p_{i,k}^{(n)} = \mu_k, \text{ for } k \in I \text{ and } n \geq 1.$$

Therefore, all the random variables X_n have the same law, $\forall n \geq 1$.

Taking the previous observation into consideration, we get the following result.

Theorem 5.1.3. Assume that $\{X_n : n \geq 0\}$ is a $HMC\,(\mu, \Pi)$ and let μ be an invariant distribution for Π. Then, $\{X_{n+m} : n \geq 0\}$ is also a $HMC\,(\mu, \Pi)$.

Proof. The preceding observation asserts that $P\,(X_m = k) = \mu_k$, $\forall k \in I$. As we have as well that X_{m+n+1} is independent of $X_m, X_{m+1}, \ldots, X_{m+n}$ conditional on $X_{m+n} = k$ and has distribution $(p_{i,j} : j \in I)$, we obtain the required result. $\square$

From now on, we suppose that the state space I is finite. Next we study the existence of stationary distributions.

Theorem 5.1.4. Consider a HMC with finite state space I and transition matrix Π. Then, Π has, at least, one invariant distribution.

Proof. Let v be a probability over I; it means that $v = \left(v_1, \dots, v_{card(I)}\right)$, with $v_i \in [0,1]$ and $\sum_{i \in I} v_i = 1$.

Now, $\forall n \geq 1$, we define $v_n^t = \frac{1}{n} \sum_{k=0}^{n-1} v^t \Pi^k$.

Then, we get that v_n^t establishes a probability over I because we have that

- $v_n^i \geq 0$, $\forall i \in I$.
- $\sum_{i \in I} v_n^i = \frac{1}{n} \sum_{i \in I} \sum_{k=0}^{n-1} \sum_{h \in I} v_h p_{h,i}^{(k)} = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{h \in I} v_h \left[\sum_{i \in I} p_{h,i}^{(k)}\right] = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{h \in I} v_h = \sum_{h \in I} v_h = 1.$

Furthermore, the set of probabilities over I is a closed and bounded set of $[0,1]^{card(I)}$; so that, there exists a convergent subsequence, whose limit is an element included in that set. That is to say that there exists a probability μ over I such that, for some subsequence $\{v_{n_k} : k \geq 1\}$, it satisfies that the limit of v_{n_k} as k approaches infinity equals μ.

Then, we have that

$$v_{n_k}^t - v_{n_k}^t \Pi = \frac{1}{n_k} \left[\sum_{l=0}^{n_k - 1} v^t \Pi^l - \sum_{l=0}^{n_k - 1} v^t \Pi^{l+1}\right] = \frac{1}{n_k} \left(v^t - v^t \Pi^{n_k}\right).$$

Finally, considering the limit of the previous result as k approaches infinity, we obtain

$$\mu^t - \mu^t \Pi = \lim_{k \to \infty} \frac{1}{n_k} \left(v^t - v^t \Pi^{n_k}\right) = 0,$$

because $\left\{v^t - v^t \Pi^{n_k} : k \geq 1\right\}$ is a bounded sequence. In consequence, μ is an invariant distribution for Π. $\square$

Observation 5.1.5. The system of linear equations $\mu^t = \mu^t \Pi$ has solution. To see this fact, we verify that $\mu^t (I - \Pi) = 0$, which is also a system of linear equations with determinant

$$\begin{vmatrix} 1 - p_{1,1} & -p_{1,2} & \dots & -p_{1,|I|} \\ -p_{2,1} & 1 - p_{2,2} & \dots & -p_{2,|I|} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{|I|,1} & -p_{|I|,2} & \dots & 1 - p_{|I|,|I|} \end{vmatrix}$$

where $|I|$ denotes the cardinality of the state space I.

The sum of any row of the previous determinant equals $1 - \sum_{j \in I} p_{i,j}$, $\forall i \in I$. So that, the system has solution, although we can not assert that the possible solution is a probability. In addiction, in general there is not uniqueness.

Examples:

(i) If μ_1 and μ_2 are two invariant distributions for Π; then, any convex linear combination of both distributions, that is to say that it can be expressed as $\lambda\mu_1 + (1-\lambda)\,\mu_2$, with $\lambda \in [0,1]$, is a stationary distribution as well.

(ii) If $\Pi = I$; then, all the distributions are invariant.

(iii) Random walk on $\mathbb{Z}$ with absorbing barriers.

We have already seen that in this case the state space is $I = \{0,1,\ldots,N-1,N\}$ and the transition matrix is

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

Then, $\mu^t = \mu^t \Pi$, if, and only if,

$$\left.\begin{array}{lll} \mu_0 & = \mu_0 + \mu_1 q, & \\ \mu_1 & = \mu_2 q, & \\ \vdots & & \\ \mu_j & = \mu_{j-1}p + \mu_{j+1}q, & \text{for } j = 2,\ldots,N-2, \\ \vdots & & \\ \mu_{N-1} & = \mu_{N-2}p, & \\ \mu_N & = \mu_{N-1}p + \mu_N. & \end{array}\right\}$$

Using as well the condition that $\mu_0 + \mu_1 + \ldots + \mu_{N-1} + \mu_N = 1$, we have that the solution of the previous system of linear equations is

$\mu_1 = \mu_2 = \ldots = \mu_{N-2} = \mu_{N-1} = 0$, and μ_0 and μ_N such that verify that $\mu_0 + \mu_N = 1$, with $\mu_0, \mu_N \geq 0$.

Therefore, we have obtained a collection of invariant distributions:

$$\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}, \mu_N) = (\lambda, 0, \ldots, 0, 1-\lambda)\,, \text{ with } \lambda \in [0,1]\,.$$

Now, we study the value of the limit of $p_{i,j}^{(n)}$ as n approaches infinity.

Theorem 5.1.6. Suppose that the state space I is finite and for some $i \in I$ it is satisfied that

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \Pi_j, \forall j \in I.$$

Then, $\Pi = \{\Pi_j : j \in I\}$ is an invariant distribution.

Proof. Using that the summation and the limit commute due to the hypothesis that the state space I is finite, we have that

$$\sum_{j\in I} \Pi_j = \sum_{j\in I} \lim_{n\to\infty} p_{i,j}^{(n)} = \lim_{n\to\infty} \sum_{j\in I} p_{i,j}^{(n)} = 1.$$

Therefore, we obtain

$$\Pi_j = \lim_{n\to\infty} p_{i,j}^{(n)} = \lim_{n\to\infty} \sum_{k\in I} p_{i,k}^{(n-1)} p_{k,j} = \sum_{k\in I} \lim_{n\to\infty} p_{i,k}^{(n-1)} p_{k,j} = \sum_{k\in I} \Pi_k p_{k,j}.$$

□

Observation 5.1.7. In the example of random walk on $\mathbb{Z}$, if we consider the probability of, starting in state i, return to this state in n stages, that is to say $p_{i,i}^{(n)}$, then we have seen previously that $\lim_{n\to\infty} p_{0,0}^{(n)} = 0$. Moreover, as this argument is valid for every $i \in I$, then

$$\lim_{n\to\infty} p_{i,i}^{(n)} = 0, \forall i \in I.$$

But, on the other hand, it is not an invariant distribution because the state space I is not finite in this case.

Example: Consider the HMC, with $0 < \alpha, \beta < 1$, with transition matrix

$$\Pi = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Therefore, the state diagram in this case is

$$\begin{tikzcd} 1 \arrow[loop left, "1-\alpha"] \arrow[r, bend left, "\alpha"] & 2 \arrow[l, bend left, "\beta"] \arrow[loop right, "1-\beta"] \end{tikzcd}$$

We have already worked on the expression of Π^n, which was

$$\Pi^n = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta} \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}$$

Now, taking the limit of Π^n as n approaches infinity, we get that

$$\lim_{n\to\infty} \Pi^n = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}$$

Therefore, we have

$$\lim_{n\to\infty} p_{i,0}^{(n)} = \frac{\beta}{\alpha+\beta} = \Pi_0, \text{ for } i = 0, 1,$$
$$\lim_{n\to\infty} p_{i,1}^{(n)} = \frac{\alpha}{\alpha+\beta} = \Pi_1, \text{ for } i = 0, 1.$$

In consequence, applying the preceding theorem, we obtain that

$$\Pi = (\Pi_0, \Pi_1) = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$$

is an invariant distribution.

5.2 Ergodicity

Previously, we have studied the existence of stationary distributions but, in this section, we work on the uniqueness of this type of distributions. Furthermore, we will see that this concept is related to what we call ergodicity and we study the notion of regularity as well.

Definition 5.2.1. A time-homogeneous Markov chain is called ergodic if there exists the limit of $p_{i,j}^{(n)}$ as n approaches infinity, for all $i \in I$, and the result of this limit is independent of i. In addiction, this limit defines a non-degenerate probability distribution on I; that is to say that

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \Pi_j, \text{ with } \sum_{j\in I} \Pi_j = 1, \text{ for } \Pi_j > 0, \forall j \in I.$$

Example: In the last example, where the transition matrix was

$$\Pi = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Using the results seen, as we have that $\alpha, \beta \in (0,1)$, then $\Pi_0, \Pi_1 \in (0,1)$ and $\Pi_0 + \Pi_1 = 1$. In consequence, this is an ergodic chain.

In the following result we study the ergodicity for finite time-homogeneous Markov chains.

Theorem 5.2.2. Let $\Pi = (p_{i,j} : i,j \in I)$ be a transition matrix of a HMC with finite state space $I = \{1, 2, \dots, N\}$.

(a) Assume that there exists $n_0 \geq 1$ such that $\min_{i,j\in I} p_{i,j}^{(n_0)} > 0$. Then, there exists a non-degenerate probability distribution on I, $\tilde{\Pi} = \left(\tilde{\Pi}_j : j \in I\right)$, such that

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \tilde{\Pi}_j, \forall i \in I.$$

Moreover, we also have that $\tilde{\Pi}_j = \sum_{k\in I} \tilde{\Pi}_k p_{k,j}$, for $j \in I$. In other words, $\tilde{\Pi}$ is a stationary distribution for Π.

(b) The reciprocal is also satisfied; that is to say that if there exists a non-degenerate probability distribution on $I, \tilde{\Pi} = \left(\tilde{\Pi}_j : j \in I\right)$, such that

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \tilde{\Pi}_j, \forall i \in I;$$

then, there exists $n_0 \geq 1$ such that $\min_{i,j\in I} p_{i,j}^{(n_0)} > 0$.

Proof.

(a) Firstly, we define $m_j^{(n)} = \min_{i\in I} p_{i,j}^{(n)}$ and $M_j^{(n)} = \max_{i\in I} p_{i,j}^{(n)}$. Using the Chapman-Kolmogorov equation, we get that $p_{i,j}^{(n+1)} = \sum_{k\in I} p_{i,k} p_{k,j}^{(n)}$. Therefore, $\forall j \in I$, we have

- $m_j^{(n+1)} = \min_{i\in I} p_{i,j}^{(n+1)} = \min_{i\in I} \sum_{k\in I} p_{i,k} p_{k,j}^{(n)} \geq \min_{i\in I} \sum_{k\in I} p_{i,k} \min_{k\in I} p_{k,j}^{(n)}$
$= \min_{k\in I} p_{k,j}^{(n)} \min_{i\in I} \sum_{k\in I} p_{i,k} = \min_{k\in I} p_{k,j}^{(n)} = m_j^{(n)}.$
- $M_j^{(n+1)} = \max_{i\in I} p_{i,j}^{(n+1)} = \max_{i\in I} \sum_{k\in I} p_{i,k} p_{k,j}^{(n)} \leq \max_{i\in I} \sum_{k\in I} p_{i,k} \max_{k\in I} p_{k,j}^{(n)}$
$= \max_{k\in I} p_{k,j}^{(n)} \max_{i\in I} \sum_{k\in I} p_{i,k} = \max_{k\in I} p_{k,j}^{(n)} = M_j^{(n)}.$

So that, the sequence $\left\{m_j^{(n)}\right\}$ is increasing and the sequence $\left\{M_j^{(n)}\right\}$ is decreasing. As both sequences are bounded, then they have limit and we call these limits m_j and M_j, respectively. So, to show the result we want, is enough to prove that

$$\lim_{n\to\infty} \left(M_j^{(n)} - m_j^{(n)}\right) = 0, \text{ for } j \in I.$$

Because, if the previous property is satisfied, then we get that $m_j = M_j$. But, we also have that $m_j^{(n)} \leq p_{i,j}^{(n)} \leq M_j^{(n)}$; therefore

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \tilde{\Pi}_j, \text{ where } \tilde{\Pi}_j = m_j = M_j, \text{ for } j \in I.$$

So, now we prove that $\lim_{n\to\infty} \left(M_j^{(n)} - m_j^{(n)}\right) = 0, \text{ for } j \in I.$

We define $\epsilon = \min_{i,j\in I} p_{i,j}^{(n_0)} > 0$. Then, using the Chapman-Kolmogorov equation, we obtain

$$\begin{aligned} p_{i,j}^{(n_0+n)} &= \sum_{k\in I} p_{i,k}^{(n_0)} p_{k,j}^{(n)} = \sum_{k\in I} \left(p_{i,k}^{(n_0)} - \epsilon p_{j,k}^{(n)}\right) p_{k,j}^{(n)} + \epsilon \sum_{k\in I} p_{j,k}^{(n)} p_{k,j}^{(n)} \\ &= \sum_{k\in I} \left(p_{i,k}^{(n_0)} - \epsilon p_{j,k}^{(n)}\right) p_{k,j}^{(n)} + \epsilon p_{j,j}^{(2n)}. \end{aligned}$$

On the other hand, $p_{i,k}^{(n_0)} - \epsilon p_{j,k}^{(n)} \geq 0$, because $\epsilon \leq p_{i,k}^{(n_0)}$ and $p_{j,k}^{(n)} \leq 1$. So that, we get

$$\begin{aligned} p_{i,j}^{(n_0+n)} &= \sum_{k\in I} \left(p_{i,k}^{(n_0)} - \epsilon p_{j,k}^{(n)}\right) p_{k,j}^{(n)} + \epsilon p_{j,j}^{(2n)} \geq \sum_{k\in I} \left(p_{i,k}^{(n_0)} - \epsilon p_{j,k}^{(n)}\right) \min_{k\in I} p_{k,j}^{(n)} + \epsilon p_{j,j}^{(2n)} \\ &= (1-\epsilon)\, m_j^{(n)} + \epsilon p_{j,j}^{(2n)}. \end{aligned}$$

Therefore, $m_j^{(n_0+n)} \geq (1-\epsilon)\, m_j^{(n)} + \epsilon p_{j,j}^{(2n)}$.

Using a similar argument, we can get that $M_j^{(n_0+n)} \leq (1-\epsilon)\, M_j^{(n)} + \epsilon p_{j,j}^{(2n)}$.

In consequence, we have that $M_j^{(n_0+n)} - m_j^{(n_0+n)} \leq (1-\epsilon) \left(M_j^{(n)} - m_j^{(n)}\right)$.

By iteration and using that $\epsilon > 0$, for $k \geq 1$, we obtain

$$0 \leq M_j^{(kn_0+n)} - m_j^{(kn_0+n)} \leq (1-\epsilon)^k \left(M_j^{(n)} - m_j^{(n)}\right).$$

The last condition decreases to 0 as k approaches infinity. So that, there exists a subsequence $\{n_k : k \geq 1\}$ such that

$$\lim_{k\to\infty} \left(M_j^{(n_k)} - m_j^{(n_k)}\right) = 0.$$

As we also have that the sequence $\left\{M_j^{(n)} - m_j^{(n)} : n \geq 1\right\}$ is monotone; then, we obtain that

$$\lim_{n\to\infty}\left(M_j^{(n)} - m_j^{(n)}\right) = 0,$$

which is the result we wanted to prove.

Now, on one hand, we have

$$m_j^{(n)} \geq m_j^{(n_0)} \geq \epsilon > 0, \quad \forall n \geq n_0;$$

so that

$$\tilde{\Pi}_j = \lim_{n\to\infty} m_j^{(n)} \geq \epsilon > 0.$$

On the other hand, using that $\lim_{n\to\infty} \sum_{j\in I} p_{i,j}^{(n)} = 1$ and the hypothesis that $\lim_{n\to\infty} p_{i,j}^{(n)} = \tilde{\Pi}_j$, $\forall i \in I$, with I the finite state space that we are considering, we get

$$\sum_{j\in I} \tilde{\Pi}_j = 1.$$

In addiction, $\left\{\tilde{\Pi}_j : j \in I\right\}$ is a non-degenerate distribution of probability.

Finally, using again that $\tilde{\Pi}_j = \lim_{n\to\infty} p_{i,j}^{(n)}$ and considering the limit of the Chapman-Kolmogorov equation $p_{i,j}^{(n+1)} = \sum_{k\in I} p_{i,k}^{(n)} p_{k,j}$ as n approaches infinity, we obtain

$$\tilde{\Pi}_j = \lim_{n\to\infty} p_{i,j}^{(n+1)} = \lim_{n\to\infty} \sum_{k\in I} p_{i,k}^{(n)} p_{k,j} = \sum_{k\in I} p_{k,j} \lim_{n\to\infty} p_{i,k}^{(n)} = \sum_{k\in I} \tilde{\Pi}_k p_{k,j}.$$

(b) Using that $\tilde{\Pi}_j = \lim_{n\to\infty} m_j^{(n)} \geq \epsilon > 0$, we have that $\forall j \in I$ there exists $n_j \geq 1$ such that $\forall n \geq n_j, p_{i,j}^{(n)} > 0, \forall i \in I$.

So that, $\min_{i\in I} p_{i,j}^{(n)} = m_j^{(n)} > 0$, $\forall j \in I$. If we define $n_0 = \max(n_1, \ldots, n_N) \geq 1$, then

$$\min_{i\in I} p_{i,j}^{(n_0)} = m_j^{(n_0)} > 0, \forall j \in I.$$

$\square$

Definition 5.2.3. A transition matrix $\Pi = (p_{i,j} : i, j \in I)$ is regular if there exists an integer number $n_0 \geq 1$ such that

$$\min_{i,j\in I} p_{i,j}^{(n_0)} > 0.$$

Furthermore, a time-homogeneous Markov chain is called regular if its transition matrix is regular in the sense of the definition. In other words, it is regular if some power of its transition matrix has only strictly positive elements.

Proposition 5.2.4. Suppose that the state space I is finite. Then:

(a) If $\Pi = (p_{i,j} : i, j \in I)$ is regular, then the chain has only one stationary distribution, which is the one obtained in **Theorem 5.2.2.**.

(b) Let $\{X_n : n \geq 0\}$ be a regular and finite HMC. Then, the law of X_n converges weakly[2] as n approaches infinity to the stationary distribution of the chain, independently of the initial distribution considered.

Proof. Assuming that the state space I is finite, then:

(a) **Theorem 5.2.2.** shows the existence of a stationary distribution in this case. Now we want to prove the uniqueness. Suppose that there exists another stationary distribution, $\hat{\Pi}_j = \left(\hat{\Pi}_j : j \in I\right)$. Then, that invariant distribution has to verify that

$$\hat{\Pi}^t = \hat{\Pi}^t \Pi^n \Leftrightarrow \hat{\Pi}_j = \sum_{k \in I} \hat{\Pi}_k p_{k,j}^{(n)}, \text{ for } j \in I.$$

Considering the limit of the previous result as n approaches infinity, we obtain

$$\hat{\Pi}_j = \sum_{k \in I} \hat{\Pi}_k \lim_{n \to \infty} p_{k,j}^{(n)} = \sum_{k \in I} \hat{\Pi}_k \tilde{\Pi}_j,$$

where $\tilde{\Pi} = \left(\tilde{\Pi}_j : j \in I\right)$ is the invariant distribution given by **Theorem 5.2.2.**.

(b) Using the property that the invariant distribution verifies: $\sum_{i \in I} \mu_i = 1$, as well as the first part of this proposition, we have

$$\lim_{n \to \infty} P\left(X_n = j\right) = \lim_{n \to \infty} \sum_{i \in I} \mu_i p_{i,j}^{(n)} = \sum_{i \in I} \mu_i \lim_{n \to \infty} p_{i,j}^{(n)} = \left(\sum_{i \in I} \mu_i\right) \tilde{\Pi}_j = \tilde{\Pi}_j.$$

□

Observation 5.2.5. The second part of the last proposition implies that if we have a regular and finite time-homogeneous Markov chain, then we can compute approximately the law of X_n solving the system of linear equations

$$\mu^t \left(I - \Pi\right) = 0.$$

In the next result we study the property of regularity, which was defined previously:

Lemma 5.2.6. Let $\{X_n : n \geq 0\}$ be an irreducible finite HMC. If there exists $h \in I$ such that $p_{h,h} > 0$, then the chain is regular.

Proof. Consider $i, j \in I$. Using that the chain is irreducible and finite, we have that there exists $n\left(i,j\right) \geq 1$ such that $p_{i,j}^{(n(i,j))} > 0$.

If we define $m = \max_{i,j \in I} n\left(i,j\right)$, then the matrix Π^{2m} has only strictly positive elements, as we show

$$p_{i,j}^{(2m)} \geq p_{i,h}^{(n(i,h))} \underbrace{p_{h,h} \quad \cdots \quad p_{h,h}}_{2m - n(i,h) - n(h,j) \text{ times}} p_{h,j}^{(n(h,j))} > 0, \text{ for } h \in I.$$

□

[2] A sequence of random variables $\{X_n : n \geq 1\}$ converges weakly (or in law or in distribution) to a random variable X if $\lim_{n \to \infty} E\left(f\left(X_n\right)\right) = E\left(f\left(X\right)\right)$, with f a bounded continuous function.

To conclude this section, we focus on the relation between transient states and invariant distributions.

Proposition 5.2.7. Consider a HMC with finite state space I. Suppose that $i \in I$ is a transient state and μ is a stationary distribution; then, $\mu_i = 0$.

Proof. If I is finite, then

$$\mu_i = \sum_{j \in I} \mu_j p_{j,i}^{(n)}.$$

On the other hand, as i is transient, then

$$\lim_{n \to \infty} p_{j,i}^{(n)} = 0,$$

because once you have left the state i, you can never return to it.

So that, using again that the state space I is finite, we obtain

$$\mu_i = \lim_{n \to \infty} \mu_i = \lim_{n \to \infty} \sum_{j \in I} \mu_j p_{j,i}^{(n)} = \sum_{j \in I} \mu_j \lim_{n \to \infty} p_{j,i}^{(n)} = 0.$$

□

5.3 Examples

In this section we provide some examples in order to understand better the concepts studied in the previous two sections.

(i) A transition matrix is called bistochastic matrix or doubly stochastic matrix if the sum of the elements of its rows equals 1 and the sum of the elements of its columns equals 1 as well. That is to say that if $\Pi = (p_{i,j} : i, j \in I)$; then,

$$\sum_{j \in I} p_{i_0,j} = 1 \quad \text{and} \quad \sum_{i \in I} p_{i,j_0} = 1, \text{ for } i_0, j_0 \in I.$$

In addiction, if a transition matrix is symmetric, then it satisfies this property.

In fact, a matrix of this type has a stationary distribution because the system $\mu_j = \sum_{i \in I} \mu_i p_{i,j}$ verifies that

$$\mu_j = \frac{1}{|I|}, \text{ for } j \in I,$$

where $|I|$ denotes the cardinality of the state space I. The last equality is true for the reason that μ_j is independent of j and $\sum_{i \in I} p_{i,j} = 1$.

(ii) An example of existence and uniqueness of stationary distribution.

Consider the HMC with finite state space $I = \{1, 2, 3, 4, 5\}$ and transition matrix

$$\Pi = \begin{pmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix}$$

As all the elements of the main diagonal of the transition matrix Π are zero, we can not use **Lemma 5.2.6.** to prove the regularity, so we can not use **Proposition 5.2.4.** to show the existence of an only one invariant distribution either.

Using the results of the **section 5.1.**, we can write the following system of linear equations

$$\left.\begin{aligned} \mu_1 &= \tfrac{1}{3}(\mu_2+\mu_4)+\tfrac{1}{4}\mu_5, \\ \mu_2 &= \tfrac{1}{3}(\mu_1+\mu_3)+\tfrac{1}{4}\mu_5, \\ \mu_3 &= \tfrac{1}{3}(\mu_2+\mu_4)+\tfrac{1}{4}\mu_5, \\ \mu_4 &= \tfrac{1}{3}(\mu_1+\mu_3)+\tfrac{1}{4}\mu_5, \\ \mu_5 &= \tfrac{1}{3}(\mu_1+\mu_2+\mu_3+\mu_4). \end{aligned}\right\}$$

Using as well the condition that $\mu_1+\mu_2+\mu_3+\mu_4+\mu_5=1$, we obtain that the solution of the previous system of linear equations is

$$\mu_1=\mu_2=\mu_3=\mu_4=\frac{3}{16} \text{ and } \mu_5=\frac{1}{4}.$$

Furthermore, Π is regular because, for instance, Π^2 has all its elements strictly positive.

Now, the hypothesis of **Theorem 5.2.2.** and **Proposition 5.2.4.** are satisfied and we can apply both of them; therefore, we obtain that

$$\mu=(\mu_1,\mu_2,\mu_3,\mu_4,\mu_5)=\left(\frac{3}{16},\frac{3}{16},\frac{3}{16},\frac{3}{16},\frac{1}{4}\right)$$

is the only stationary distribution that has this chain.

(iii) We have already seen that if a HMC is ergodic, then there exists a stationary distribution. But the reciprocal is not true; in other words, it is possible that in a time-homogeneous Markov chain can exist an invariant distribution, but the chain can not be ergodic. We show this result giving an example.

Consider a HMC with transtion matrix

$$\Pi=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

So that, by induction we get

$$\Pi^{n+1}=\Pi=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \Pi^{n}=\Pi^2=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

That is to say that $\lim_{n\to\infty} p_{i,j}^{(n)}$ does not exist; this result proves that this chain is not ergodic.

On the other hand, using the results of the **section 5.1.**, we can write the following system of linear equations

$$\left.\begin{aligned} \mu_1 &= \mu_2, \\ \mu_2 &= \mu_1. \end{aligned}\right\}$$

Using as well the condition that $\mu_1+\mu_2=1$, we obtain that the solution of the previous system of linear equations

$$\mu=(\mu_1,\mu_2)=\left(\frac{1}{2},\frac{1}{2}\right)$$

determines a distribution of probability over the set of states I, which is invariant for Π.

6 Limit of a time-homogeneous Markov chain

In this chapter, we work on the limit of $p_{i,j}^{(n)}$ as n approaches infinity, a concept that we have already introduced in the previous chapter. Furthermore, we define the idea of mean recurrence time and we study the relation between that concept and invariant distributions. After that, we show the relation that exists between stationary distributions and the limiting probability distribution that has the chain. Finally, we define the concept of reversibility and we see an important outcome related with it.

Throughout all this chapter we consider a probability space $(\Omega, \mathcal{A}, P)$ and a collection of discrete random variables $X_n : \Omega \to I$, where I is a countable set and $\{X_n : n \geq 0\}$ is a $HMC\,(\gamma, \Pi)$ with $\Pi = (p_{i,j} : i, j \in I)$.

6.1 Behaviour in mean

Let $\{X_n : n \geq 0\}$ be an irreducible and recurrent HMC. In this section, we want to study the value of the following limit

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{l=0}^{n} p_{i,j}^{(l)}.$$

Firstly, we define the discrete random variable

$$N_j^{(n)} = \#\{X_k = j : 0 \leq k \leq n\} = \sum_{k=0}^{n} \mathbb{1}_{\{X_k = j\}}.$$

We observe that this concept is similar to the notion of number of times that the chain is located in state $j \in I$, a definition made in **chapter 4**. Furthermore, we have

$$\begin{aligned} E_i\left(N_j^{(n)}\right) &= E_i\left(\sum_{k=0}^{n} \mathbb{1}_{\{X_k=j\}}\right) = \sum_{k=0}^{n} E_i\left(\mathbb{1}_{\{X_k=j\}}\right) = \sum_{k=0}^{n} P\left(X_k = j \mid X_0 = i\right) \\ &= \sum_{k=0}^{n} p_{i,j}^{(k)}. \end{aligned}$$

Therefore, we obtain

$$\frac{1}{n+1} \sum_{k=0}^{n} p_{i,j}^{(k)} = \frac{1}{n+1} E_i\left(N_j^{(n)}\right),$$

which can be thought as the mean of the expected value of the number of times that the chain is located in state $j \in I$; it is also known as the occupation time of the state $j \in I$.

We focus on the case where $i = j$ but, firstly, we define a useful concept for this chapter.

Definition 6.1.1. Starting in a state i, the probability of going to the first time to state j in n steps is defined as

$$f_{i,j}^{(n)} = P\left(T_j = n \mid X_0 = i\right),$$

where T_j is the first instant that the chain is located in state $j \in I$, a notion defined in **chapter 4**.

We also study an important concept related to the previous definition.

Proposition 6.1.2. If $i, j \in I$, then for all $n \geq 1$ we have

$$p_{i,j}^{(n)} = \sum_{k=1}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)}.$$

Proof. Using the Markov property we get

$$\begin{aligned} p_{i,j}^{(n)} &= P_i (X_n = j) = P_i (X_n = j, T_j \leq n) = \sum_{k=1}^{n} P_i (X_n = j, T_j = k) \\ &= \sum_{k=1}^{n} P_i (X_n = j \mid T_j = k) P_i (T_j = k) \\ &= \sum_{k=1}^{n} P_i (X_n = j \mid X_1 \neq j, \ldots, X_{k-1} \neq j, X_k = j) P_i (T_j = k) \\ &= \sum_{k=1}^{n} P (X_n = j \mid X_k = j) P_i (T_j = k) = \sum_{k=1}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)}. \end{aligned}$$

□

So, now, we can study the following concept.

Definition 6.1.3. We define the expected return time of the state $j \in I$, also known as the mean recurrence time, as

$$m_{j,j} = E_j (T_j) = \sum_{k=1}^{\infty} k P (T_j = k | X_0 = j) = \sum_{k=1}^{\infty} k f_{j,j}^{(k)}.$$

Observation 6.1.4. The relation between $E_i \left(N_j^{(n)} \right)$ and $m_{j,j}$ can also be shown as follows: as $m_{j,j}$ is the mean time of one return to state j, then $\frac{n}{m_{j,j}}$ is the mean number of returns to j in n time unities. So that, in n time unities, the chain is located, in mean, $\frac{n}{m_{j,j}}$ time unities in j. Therefore, we have

$$E_j \left(N_j^{(n)} \right) \approx \frac{n}{m_{j,j}} \Rightarrow \frac{1}{n+1} E_j \left(N_j^{(n)} \right) \approx \frac{1}{m_{j,j}}.$$

Observation 6.1.5. We have considered the case where $i = j$ but, if we start in a state $i \neq j$, the result is the same. This is because after the first time that the chain is located in state j, it is not important what has happened before.

Now, we formulate a result that we will need afterward. Although we do not prove it, its demonstration in a general version can be found in reference [8].

Theorem 6.1.6. Consider the series $A(z) = \sum\limits_{n=0}^{\infty} a_n z^n$, with $a_n \geq 0$, and suppose that the series converges for $0 \leq z \leq 1$. Then

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{l=0}^{n} a_l = \lim_{z \to 1} (1 - z) A(z).$$

So, we have the theorem below.

Theorem 6.1.7. Let $\{X_n : n \geq 0\}$ be an irreducible and recurrent HMC. Then

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{l=0}^{n} p_{i,j}^{(l)} = \frac{1}{m_{j,j}}.$$

Proof. This result can be shown applying the strong law of large numbers[3] to the successive return times, which form a sequence of independent, identically distributed random variables. Another way of proving it is using **Theorem 6.1.6.**, as we do next.

In the first place, for $z \in \mathbb{R}$, we define the following function series

$$P_{i,j}(z) = \sum_{n=0}^{\infty} p_{i,j}^{(n)} z^n,$$

$$F_{i,j}(z) = \sum_{n=1}^{\infty} f_{i,j}^{(n)} z^n,$$

It can be proved that, for $|z| < 1$, both series are convergent.

Using **Proposition 6.1.2.**, we get

$$\begin{aligned} P_{i,j}(z) &= p_{i,j}^{(0)} + \sum_{n=1}^{\infty} \left[\sum_{k=1}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)} z^k \right] z^{n-k} = p_{i,j}^{(0)} + \sum_{k=1}^{\infty} f_{i,j}^{(k)} z^k \left[\sum_{n=k}^{\infty} p_{j,j}^{(n-k)} z^{n-k} \right] \\ &= p_{i,j}^{(0)} + \sum_{k=1}^{\infty} f_{i,j}^{(k)} z^k \sum_{n=0}^{\infty} p_{j,j}^{(n)} z^n = p_{i,j}^{(0)} + F_{i,j}(z) P_{j,j}(z). \end{aligned}$$

We observe that, on one hand, for $z = 1$, we have

$$F_{i,j}(1) = \sum_{n=1}^{\infty} f_{i,j}^{(n)} = P(T_j < \infty \mid X_0 = i) = \rho_{i,j},$$

where $\rho_{i,j}$ denotes a return probability, a concept defined in **chapter 4**.

And, on the other hand, for $i = j$, we get

$$P_{i,i}(z) = p_{i,i}^{(0)} + F_{i,i}(z) P_{i,i}(z) \Leftrightarrow P_{i,i}(z) [1 - F_{i,i}(z)] = 1.$$

Now, we define $A(z) = P_{i,j}(z)$. Then, we have

$$A(z) = P_{i,j}(z) = \sum_{n=0}^{\infty} p_{i,j}^{(n)} z^n = p_{i,j}^{(0)} + F_{i,j}(z) P_{j,j}(z) = p_{i,j}^{(0)} + \frac{F_{i,j}(z)}{1 - F_{j,j}(z)}.$$

[3] **Theorem (Strong law of large numbers):** Let $\{X_n : n \geq 1\}$ be a sequence of random variables with no correlation and with uniformly bounded second moment, that is to say that there exists $c \in \mathbb{R}_+$ such that $E[X_n^2] < c$, $\forall n \geq 1$. If $S_n = \sum_{i=1}^{n} X_i$, then

$$\lim_{n\to\infty} \frac{S_n - E[S_n]}{S_n} = 0,$$

where the convergence considered is the almost sure convergence.

Therefore, using that for $i, j \in I$ recurrent we have that

$$\lim_{z \to 1} F_{i,j}(z) = F_{i,j}(1) = \sum_{n=1}^{\infty} f_{i,j}^{(n)} = \rho_{i,j} = 1,$$

we get

$$\begin{aligned}\lim_{z \to 1} (1-z) A(z) = \lim_{z \to 1} (1-z) P_{i,j}(z) &= \lim_{z \to 1} (1-z) \left[p_{i,j}^{(0)} + \frac{F_{i,j}(z)}{1 - F_{j,j}(z)} \right] \\ &= p_{i,j}^{(0)} \lim_{z \to 1} (1-z) + F_{i,j}(1) \lim_{z \to 1} \frac{1-z}{1 - F_{j,j}(z)} = \lim_{z \to 1} \frac{1-z}{1 - F_{j,j}(z)}.\end{aligned}$$

Finally, using l'Hôpital's rule and applying Abel's Theorem[4] to the series $F'_{j,j}(z) = \sum_{l=1}^{\infty} l f_{j,j}^{(l)} z^{l-1}$; that is to say that we have that $\lim_{z \to 1} F'_{j,j}(z) = \sum_{l=1}^{\infty} l f_{j,j}^{(l)} = m_{j,j}$, we obtain

$$\lim_{z \to 1} \frac{1-z}{1 - F_{j,j}(z)} = \frac{1}{m_{j,j}}.$$

□

6.2 Mean recurrence time and invariant distributions

In this section, we study the relation between the mean recurrence time and the stationary distribution. To show that, we prove an important result that involves some concepts defined previously.

Definition 6.2.1. Suppose a finite state space $I = \{1, \ldots, M\}$. We define

$$\omega_j = \frac{1}{m_{j,j}}, \text{ for } j \in I.$$

Theorem 6.2.2. If a HMC is irreducible (so that, in particular, it is recurrent) with finite state space $I = \{1, \ldots, M\}$, then:

(a) $\omega = (\omega_j : j \in I)$ is a solution of $\omega^t = \omega^t \Pi$.

(b) $\sum_{j=1}^{M} \omega_j = 1$.

(c) $\omega_j > 0, \forall j \in I$.

(d) All the solutions of $\mu^t = \mu^t \Pi$ are multiples of ω.

---

[4]**Theorem (Abel's Theorem):** Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be a power series which converges for $|z| < 1$. If $\sum_{n=0}^{\infty} a_n$ converges, then

$$\lim_{z \to 1^-} f(z) = \sum_{n=0}^{\infty} a_n.$$

Proof.

(a) Firstly, using the Chapman-Kolmogorov equation, we have

$$\frac{1}{n+1}\sum_{l=0}^{n} p_{i,k}^{(l+1)} = \frac{1}{n+1}\sum_{l=0}^{n}\left(\sum_{j\in I} p_{i,j}^{(l)} p_{j,k}\right) = \sum_{j\in I}\left[\frac{1}{n+1}\sum_{l=0}^{n} p_{i,j}^{(l)}\right] p_{j,k}.$$

Now, using that the state space I is finite as well as **Theorem 6.1.7.**, we take the limit as n approaches infinity of the previous result.

On one hand, we work on the first part of the previous equality, which can be written as

$$\begin{aligned}
\frac{1}{n+1}\sum_{l=0}^{n} p_{i,k}^{(l+1)} &= \frac{1}{n+1}p_{i,k}^{(n+1)} + \frac{1}{n+1}\sum_{l=0}^{n-1} p_{i,k}^{(l+1)} \\
&= \frac{1}{n+1}p_{i,k}^{(n+1)} + \frac{1}{n+1}\sum_{l=1}^{n} p_{i,k}^{(l)} \\
&= \frac{1}{n+1}p_{i,k}^{(n+1)} + \frac{1}{n+1}\sum_{l=0}^{n} p_{i,k}^{(l)} - \frac{1}{n+1}p_{i,k}^{(0)}.
\end{aligned}$$

Therefore, taking the limit as n approaches infinity, we get

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n+1}\sum_{l=0}^{n} p_{i,k}^{(l+1)} &= \lim_{n\to\infty}\frac{1}{n+1}\sum_{l=0}^{n} p_{i,k}^{(l)} + \lim_{n\to\infty}\frac{1}{n+1}\left(p_{i,k}^{(n+1)} - p_{i,k}^{(0)}\right) \\
&= \frac{1}{m_{k,k}} = \omega_k.
\end{aligned}$$

On the other hand, we take the limit as n approaches infinity of the second part of the equality that we are considering; so that, we have

$$\begin{aligned}
\lim_{n\to\infty}\sum_{j\in I}\left[\frac{1}{n+1}\sum_{l=0}^{n} p_{i,j}^{(l)}\right] p_{j,k} &= \sum_{j\in I}\left[\lim_{n\to\infty}\frac{1}{n+1}\sum_{l=0}^{n} p_{i,j}^{(l)}\right] p_{j,k} = \sum_{j\in I}\frac{1}{m_{j,j}}p_{j,k} \\
&= \sum_{j\in I}\omega_j p_{j,k}.
\end{aligned}$$

In consequence, we obtain

$$\omega_k = \sum_{j\in I}\omega_j p_{j,k}.$$

(b) Using that the state space I is finite and that $\frac{1}{n+1}\sum_{l=0}^{n}\sum_{j=1}^{M} p_{i,j}^{(l)} = 1$, we have

$$\sum_{j=1}^{M}\omega_j = \sum_{j=1}^{M}\frac{1}{m_{j,j}} = \sum_{j=1}^{M}\left(\lim_{n\to\infty}\frac{1}{n+1}\sum_{l=0}^{n} p_{i,j}^{(l)}\right) = \lim_{n\to\infty}\frac{1}{n+1}\sum_{l=0}^{n}\sum_{j=1}^{M} p_{i,j}^{(l)} = 1.$$

(c) The previous property of this theorem implies that there exists $j_0 \in I$ such that $\omega_{j_0} > 0$. In addiction, if we take $k \in I$, then there exists $n = n(k)$ such that $p_{j_0,k}^{(n)} > 0$. So that, we obtain

$$\omega_k = \omega_{j_0} p_{j_0,k}^{(n)} + \sum_{j \neq j_0} \omega_j p_{j,k}^{(n)} > 0.$$

As this argument can be done for every state of I, we get that

$$\omega_j > 0, \forall j \in I.$$

(d) Let $x \in \mathbb{R}^M$ be a solution of $x^t = x^t \Pi$. Then, we have

$$x^t = \frac{1}{n+1} \sum_{l=0}^{n} x^t \Pi^l \Leftrightarrow x_k = \sum_{j \in I} x_j \left(\frac{1}{n+1} \sum_{l=0}^{n} p_{j,k}^{(l)} \right), \text{ for } k \in I.$$

Considering the limit of the previous result as n approaches infinity, we obtain

$$x_k = \sum_{j \in I} x_j \left[\lim_{n \to \infty} \frac{1}{n+1} \sum_{l=0}^{n} p_{j,k}^{(l)} \right] = \left(\sum_{j \in I} x_j \right) \omega_k, \text{ for } k \in I.$$

That is to say that $x_k = c\omega_k$, with $c = \sum\limits_{j=1}^{M} x_j$, for $k \in I$. $\square$

Observation 6.2.3. The previous theorem shows that in an irreducible and finite HMC:

- $\omega = (\omega_j : j \in I)$ defined in **Definition 6.2.1.**, is the only stationary distribution that has the chain.
- All the states are positive recurrent; that is due to the property **(c)** of the theorem: $\omega_j = \frac{1}{m_{j,j}} > 0$, $\forall j \in I$.

6.3 Limiting probability distribution and stationary distribution

The aim of this section is to study the relation between invariant distributions and the limiting probability distribution that has the chain. To see that, we work on a relevant outcome related to the existence of stationary distributions.

Proposition 6.3.1. Suppose that there exists the limit of $p_{i,j}^{(n)}$ as n approaches infinity, for all $i, j \in I$, and assume that the value of this limit does not depend on i; that is to say that

$$\lim_{n \to \infty} p_{i,j}^{(n)} = q_j, \forall j \in I.$$

Then, it is satisfied that

(a) $q^t = q^t \Pi$, with $0 \leq \sum\limits_{j \in I} q_j \leq 1$, where $q = (q_j : j \in I)$.

(b) $\sum\limits_{j \in I} q_j \in \{0, 1\}$.

(c)
- If $\sum\limits_{j \in I} q_j = 0$, then does not exist an invariant distribution.
- If $\sum\limits_{j \in I} q_j = 1$, then $q = (q_j : j \in I)$ is the only stationary distribution that has the chain.

Proof.

(a) On one hand, using the definition of inferior limit and the Fatou's lemma for sums[5], we have

$$0 \leq \sum_{j\in I} q_j = \sum_{j\in I} \lim_{n\to\infty} p_{i,j}^{(n)} = \sum_{j\in I} \liminf_{n\to\infty} p_{i,j}^{(n)} \leq \liminf_{n\to\infty} \sum_{j\in I} p_{i,j}^{(n)} = 1.$$

On the other hand, using again the definition of inferior limit and the Fatou's Lemma for sums, as well as the Chapman-Kolmogorov equation, we get

$$\begin{aligned}\sum_{i\in I} q_i p_{i,j} &= \sum_{i\in I} \lim_{n\to\infty} p_{k,i}^{(n)} p_{i,j} = \sum_{i\in I} \liminf_{n\to\infty} p_{k,i}^{(n)} p_{i,j} \leq \liminf_{n\to\infty} \sum_{i\in I} p_{k,i}^{(n)} p_{i,j} \\ &= \liminf_{n\to\infty} p_{k,j}^{(n+1)} = q_j.\end{aligned}$$

Now, if we suppose that there exists $j_0 \in I$ such that $\sum\limits_{i\in I} q_i p_{i,j_0} < q_{j_0}$, then we obtain

$$\sum_{j\in I} q_j > \sum_{j\in I}\sum_{i\in I} q_i p_{i,j} = \sum_{i\in I} q_i \sum_{j\in I} p_{i,j} = \sum_{i\in I} q_i,$$

which is a contradiction. Therefore, we get

$$q_j = \sum_{i\in I} q_i p_{i,j}, \forall j \in I.$$

(b) Using the previous property, we have that $q^t = q^t \Pi^n$; if we express it by coordinates and we use the dominated convergence theorem for sums[6], for $j \in I$ we get

$$q_j = \lim_{n\to\infty} \sum_{i\in I} q_i p_{i,j}^{(n)} = \sum_{i\in I} q_i \lim_{n\to\infty} p_{i,j}^{(n)} = \left(\sum_{i\in I} q_i\right) q_j.$$

So that, we obtain that

$$\sum_{i\in I} q_i \in \{0,1\}.$$

[5]**Lemma (Fatou's lemma for sums)**: Suppose that $f_n : X \to [0,\infty)$ is a sequence of functions; then, $\forall x \in X$, we have that

$$\sum_{x\in X} \liminf_{n\to\infty} f_n(x) \leq \liminf_{n\to\infty} \sum_{x\in X} f_n(x).$$

[6]**Theorem (Dominated convergence theorem for sums):** Suppose that $f_n : X \to \mathbb{C}$ is a sequence of functions on X such that $f(X) = \lim\limits_{n\to\infty} f_n(x) \in \mathbb{C}$ exists for all $x \in X$. Moreover, assume that there is a function $g : X \to [0,\infty)$, called dominating function, such that

$$|f_n(x)| \leq g(x), \forall x \in X \text{ and } \forall n \in \mathbb{N},$$

and that g also verifies that $\sum\limits_{x\in X} |g(x)| < \infty$. Then, we have that

$$\lim_{n\to\infty} \sum_{x\in X} f_n(x) = \sum_{x\in X} f(X).$$

(c) Let v be another invariant distribution, different from $q = (q_j : j \in I)$. Therefore, we have

$$v_j = \sum_{i \in I} v_i p_{i,j}^{(n)}.$$

Now, taking the limit as n approaches infinity in the previous expression, and using again the dominated convergence theorem for sums, we obtain

$$v_j = \lim_{n \to \infty} \sum_{i \in I} v_i p_{i,j}^{(n)} = \sum_{i \in I} v_i \lim_{n \to \infty} p_{i,j}^{(n)} = \left(\sum_{i \in I} v_i \right) q_j = q_j, \quad \forall j \in I.$$

□

Observation 6.3.2. If the state space I is finite and it is satisfied that $\lim_{n \to \infty} p_{i,j}^{(n)} = q_j$, then

$$\sum_{j \in I} q_j = \sum_{j \in I} \lim_{n \to \infty} p_{i,j}^{(n)} = \lim_{n \to \infty} \sum_{j \in I} p_{i,j}^{(n)} = 1.$$

Therefore, in this case there is an only one stationary distribution.

6.4 Reversibility. Metropolis-Hastings algorithm

In this section, at the beginning we define what is a reversible Markov chain. After that, we work on the following problem: given an irreducible and symmetric transition matrix, how to construct an stochastic matrix that has an invariant distribution, previously set and different from the uniform distribution. That result is known as Metropolis-Hastings algorithm.

Definition 6.4.1. A probability π over I is called reversible with respect to a matrix of transition probabilities $\Pi = (p_{i,j} : i, j \in I)$ if

$$\pi_i p_{i,j} = \pi_j p_{j,i}, \forall i, j \in I.$$

Moreover, a Markov chain is called reversible if there exists a reversible probability π over I.

Observation 6.4.2. All the reversible probabilities are invariant, as we see next

$$\sum_{i \in I} \pi_i p_{i,j} = \sum_{i \in I} \pi_j p_{j,i} = \pi_j \sum_{i \in I} p_{j,i} = \pi_j.$$

Now, we can show the following result, which is called Metropolis-Hastings algorithm.

Theorem 6.4.3. Let $Q = (q_{i,j} : i, j \in I)$ be a matrix of transition probabilities of an irreducible and symmetric Markov chain and suppose that the state space I is finite. Let π be a distribution over I, different from the uniform distribution. Then, there exists a regular transition matrix Π and a stationary distribution π for Π.

Proof. Firstly, we construct the transition matrix Π.

We consider a probability π over I such that $\pi_i > 0$, for $i \in I$. We define

$$p_{i,j} = \begin{cases} q_{i,j}, & \text{if } \ i \neq j \text{ and } \ \pi_j \geq \pi_i, \\ q_{i,j} \frac{\pi_j}{\pi_i}, & \text{if } \ i \neq j \text{ and } \ \pi_j < \pi_i, \\ 1 - \sum_{j \neq i} p_{i,j}, & \text{if } \ i = j. \end{cases}$$

Then, the matrix $\Pi = (p_{i,j} : i, j \in I)$ is a stochastic matrix, because it satisfies that

- $p_{i,j} \in [0, 1]$.
- $\sum\limits_{j \in I} p_{i,j} = 1$, due to the construction of $p_{i,j}$.

So, using that the transition matrix Q is symmetric, we have that

- For $i \neq j$ and $\pi_j \geq \pi_i$:
$$\pi_j p_{j,i} = \pi_j q_{j,i} \frac{\pi_i}{\pi_j} = q_{j,i}\pi_i = q_{i,j}\pi_i = p_{i,j}\pi_i.$$
- For $i \neq j$ and $\pi_j < \pi_i$:
$$\pi_i p_{i,j} = \pi_i q_{i,j} \frac{\pi_j}{\pi_i} = q_{i,j}\pi_j = q_{j,i}\pi_j = p_{j,i}\pi_j.$$
- If $i = j$, then the equality $\pi_i p_{i,j} = \pi_j p_{j,i}$ is always satisfied.

So that, the probability π is reversible for Π.

Now, we prove the regularity of Π.

Using **Lemma 5.2.6.**, we only need to find a state $i_0 \in I$, such that $p_{i_0,i_0} > 0$. In fact, there exist two states $i_0, j_0 \in I$, such that $q_{i_0,j_0} > 0$ and $\pi_{j_0} < \pi_{i_0}$. To see that, we consider the set M formed by the states $i \in I$ such that $\pi_i = \max\limits_{j \in I} \pi_j > 0$. As Q is irreducible, there exist $i_0 \in M$ and $j_0 \in M^c$ such that $q_{i_0,j_0} > 0$, otherwise, we have that M is a closed class. In addiction, the definition of M implies that $\pi_{j_0} < \pi_{i_0}$.

In consequence, using that if $i \neq j$, then $p_{i,j} \leq q_{i,j}$, and that the transition matrix Q is symmetric, we obtain

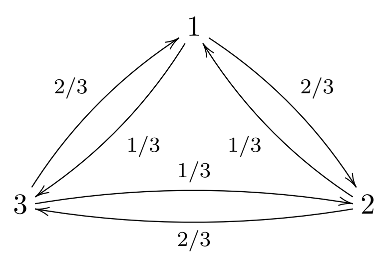
$$\begin{aligned}
p_{i_0,i_0} &= 1 - \sum_{j \neq i_0} p_{i_0,j} = 1 - \sum_{j \neq i_0, j_0} p_{i_0,j} - p_{i_0,j_0} \geq 1 - \sum_{j \neq i_0, j_0} q_{i_0,j} - q_{i_0,j_0}\frac{\pi_{j_0}}{\pi_{i_0}} \\
&= 1 - \sum_{j \neq i_0, j_0} q_{i_0,j} - q_{i_0,j_0}\frac{\pi_{j_0}}{\pi_{i_0}} + q_{i_0,j_0} - q_{i_0,j_0} = 1 - \sum_{j \neq i_0} q_{i_0,j} + q_{i_0,j_0}\left(1 - \frac{\pi_{j_0}}{\pi_{i_0}}\right) \\
&= q_{i_0,i_0} + q_{i_0,j_0}\left(1 - \frac{\pi_{j_0}}{\pi_{i_0}}\right) \geq q_{i_0,j_0}\left(1 - \frac{\pi_{j_0}}{\pi_{i_0}}\right) > 0.
\end{aligned}$$

□

The condition of symmetry in the previous theorem is fundamental as we see next.

Example: Here we see that although there exists a stationary distribution, if the chain is not symmetric, then it could happen that the chain is not reversible. So that, without the reversibility property, the proof could not have been done as before.

Consider the Markov chain with state diagram



The transition matrix in this case is

$$\Pi = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{pmatrix}$$

We have that $\pi = (\pi_1, \pi_2, \pi_3) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is an invariant distribution and Π is not symmetric. Furthermore, as for instance we have

$$\pi_1 p_{1,2} = \frac{1}{3}\,\frac{2}{3} = \frac{2}{9} \neq \frac{1}{9} = \frac{1}{3}\,\frac{1}{3} = \pi_2 p_{2,1},$$

we obtain that the chain is not reversible.

7 Conclusions

In the first place, after having done this final project, I could assert that as much as the field of it as the specific topic have been completely appropriate. My knowledge in the probabilistic area have been able to grow by learning how finite Markov chains work. Nevertheless, I am aware of the difficulty and importance that continuous-time Markov chains have because, in that case, all the results are totally generals.

This project has allowed me to increase my mathematical knowledge and I have improved the accuracy of writing it. Moreover, while I have been researching and writing the project, I realize of how to be rigorous when we are working with scientific texts. I would also like to emphasize that doing it in English has increased its level of difficulty, although my global evaluation of it is completely positive because my mathematical and English skills have both risen.

A Probability review

A.1 General outcomes

In this section, we summarize the most important results of probability that are necessary for a correct understanding of this work.

Definition A.1.1. A probability space is a triple $(\Omega, \mathcal{A}, P)$ comprised of the following three elements

1. Ω is the sample space, the set of possible outcomes.
2. $\mathcal{A}$ is a σ-algebra or a σ-field, a collection of subsets of Ω. It means that
   (a) $\Omega \in \mathcal{A}$.
   (b) If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$.
   (c) If $\{A_i : i \in I\} \subseteq \mathcal{A}$, then $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{A}$.
3. P is a mapping $P : \mathcal{A} \to [0, 1]$, called probability, which satisfies
   (a) $P(\Omega) = 1$.
   (b) For any sequence $\{A_n : n \geq 1\} \subseteq \mathcal{A}$ mutually disjoint (in other words, $A_i \cap A_j = \emptyset$, for all $i \neq j$)
   $$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Now we study two particular cases. One one hand, the finite or numerable case; on the other hand, the equiprobable case.

Definition A.1.2. Let us assume that $\Omega = \{\omega_i : i \in I\}$, where I is finite or countable. Then, $\{p_i, i \in I\}$ is a probability if $0 \leq p_i \leq 1$, $\forall i \in I$, with $\sum\limits_{i \in I} p_i = 1$. Moreover, the probability of any event $A \subseteq \Omega$ is defined as

$$P(A) = \sum_{i, w_i \in A} P(\{w_i\}) = \sum_{i, w_i \in A} p_i.$$

Definition A.1.3. Suppose that $\Omega = \{\omega_i; i \in I\}$, where I is finite, that is to say that $I = \{1, 2, \ldots, n\}$, and let us assume that the sample space is equiprobable. Then, the probabilities $\{p_1, \ldots, p_n\}$ verify that $p_i = p_j$, $\forall i \neq j$; so that, $p_i = \frac{1}{n}$, $\forall i = 1, \ldots, n$. Furthermore, the probability of any event $A \subseteq \Omega$ is defined as

$$P(A) = \sum_{i, w_i \in A} p_i = \sum_{i, w_i \in A} \frac{1}{n} = \frac{1}{\#\Omega} \sum_{i, w_i \in A} 1 = \frac{\#A}{\#\Omega}.$$

In the following result we enumerate the most important properties of probability.

Proposition A.1.4.

1. $P(\emptyset) = 0$.
2. If $A, B \in \mathcal{A}$ are disjoint, then $P(A \cup B) = P(A) + P(B)$.
3. If $A_1, A_2, \ldots, A_n \in \mathcal{A}$ are mutually disjoint, then $P\left(\bigcup\limits_{i=1}^{n} A_i\right) = \sum\limits_{i=1}^{n} P(A_i)$.
4. $P(A^c) = 1 - P(A)$, for all $A \in \mathcal{A}$.
5. If $A, B \in \mathcal{A}$, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
6. If $A, B \in \mathcal{A}$ and $A \subseteq B$, then $P(A) \leq P(B)$.
7. If $A_1, A_2, \ldots, A_n \in \mathcal{A}$, then $P\left(\bigcup\limits_{i=1}^{n} A_i\right) \leq \sum\limits_{i=1}^{n} P(A_i)$.
8. Let $\{A_n : n \geq 1\} \subseteq \mathcal{A}$ be a sequence of sets. Then
   (a) If the sequence is increasing, then $P\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = \lim\limits_{n\to\infty} P(A_n)$.
   (b) If the sequence is decreasing, then $P\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = \lim\limits_{n\to\infty} P(A_n)$.

A.2 Conditional probability

In this section we define the concept of conditional probability and we explain the most important results related to this notion.

Definition A.2.1. Given $A, B \in \mathcal{A}$ with $P(B) > 0$, the conditional probability of A given B is defined as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

The following result is known as chain rule or general product rule.

Theorem A.2.2. If we consider $A_1, A_2, \ldots, A_n \in \mathcal{A}$ with $P(A_1 \cap \ldots \cap A_{n-1}) > 0$, then

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) P(A_2 \mid A_1) \ldots P(A_n \mid A_1 \cap \ldots \cap A_{n-1}).$$

The law of total probability is explained in the next outcome.

Theorem A.2.3. Let $\{B_i : 1 \leq i \leq n\} \subseteq \mathcal{A}$ be a partition of Ω, where each set of the partition has strictly positive probability. Then, for all $A \in \mathcal{A}$

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i).$$

To conclude this section, we study the Bayes' theorem or Bayes' law or Bayes' rule.

Theorem A.2.4. Let $\{A_i : 1 \leq i \leq n\} \subseteq \mathcal{A}$ and $\{B_j : 1 \leq j \leq m\} \subseteq \mathcal{A}$ be two partitions of Ω, where each set of both partitions has strictly positive probability. Then

$$P(A_i \mid B_j) = \frac{P(A_i) P(B_j \mid A_i)}{P(B_j)} = \frac{P(A_i) P(B_j \mid A_i)}{\sum_{k=1}^{n} P(B_j \mid A_k) P(A_k)}.$$

A.3 Independence

In this section, we define the idea that two or more sets are independent. Moreover, we explain the concept of conditional independence and the relation that it has to what we call a Markov process.

Definition A.3.1. Two sets $A, B \in \mathcal{A}$ are independent if

$$P(A \cap B) = P(A) P(B).$$

This definition can be extended as we show next.

Definition A.3.2. The sets $A_1, A_2, \ldots, A_n \in \mathcal{A}$ are independent if

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \ldots P(A_{i_k}),$$

for all $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$.

Definition A.3.3. Assume that $A, B, C \in \mathcal{A}$ with $P(C) > 0$. The sets A and B are conditionally independent given C if

$$P(A \cap B \mid C) = P(A \mid C) P(B \mid C).$$

Definition A.3.4. The events $\{A, B, C\}$ with $P(A \cap B) > 0$ form a Markov process if

$$P(C \mid A \cap B) = P(C \mid B).$$

Observation A.3.5. A, B and C should be considered as a chronological sequence: C as the future, B as the present and A as the past. Thus, the previous relation shows that there is only dependence on the present, and there is not on the past.

Proposition A.3.6. A and C are conditionally independent given B if, and only if, the events $\{A, B, C\}$ form a Markov process. That is to say that

$$P(A \cap C \mid B) = P(A \mid B) P(C \mid B) \Leftrightarrow P(C \mid A \cap B) = P(C \mid B).$$

A.4 Discrete Random Variables

In this section we define what is a random variable and we focus on an important type of them. In addiction, we explain some properties that random variable have.

Definition A.4.1. A random variable is a mapping $X : \Omega \to \mathbb{R}$ such that

$$X^{-1}(B) = \{\omega \in \Omega; X(\omega) \in B\} \in \mathcal{A}, \forall B \in \mathcal{B}(\mathbb{R})$$

where $\mathcal{B}(\mathbb{R})$ means the Borel set of $\mathbb{R}$.

Definition A.4.2. The law or distribution of a random variable X, written as P_X, is the probability over $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, and it is defined as

$$P_X(B) = P\left(X^{-1}(B)\right), \forall B \in \mathcal{B}(\mathbb{R}).$$

Definition A.4.3. A random variable X is discrete if the set $X(\Omega)$ is finite or countable. Moreover, this set is represented by $\{x_i : i \in I \subseteq \mathbb{N}\}$.

Therefore, we can establish the law or distribution of a discrete random variable, which is defined by: $P_X(\{x_i\}) = P(X = x_i) = p(x_i)$, for $i \in I$, and it is called probability (mass) function.

Definition A.4.4. A collection $X_i : \Omega \to \mathbb{N}$, with $i = 1, \ldots, n$, of random variables is independent if, for any $x_1, x_2, \ldots, x_n \in \mathbb{N}$, we have

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{j=1}^{n} P(X_j = x_j).$$

A.5 Expected value

Definition A.5.1. The expected value of a discrete random variable $X : \Omega \to \{x_i : i \in I \subseteq \mathbb{N}\}$ is well-defined if, and only if,

$$\sum_{i \in I} |x_i| P(X = x_i) < \infty.$$

In that case, the expected value of X is

$$E(X) = \sum_{i \in I} x_i P(X = x_i).$$

B Problems and applications

1. Let A and B be two sets of the σ-algebra $\mathcal{A}$. Prove that:

$$P(A \cap B) \geq P(A) + P(B) - 1.$$

2. Consider three groups of people: A, B and C, with 57, 49 and 43 members, respectively. A and B have 13 members in common, A and C have 7 members in common and B and C have 4 members in common. Finally, there is one person that belongs to the three groups. We choose one member randomly.

   (a) Write what is the associated probabilistic model in this case.

   Consider the events $A =$ be a member of group A, $B =$ be a member of group B and $C =$ be a member of group C.

   (b) Are A and B independent?

   (c) Are A and B independent given C?

3. An event A is favorable to another event B if and only if

$$P(A \cap B) \geq P(A) P(B).$$

   We express this relation with $A \| B$. Prove that:

   (a) If A and B are independent, then $A \| B$.

   (b) The relation $\|$ is reflexive and symmetric, but it is not transitive.

4. Prove that if $A \cap B = \emptyset$, then:

   (a) If $A \| C$ and $B \| C$, then $(A \cup B) \| C$.

   (b) If $C \| A$ and $C \| B$, then $C \| (A \cup B)$.

5. Prove that if

$$P(A \mid C) \geq P(B \mid C) \quad \text{and} \quad P(A \mid C^c) \geq P(B \mid C^c),$$

   then $P(A) \geq P(B)$.

6. We roll a fair die twice.

   (a) If the outcome of the first die is 4, what is the probability of getting the value of k, with $1 \leq k \leq 6$, on the second die?

   (b) We know that the two faces add up to 7, what is the probability of having get the value of k, with $1 \leq k \leq 6$, on the first die?

7. Consider a particle which moves along the integer number line, one unity to the left or one unity to the right with probability $\frac{1}{2}$. Each movement is independent of the previous ones. The position after n movements is denoted as Y_n. Compute the following probabilities:

   (a) $P(Y_n \geq 0 \text{ for } 1 \leq n \leq 4)$.

   (b) $P(|Y_n| \leq 2 \text{ for } 1 \leq n \leq 4)$.

   (c) $P(Y_n \geq 0 \text{ for } 1 \leq n \leq 4 \mid Y_4 = 0)$.

8. In the same situation as in the previous exercise, verify that

$$P(Y_n = 4 \mid Y_{n-2} = 2, Y_{n-3} = 1) = P(Y_n = 4 \mid Y_{n-2} = 2),$$

for any integer n.

9. Let X_n and Y_n be two independent Markov chains, both with the same discrete state space S, and with the same transition probability. We define the process $Z_n = (X_n, Y_n)$ with state space $S \times S$. Prove that Z_n is a Markov chain and obtain the matrix of transition probabilities.

10. Let $(Z_i : i \in \mathbb{N})$ be a sequence of independent and identically distributed random variables, which follow a Bernoulli(p) distribution. We define $X_n = (Z_{n-1}, Z_n)$. Study $\{X_n : n \geq 1\}$; is it a Markov chain? Compute the powers of the matrix of transition probabilities obtained.

11. Let $(Y_k : k \in \mathbb{N})$ be a sequence of independent and identically distributed random variables such that $P(Y_n = -1) = \frac{1}{2}$, $P(Y_n = 1) = \frac{1}{2}$. We define, for all $n \geq 0$,

$$X_n = \frac{Y_n + Y_{n-1}}{2}.$$

Is it a Markov chain?

12. We toss a biased coin repeatedly, let p be the probability of obtaining a head and $q = 1 - p$ the probability of getting a tail. Let H_n and T_n be the number of heads and tails obtained in the first n throws. Define $X_n = H_n - T_n$. Is it a time-homogeneous Markov chain?

13. Let $\{X_n : n \geq 1\}$ be a sequence of random variables, whose values are integers numbers. Let $S_n = \sum_{i=1}^{n} X_i$.

    (a) If the random variables are independent but not identically distributed, is $\{S_n : n \geq 1\}$ always a time-homogeneous Markov chain?

    (b) If the random variables are identically distributed but not independent, is $\{S_n : n \geq 1\}$ always a time-homogeneous Markov chain?

14. Let $\{S_n : n \geq 1\}$ be a simple random walk, with $S_0 = 0$. We define $X_n = S_n + 5$. Is $\{X_n : n \geq 1\}$ a time-homogeneous Markov chain?

15. The weather in the land of Oz.

    The Land of Oz is blessed by many things, but not by good weather. They never have two nice days in a row. If they have a nice day, they are just as likely to have snow as rain the next day. If they have snow or rain, they have an even chance of having the same the next day. If there is change from snow or rain, only half of the time is this a change to a nice day. Find the Markov chain associated with the weather in the land of Oz.

16. The drunk walk.

    A drunk man walks along the street between his home and a bench. We can denote the steps he does by the integer numbers between 0 (his home) and n (the bench) and we consider that the starting point i is a bar. The drunk walks one step to the right with probability p, with $0 < p < 1$, or one step to the left with probability $q = 1 - p$. When he arrives to the bench or home, he stays there. Find the Markov chain associated with the drunk walk.

17. The guard of the strong square.

    In order to confuse the enemies, a guard does his watch at the four corners of a strong square in the following way: after waiting 5 minutes in each corner, he toss a fair coin and, if the outcome is head, then he goes to the corner he has on his left and he remains there for 5 minutes; otherwise, if the outcome is tail, then he goes to his right. He repeat this process infinite times. Find the Markov chain associated with the guard of the strong square. In addiction, if $p = \frac{1}{2}$, where p is the probability that he obtains a head, then compute the nth power of its matrix of transition probabilities.

    (a) Suppose that the guard chooses randomly in which of the four corners he begins, all of them with the same probability. Let X_n be the random variable that shows where is the guard located after $5n$ minutes. Compute the law of X_n.
    (b) Study what is the result if the guard starts in a set corner.

18. Consider the state space $\{0, 1, 2, 3, 4, 5, 6\}$ and the Markov chain with matrix of transition probabilities

$$\begin{pmatrix} 0 & 0 & \frac{1}{7} & \frac{6}{7} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

    (a) Determine the absorbing states.
    (b) Compute the absorption probabilities.
    (c) Starting from state 0, what is the probability of finishing in state 4? And of finishing in state 6?

19. Give an example of a time-homogeneous Markov chain that verifies the following conditions: have more than 5 states, have more than 2 classes, where some of them have period different than 1 and have, at least, 2 absorbing states.

20. Two players A and B take part in a tennis game. In each play, A has a probability p of getting a point and a probability $q = 1 - p$ of loosing it.

    (a) Find the Markov chain associated with the tennis game, with absorbing and transient states.
    (b) Compute the absorption probabilities.

21. Prove that a finite chain is aperiodic and has only one class if, and only if, there exists a natural number n such that, for all the states of the chain i, k, $p^n_{i,k} > 0$.

22. We choose numbers of the set $\{1, \ldots, m\}$ randomly and with replacement. In each stage, we consider the biggest number of the ones selected until that moment. Find the Markov chain associated with this situation and study it.

23. Two players A and B take part in a game in the following way: initially, an urn contains 2 red balls, R, and 2 black balls, N. One ball is extracted and if it is N, then the ball is removed; otherwise, if it is R, then this ball is returned to the urn and a N ball is added as well. We repeat this process until one of the two players wins: the player A wins when there are 4 N balls in the urn and the player B wins when there is no one N ball in the urn.

    (a) Find the Markov chain associated with this game and determine whose matrix of transition probabilities.

    (b) Study the classes of this Markov chain.

    (c) After two extractions, what is the probability of remaining 2 N balls in the urn? And after three extractions?

    (d) What is the probability that player A wins the game? Which of both players have a higher probability of winning the game?

    (e) What is the mean time of the game?

24. Study the Markov chains given by the following stochastic matrices:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

25. Consider the following stochastic matrix

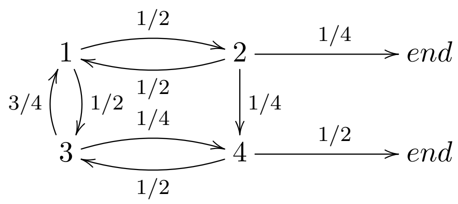
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{8} & \frac{3}{8} \\ \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

    (a) Study the states of the Markov chain associated with this matrix of transition probabilities.

    (b) Is it an irreducible chain? Is it ergodic? Is it regular?

    (c) Compute the absorption probabilities.

26. The following figure represents a program flowchart. In each time unity, the program goes from one state to another according to the graphic and the indicated probabilities below:



Is it possible to describe the evolution of that program using a Markov chain? If the answer is positive, give its matrix of transition probabilities. Starting from 1, what is the mean running time of the program? And starting from 3?

27. Consider the Markov chain with states $E = \{1, 2, 3, 4, 5\}$ and with matrix of transition probabilities

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

(a) Determine the transient and recurrent states.

(b) Compute the invariant distributions.

(c) Starting in 1, what is the probability of passing through 4 or 5?

28. Prove that if the matrix of transition probabilities of a Markov chain is

$$\Pi = \begin{pmatrix} q & 1-q \\ \beta & 1-\beta \end{pmatrix}$$

with $q \neq 1 + \beta$, then

$$\lim_{n\to\infty} \Pi^n = \frac{1}{1+\beta-q} \begin{pmatrix} \beta & 1-q \\ \beta & 1-q \end{pmatrix}.$$

29. A man has three houses: one in the city of Barcelona, one in the beach and another one in the mountain. So, each weekend he has three options: stay in the city, go to the beach or go to the mountain. He makes his decision according to the following rules: if the previous weekend he has stayed in the city, then he decides to go to the beach or to the mountain with probability $\frac{1}{2}$. However, if the last weekend he has gone to the beach, then he chooses to stay in the city. Finally, if the previous weekend he has gone to the mountain, then either he decides to go to the beach with probability $\frac{3}{4}$ or staying in the city with probability $\frac{1}{4}$. We consider the Markov chain associated with this situation.

(a) Is it an ergodic chain? Is it regular? Is it irreducible?

(b) Compute a stationary distribution.

(c) What percentage of weekends does he stay in the city?

30. Compute the mean recurrence time for the Markov chains with stochastic matrices:

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 1-2p & 2p & 0 \\ p & 1-2p & 0 \\ 0 & 2p & 1-2p \end{pmatrix}, \begin{pmatrix} 0 & p & 0 & 1-p \\ 1-p & 0 & p & 0 \\ 0 & 1-p & 0 & p \\ p & 0 & 1-p & 0 \end{pmatrix}$$

References

[1] Brémaud, P.: *Markov chains: Gibbs fields and Monte Carlo simulation, and queues*, New York: Springer, cop 1999.

[2] Chung, K.L.: *Elementary probability theory with stochastic processes*, New York: Springer, 1974.

[3] Grimmett, G.R. and Stirzaker, D.R.: *Probability and random processes*, Oxford: Oxford University Press, 2001.

[4] Grinstead, Ch. M. and Laurie Snell, J.: *Introduction to probability*, Providence (R.I.) : American Mathematical Society, 1997.

[5] Kemeny, J.G. and Laurie Snell, J.: *Finite Markov chains*, New York : Springer, 1976.

[6] Norris, J.R.: *Markov chains*, Cambridge: Cambridge University Press, 1997.

[7] Sanz, M.: *Probabilitats*, Barcelona: Edicions Universitat de Barcelona, 1999.

[8] Zygmund, A.: *Trigonometric series*, Cambridge: Cambridge University Press, 1959.