Markov Chains

INDER K. RANA
Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai 400076, India
email: ikrana@iitb.ac.in

Abstract. These notes were originally prepared for a College Teacher's Refresher course at University of Mumbai. The current revised version is for the participants of the Summer school on "Probability Theory" at Kerala School of Mathematics,2010.

Contents

Basic Probability Theory

0.1. Probability space

A mathematical model for analyzing statistical experiments is given by a **probability space**. A probability space is a triple $(\Omega, \mathcal{S}, P)$ where:

- Ω is a set representing the set of all possible outcomes of the experiment.
- $\mathcal{S}$ is a σ-algebra of subsets of Ω. Subsets of Ω are called **events** of the experiment. Elements of $\mathcal{S}$ represents the collection of events of interest in that experiment.
- For every $E \in \mathcal{S}$, the nonnegative number $P(E)$ is the probability that the event E occurs. The map $E \mapsto P(E)$, called a **probability**, is $P : \mathcal{S} \to [0, 1]$, with the following properties:

(i) $P(\emptyset) = 0$, and $P(\Omega) = 1$.

(ii) P is **countably additive**,i.e., for countable sequence $A_1, A_2, \ldots, A_n, \ldots$ in $\mathcal{S}$, which is pairwise disjoint: $A_i \cap A_j = \emptyset$,

$$P(\cup_{n=1}^{\infty}(A_i)) = \sum_{i=1}^{\infty} P(A_i).$$

0.2. Conditional probability

Let $(\Omega, \mathcal{S}, P)$ be a probability space. If B is an event with $P(B) > 0$, then for every $A \in S$, the **conditional probability** of A given B, denoted by $P(A|B)$, is defined by

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

Intuitively,$P_B(A) := P(A|B)$ is as how likely is the event A to occur, given the knowledge that B has occurred.

Some properties of conditional probability are:

(i) For countable sequence $A_1, A_2, \ldots, A_n, \ldots$ in $\mathcal{S}$, which is pairwise disjoint

$$P_B(\cup_{n=1}^{\infty}(A_i)) = \sum_{i=1}^{\infty} P_B(A_i).$$

(ii) **Chain rule**

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

In general, for $A_1, A_2, \ldots, A_n \in \mathcal{S}$,

$$P(A_1 \cap A_2 \cap \ldots \cap A_n)$$
$$= P(A_1|A_2 \cap A_2 \cap \ldots \cap A_n)\, P(A_2|A_3 \cap A_2 \cap \ldots \cap A_n) \ldots P(A_{n-1}|A_n),$$

and for $B \in \mathcal{S}$,

$$P(A_1 \cap A_2 \cap \ldots \cap A_n|B)$$
$$= P(A_1|A_2 \ldots A_n \cap B)\, P(A_2|A_3 \cap A_2 \ldots A_n \cap B) \ldots P(A_n|B).$$

(iii) **Bay's formula** If $A_1, A_2, \ldots, A_n, \ldots$ in $\mathcal{S}$, are pairwise disjoint and $\Omega = \cup_{n=1}^{\infty} A_i$, then for $B \in \mathcal{S}$,

$$P(A_i|B) = \frac{P(B|A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}$$

(iv) **Conditional impendence** Let $A_1, A_2, \ldots, A_n, \ldots$ in $\mathcal{S}$, be pairwise disjoint such that

$$P(A|A_i) = P(A|A_j) := p \text{ for every } 1, j$$

then $P(A| \cup_{i=1}^{\infty} A_i) = p$.

(v) If $A_1, A_2, \ldots, A_n, \ldots$ in $\mathcal{S}$, are pairwise disjoint and $\Omega = \cup_{n=1}^{\infty} A_i$, then for $B, C \in \mathcal{S}$,

$$P(C|B) = \sum_{i=1}^{\infty} P(A_i|B)\, P(C|A_i \cap D)$$

Chapter 1

Basics

1.1. Introduction

The aim of our lectures is to analyze the following situation: Consider an experiment/system under observation and let $s_1, s_2, ..., s_n, ...$ be the possible states in which the system can be. Let us suppose that the system is being observed at every unit of time: $n = 0, 1, 2, \ldots$. Let X_n denote the observation at time $n \geq 0$. Thus each X_n can take either of the values $s_1, s_2, ..., s_n, ...$ We further assume that the observations $X_{n's}$ are not 'deterministic', i.e., X_n can take value s_i with some probability. In other words, each X_n is a random variables on some probability space $(\Omega, \mathcal{A}, P)$. In case,the observations $X_0, X_1, ...$ are independent, we know how to compute the probabilities of various events. The situation we are going to look at is slightly more general. Let us look at some examples.

1.1.1 Example:
Consider observing the working of a particular machine in a factory. On any day, either the machine will be broken or it will be working. So our system can be in any one of the two states: 'broken' - represented by 0 , or 'working'- represented by 1. Let X_n be the observation about the machine on n^{th} day. Clearly, there is no reason to assume that X_n will be independent of $X_{n-1}, ..., X_0$.

1.1.2 Example:
Consider a gambler making bets in a gambling house. He starts with some amount say A rupees and makes a series of one rupee bets against the house. Let $X_n,\ n \geq 0$ denote the gambler's capital at time n, say after n bets. Then, the states of the system, the possible values each X_n can take, are $0, 1, 2, ...$ Clearly, the values of X_n depends upon the values of X_{n-1}.

1.1.3 Example:
Consider a bill collection office where people come to pay their bills. People arrive at the paying counter at various time points and are being served eventually. Let us suppose that we measure time in minutes. Then the number of persons that arrive during one minute are taken as the ones which arrive at that minute and let us say at most one person can be/will be served in a minute. Let ξ_n denote the number of persons that arrive at the n^{th} minute. Let X_0 denote the number of persons that were waiting initially, (i.e., when the office opened) and for $n \geq 1$, let X_n denote the number

of customers at the n^{th} minute. Thus, for all $n \geq 0$,

$$\begin{aligned} X_{n+1} &= \xi_{n+1}, \text{ if } X_n = 0, \\ X_{n+1} &= X_n + \xi_{n+1} - 1, \text{ if } X_n \geq 0, \end{aligned}$$

because one person will be served in that minute. The states of the system are $0, 1, 2, ...,$ and clearly X_{n+1} depends upon X_n.

Thus, we are going to look at a sequence of random variables $\{X_n\}_{n\geq 0}$ defined on a probability space $(\Omega, \mathcal{A}, P)$, such that each X_n can take at most countable number of values. As mentioned in the beginning, if $X_{n's}$ are independent, then one knows how to analyze the system. If $X_{n's}$ are not independent, what kind of relation $X_{n's}$ can have? For example, let us consider the system of example 1.1.1: observing the working of a machine on each day. Clearly, the observation that the machine will be "in order" or "not in order" on a particular day depends only upon the fact that it was working or was out of order on previous day. Or in example 1.1.2, the example of gambler, his capital on n^{th} day will be depend only upon his capital on the $(n-1)^{th}$ day.

This motivates the following assumption about our system.

1.1.4 Definition:
Let $\{X_n\}_{n\geq 0}$ be a sequence of random variables taking values in a set S, called **state space**, which is at most a countable set. We say that has $\{X_n\}_{n\geq 0}$ has the **markov property** if for every $n \geq 0$ and $i_0, i_1, \dots i_n \in S$,

$$P\{X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, ...X_n = i_n\}$$
$$= P\{X_{n+1} = i_{n+1} | X_n = i_n\} \text{ for all } n \geq 0.$$

That is, the observation/outcome at the $(n+1)^{th}$ stage of the experiment depends only on the outcome immediate past. Thus, if $n \geq 0$, and $i, j \in S$, then the numbers

$$P(i, j, n) := P\{X_{n+1} = j | X_n = i\}$$

are going to be important for the system. This is the probability that the system will be in state j at stage $n+1$ given that it was in state i at stage n.

Note that saying that a sequence $\{X_n\}_{n\geq 1,}$ has markov property means that given X_{n-1}, the random variable X_n is conditionally independent of $X_{n-2}, \dots, X_1, X_0$. It means that the distribution of the sequence to go to next step depends only upon where the system is now and not where it has been in the past.

1.1.5 Definition:
Let $\{X_n\}_{n\geq 1,}$ be a markov system with state space S.

(i) For $n \geq 0$, and $i, j \in S$, the number $P(i, j, n)$ is called the **one step transition probability** for the system at stage n to go from state i to the state j at the next stage.

(ii) The system is said to have the **stationary property** or the **homogeneous property** if $P(i, j, n)$ is independent of n, i.e.,

$$P(i, j, n+1) = P(i, j, n) \text{ for every } i, j \in S, n \geq 1.$$

That is the probability that the system will be in state j at stage $n+1$ given that it is in state i at stage n is independent of n. Thus, the probability of the system in going from state i to j does not depend upon the time at which this happens.

(iii) A markov system $\{X_n\}_{n\geq 1}$ is called a **markov chain** if it is stationary.

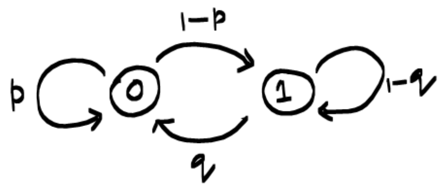
1.1.6 Definition:
Given a markov chain $\{X_n\}_{n\geq 1}$,

$$\Pi_0(i) := P\{X_0 = i\},\ i \in S$$

is called the **initial distribution vector** or the distribution of X_0.

1.1.7 Graphical representation:
A pictorial way to represent a markov chain is by its **transition graph**. It consists of nodes representing the states of the chain and arrows between the nodes representing the transition probabilities. The transition graphs of examples markov chain in example 1.1.1 is as follows:

$$\begin{aligned} p(0,0) &= p, \\ p(0,1) &= 1-p, \\ p(1,0) &= q, \\ p(1,1) &= 1-q. \end{aligned}$$



1.1.8 Theorem:
Let $\{X_n\}_{n\geq 1}$, be the markov chain with state space S, transition probabilities $p(i,j)$, and initial distribution vector $\Pi_0(i)$. Let P be the matrix

$$P = [p_{ij}]_{i\times j}.$$

Then the following hold:

(i) $0 \leq p(i,j),\ \Pi_0(i) \leq 1$.

(ii) For every i, $\sum_{j\in S} p(i,j) = 1$.

(iii) For every j, $\sum_{i\in S} \Pi_0(i) = 1$.

1.1.9 Definition:
The matrix $P = [p(i,j)]_{i\times j}$ is called the **transition matrix** of the markov chain. It has the property that each entry is a non negative number between 0 and 1, sum of each row and each column is 1.
Let us look at some examples.

1.1.10 Example:
Consider the example 1.1.1, observing the working of a machine. Here $S = \{0,1\}$. Let

$$P\{X_{n+1} = 1 | X_n = 0\} := p(0,1) = p,$$

$$P\{X_{n+1} = 0 | X_n = 1\} := p(1,0) = q.$$

Then,

$$P\{X_{n+1} = 0 | X_n = 0\} = 1-p \text{ and } \{X_{n+1} = 1 | X_n = 1\} = 1-q.$$

Thus, the transition matrix is

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

Another way of describing a markov chain is given by

1.1.11 Theorem:
A sequence of random variables $\{X_n\}_{n\geq 0}$ *is a markov chain with initial vector* Π_0 *and transition matrix* P*, if and only if for every* $n \geq 1$*, and* $i_0, i_1, \ldots, i_n \in S$*,*

$$P\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\} \\ = \Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_{n-1}, i_n). \tag{1.1}$$

Proof:
First suppose that $\{X_n\}_{n\geq 0}$ is a markov chain with initial vector Π_0 and transition matrix P. Then using the chain rule for conditional probability,

$$\begin{aligned} &P\{X_0 = i_0, X_1 = i_1, \ldots, X = i_n\} \\ &= P\{X_0 = i_0\}P\{X_1 = i_1|X_0 = i_0\}\cdots P\{X_n = i_n|X_0 = i_0, \cdots, X_{n-1} = i_{n-1}\} \\ &= \Pi_0(i)\, p(i_0, i_1)\, p(i_1, i_2)\, \cdots\, p(i_{n-1}, i_n), \end{aligned}$$

Conversely, if equation (1.1) holds, then summing both sides over $i_n \in S$,

$$\sum_{i_n\in S} P\{X_0 = i_0, X_1 = i_1, \ldots, X = i_n\} \\ = \sum_{i_n\in S} \Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_{n-1}, i_n).$$

Thus,

$$\begin{aligned} P\{X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\} &= \sum_{i_n\in S} P\{X_0 = i_0, X_1 = i_1, \ldots, X = i_n\} \\ &= \sum_{i_n\in S} \Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_{n-1}, i_n) \\ &= \Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_{n-2}, i_{n-1}). \end{aligned}$$

Proceeding similarly, we have for every $n = 0, 1, \ldots, i_k \in S$,

$$P\{X_0 = i_0, X_1 = i_1, \ldots, X_k = i_k\} \\ = \Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_{k-1}, i_k).$$

Thus, for $k = 0$, we have

$$P\{X_0 = i_0\} = \Pi_0(i)$$

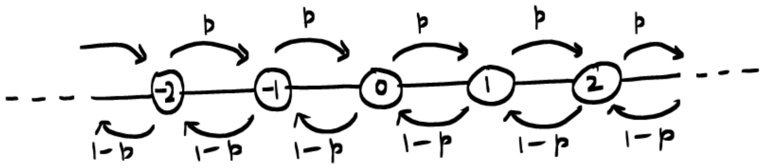
and

$$\begin{aligned} P\{X_{n+1} = i_{n+1} \,|\, X_0 = i_0, \ldots, X_n = i_n\} &= \frac{P\{X_0 = i_0, \ldots X_n = i_n, X_{n+1} = i_{n+1}\}}{P\{X_0 = i_0, \ldots X_n = i_n, X_n = i_n\}} \\ &= \frac{\Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_n, i_{n+1})}{\Pi_0(i)p(i_0, i_1)p(i_1, i_2)\cdots p(i_{n-1}, i_n)} \\ &= p(i_n, i_{n+1}). \end{aligned}$$

Hence, $\{X_n\}_{n\geq 0}$ is a markov chain with initial vector Π_0, and transition probabilities $p(i, j),\ i, j \in S$. ■

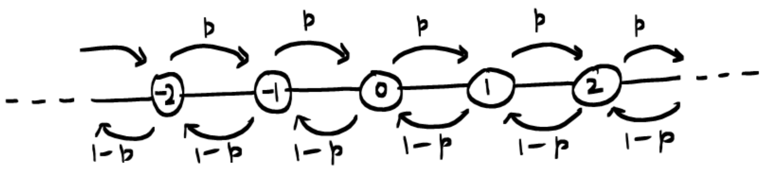
1.2. Random walks

1.2.1 Example(Unrestricted random walk on the line):
Consider a particle which moves one unit to the left with probability $1-p$ or to the right on the line with probability p. This is called **unrestricted random walk on the line**. Let X_n denote the position of the particle at time n. Then $S = \{0, \pm 1, \pm 2, ...\}$ and the markov chain has the transition graph and the transition matrix:



$$P = \begin{array}{c} \\ \vdots \\ -3 \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccccc} \dots & -3 & -2 & -1 & 0 & 1 & 2 & 3 \;\dots \end{array} \\ \left(\begin{array}{cccccccc} \dots & & \dots & \vdots & & & \dots & \dots \\ \dots & & \dots & 0 & \dots & & & \dots \\ \dots & (1-p) & 0 & p & \dots & & & \dots \\ \dots & 0 & (1-p) & 0 & \text{p} & 0 & . & \dots \\ \dots & \dots & 0 & (1-p) & 0 & p & 0 & \dots \\ \dots & & & 0 & (1-p) & p & 0 & \dots \\ \vdots & \dots & & \vdots & \vdots & & \dots & \dots \end{array} \right) \end{array}.$$

1.2.2 Random walk on the line with absorbing barriers:
We can also consider the random walk on the line in with state space $S = \{0, 1, 2, 3, ..., r\}$ and the condition that the walk ends if the particle reaches 0 or r. The states 0 and r are called **absorbing states** for the particle that reaches this state and is absorbed in it. It cannot leave the state. The transition graph and the transition probability matrix for this walk is given by



$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \\ \vdots \\ r \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & 3 & \dots & \dots & r \end{array} \\ \left(\begin{array}{ccccccc} 1 & 0 & 0 & \dots & \dots & \dots & 0 \\ (1-p) & 0 & p & 0 & \dots & \dots & 0 \\ 0 & (1-p) & 0 & p & 0 & \dots & 0 \\ \dots & \dots & & \dots & \dots & & \dots \\ 0 & \dots & & (1-p) & 0 & & p \\ 0 & 0 & \dots & \dots & \dots & 0 & 1 \end{array} \right) \end{array}$$

A typical illustration of this situation is when two players are gambling with total capital r rupees. The game ends when A looses all the money, i.e., 0 stage or B looses all the money, i.e., stage r for A, and X_n is the capital of A at n^{th} stage.

1.2.3 Random walk on the line with reflecting barriers:
Another variation of the previous example is the situation where two friends are gambling with a view to play longer. So they put the condition that every time a player loses his last rupee, the opponent returns it to him. Let X_n denote the capital of a player A at nth stage. If total money both the players have is $r+1$ rupees, then the state space for the system is $S = \{1, 2, 3, \dots, r\}$. To find the transition matrix, note that in the first row,

$$\begin{aligned} P(1,1) &= P\{X_{n+1} = 1 | X_n = 1\} \\ &= P\{\text{A's capital remains Rs.1 at next stage given that it was 1 at this stage.}\} \\ &= P\{\text{A has last rupee and loses. It will be returned}\} \\ &= (1-p). \end{aligned}$$

$$\begin{aligned} p(1,2) &= P\{\text{Capital of A becomes 2} | \text{ it is 1 now}\} \\ &= P\{\text{A wins}\} = p. \end{aligned}$$

$$p(1,j) = 0 \text{ for } j \geq 3.$$

For the i^{th} row, $1 < i < r$, and $1 \leq j \leq r$,

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

Thus, the transition matrix is given by:

$$
\begin{array}{c}
\\
1\\ \\ 2\\ 3\\ i\\ \vdots\\ r
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc}
1 & 2 & 3 & \cdots & \cdots & \cdots & i & \cdots & r
\end{array}\\
\begin{pmatrix}
(1-p) & p & 0 & \cdots & \cdots & \cdots & & 0\\
\\
(1-p) & 0 & \mathrm{p} & 0 & \cdots & \cdots & & 0\\
0 & (1-p) & 0 & p & 0 & \cdots & & 0\\
0 & \cdots & \cdots & (1-p) & 0 & p & \cdots & 0\\
\cdots & \cdots & \cdots & \cdots & & & \cdots & \cdots\\
0 & \cdots & \cdots & \cdots & & 0 & (1-p) & p
\end{pmatrix}
\end{array}.
$$

1.2.4. Birth and death chain

Let X_n denote the population of a living system at time $n, n \geq 1$. The state space for the system $\{X_n\}_{n\geq 1}$ is {0,1,2,...}. We assume that at any given stage n, if $X_n = x$, then the population increases to $x+1$, by a unit with probability p_x or decreases to $x-1$ with probability q_x, or can remain the same with probability r_x. Then,

$$
p(x,y) = \begin{cases} p_x \text{ if } & y = x+1, \\ q_x \text{ if } & y = x-1, \\ r_x \text{ if } & y = x, \\ 0 & \text{otherwise.} \end{cases}
$$

Clearly, this is a markov chain, called the **birth and death chain** and is a special case of random walks.

1.3. Queuing chains

Consider a counter where customers are being served at every unit of time. Let X_0 be the number of customers in the queue to be served when the counter opens and let ξ_n be the number of customers who arrive at the n^{th} unit of time. Then, X_{n+1} the number of customers waiting to be served at the beginning of $n+1^{th}$ time unit is

$$
X_{n+1} = \begin{cases} \xi_n & \text{if} \quad X_n = 0, \\ X_n + \xi_{n-1} & \text{if} \quad X_n \geq 1. \end{cases}
$$

The state space for the system $\{X_n\}_{n\geq 1}$ is $S = \{0, 1, 2, ...\}$. If $\{\xi_n\}_{n\geq 1}$ are independent random variables taking only nonnegative integral values, then $\{X_n\}_{n\geq 1}$ is a markov chain. In case $\{\xi_n\}_{n\geq 1}$ is also identically distributed with distribution function f, we can calculate the transition probabilities: for $x, y \in S$,

$$\begin{aligned} p(x,y) &= P\{X_{n+1} = y | X_n = x\} \\ &= \begin{cases} P\{X_{n+1} = y = \xi_n\} & \text{if } x = 0, \\ P\{X_{n+1} = y = \xi_{n-1} + X_n\} & \text{if } x \geq 1. \end{cases} \\ &= \begin{cases} P\{\xi_n = y\} & \text{if } x = 0, \\ P\{\xi_n = y - x + 1\} & \text{if } x > 1. \end{cases} \\ &= \begin{cases} f(y) & \text{if } x = 0 \\ f(y - x + 1) & \text{if } x > 1. \end{cases} \end{aligned}$$

1.4. Ehrenfest chain

Consider two isolated containers labeled as body A and body B, containing two different fluids. Let the total number of molecules of the two fluids, distributed in the containers A and B, be d, labeled as $\{1, 2, ..., d\}$. Let the observation be made on the number of the molecules in A. To start with, A has some number of molecules and B has some number of molecules. In the next stage, a number $1 \leq r \leq d$ is chosen at random and the molecule labeled r is removed from the body in which it was and is placed in the other body. This gives observation at second stage and so on. Clearly, X_n, which denotes the number of molecules that can be in A is $\{0, 1, 2, ..., d\}$ Thus, the state space is $S = \{0, 1, 2, ..d\}$. Let us find the transition probabilities $p(i, j) \;\; 0 \leq i, j \leq d$ of the system. When $i = 0$,

$$P(0, j) = P\{X_{n+1} = j | X_n = 0\},$$

i.e., A had no molecules at X_n. Therefore, clearly j can be only 1 at X_{n+1}.
Thus,

$$P(0, j) = \begin{cases} 0 & \text{if} \quad j \neq 0, \\ 1 & \text{if} \quad j = 1. \end{cases}$$

If A has to have d molecules, (i.e., all of them) at $(n+1)^{th}$ stage, then, at n^{th} stage, it should have only $d-1$ molecules. Thus, B has one molecule and that should be chosen and added to A. This can be done with probability 1. (Because B has only 1 molecule and it is to be selected at random.) Thus,

$$P(d, j) = \begin{cases} 1 & \text{if} \qquad j = d - 1, \\ 0 & \text{otherwise.} \end{cases}$$

For a fixed i, $1 < i < d$, let us look at $p(i, j)$, for $0 \leq j \leq d$. Since $p(i, j)$ is the probability that A will have j molecules, given that it had i molecules. Now if A had i molecules, then the only possibility for j is $i - 1$ or $i + 1$, (because the number of molecules in A at any next stage can increase or decrease). Thus, $p(i, j) = 0$, if $j \neq i+1$ or $i - 1$. If $j = i + 1$, i.e, A has to have $i + 1$ molecules, then B had $d - i$ molecules and one of the molecules for B should be selected and added to A. The probability for doing this is $\frac{d-i}{d}$. Thus,

$$p(i, i + 1) = \frac{d - i}{d} = 1 - \frac{i}{d} \text{ and } p(i, i - 1) = \frac{i}{d}.$$

Thus, the transition matrix for this markov chain is given by

$$
\begin{array}{c c}
 & \begin{array}{ccccccc} 0 & 1 & 2 & 3 & \cdots & \cdots & d \end{array} \\
\begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ d \end{array} &
\begin{pmatrix}
0 & 1 & 0 & \cdots & & \cdots & 0 \\
(1/d) & 0 & (1-1/d) & 0 & & \cdots & 0 \\
0 & (1/d) & 0 & (1-1/d) & 0 & \cdots & 0 \\
0 & \cdots & \vdots & 1/d & 0 & (1-1/d) & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

This model is called **Ehrenfest diffusion model**.

1.5. Some consequences of the markov property

Let $\{X_n\}_{n\geq 0}$ be a markov chain with state space S and transition probabilities $(p(i,j))$, $i,j \in S$.

1.5.1 Proposition:
Let $S_1, S_2, \ldots S_0$ be subsets of S. Then for any $n \geq 1$,

$$P\{X_n = j \mid X_{n-1} = i, X_{n-2} \in S_2, \ldots, X_0 \in S_0\} = p(i,j).$$

Proof:
The required property holds for elementary sets $S_k = i_k$, for $i_k \in S$ by the markov property:

$$P\{X_n = j|X_{n-1} = i, X_{n-2} = i_{n-2}, ..., X_0 = i_0\} = P\{X_n = j|X_{n-1} = i\}.$$

Since any subset A of S is a countable disjoint union of elementary sets and the required property follows from the property (iv) of conditional probability as in prologue. ■

1.5.2 Example:
let us compute $P\{X_3 = j|X_1 = i, X_0 = k\}$, $j,k \in S$. Using proposition 1.5.1, and markov property, we have

$$
\begin{aligned}
&P\{X_3 = j|X_1 = i, X_0 = k\} \\
&\quad = \sum_{r\in S} P\{X_3 = j|X_2 = r, X_1 = i, X_0 = k\}P\{X_2 = r|X_1 = i, X_0 = k\}. \\
&\quad = \sum_{r\in S} P\{X_3 = j|X_2 = r, X_1 = i\}P\{X_2 = r|X_1 = i\} \\
&\quad = P\{X_3 = j|X_1 = i\}.
\end{aligned}
$$

In fact above example can be extended to following:

1.5.3 Theorem:
For $n > n_s > n_{s_1} > ... > n_1 \geq 0$,

$$P\{X_n = j|X_{n_s} = i, X_{n_{s-1}} = i_{s-1}..., X_{n_1} = i_1\} = P\{X_n = j|X_{n_s} = i\}$$

Thus, for a markov chain, probability at n given past at $n_s > n_{s-1} > ... > n_1$, it depends only on the most recent past, i.e., n_s.

Thus, to every markov chain, we can associate a vector, distribution of the initial stage and a stochastic matrix whose entries give us the probabilities of moving from a state to another at the next stage. Here is the converse:

1.5.4 Theorem:
Given a stochastic matrix P and probability vector Π_0, there exists a markov chain $\{X_n\}_{n\geq 1}$ with Π_0, as initial distribution and P as transition probability matrix.
The interested reader may refer Theorem 8.1 of Billingsel[4] ■

1.5.4 Exercise
Show that $P\{X_0 = i_0 | X_1 = i-1, ..., X_n = i_n\} = P\{X_0 = i_0 | X_1 = x_1\}$.

Review Exercises

(1.1) Mark the following statements as True/False:
 (i) A Markov system can be in several states at one time.
 (ii) The $(1,3)$ entry in the transition matrix is the probability of going from state 1 to state 3 in two steps.
 (iii) The $(6,5)$ entry in the transition matrix is the probability of going from state 6 to state 5 in one step.
 (iv) The entries in each row of the transition matrix add to zero.
 (v) Let $\{X_n\}_{n\geq 1}$ be a sequence of independent identically distributed discrete random variables. Then it is a markov chain.
 (vi) If the state space is $S = \{s_1, s_2, \ldots, s_n\}$, then its transition matrix will have order n.

(1.2) Let $\{\xi_n\}_{n\geq 1}$ be a sequence of independent identically distributed discrete random variables. Define

$$X_n = \begin{cases} \xi_0 & \text{if } n = 0, \\ \xi_1 + \xi_2 + \ldots + \xi_n & \text{for } n \geq 1. \end{cases}$$

Show that $\{X_n\}_{n\geq 1}$ is a markov chain. Sketch its transition graph and compute the transition probabilities.

(1.3) Consider a person moving on a 4×4 grid. He can move only to the intersection points on the right or down, each with probability 1/2. If he starts his walk from the top left corner and $X_n,\ n \geq 1$ denotes his position after n steps. Show that $\{X_n\}_{n\geq 0}\}$ is a markov chain. Sketch its transition graph and compute the transition probability matrix. Also find the initial distribution vector.

(1.4) **Web surfing:**
Consider a person surfing the Internet, and each time he encounters a web page, he selects one of its hyperlinks at random (but uniformly). Let X_n denote the page where the person is after n selections (clicks). What do you think is the state space? Find the transition probability matrix.

(1.5) Let $\{X_n\}_{n\geq 0}$ be a markov chain with state space, initial probability distribution and transition matrix given by

$$S = \{1,2,3\},\ \Pi_0 = (1/3/1/3, 1/3),\ P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

Define

$$Y_n = \begin{cases} 0 & \text{if} X_n = 1, \\ 1 & \text{otherwise.} \end{cases}$$

Show that $\{Y_n\}_{n\geq 0}$ is not a markov chain. Thus, *function of a markov chain need not be a markov chain.*

(1.6) Let $\{X_n\}_{n\geq 0}$ be a markov chain with transition matrix P. Define

$$Y_n = X_{2n} \text{ for every } n \geq 0.$$

Show that $\{Y_n\}_{n\geq 0}$ is a markov chain with transition matrix P^2. What happens if Y_n is defined as

$$Y_n = X_{n_k} \text{ for every } n \geq 0?$$

Chapter 2

Calculation of higher order probabilities

2.1. Distribution of X_n and other joint distributions

Consider a markov chain $\{X_n\}_{n\geq 1}$ with initial vector Π_0, and transition probabilities matrix $P = [p(i,j)], i \times j$. We want to find the probability that after n steps, the system will be in a given state, say $j \in S$? For a matrix A, its n-fold product with itself will be denoted by A^n.

2.1.1 Theorem:

(i) The joint distribution of $X_0, X_1, X_2, \ldots, X_n$, is given by

$$P\{X_0 = i_0, X_1 = i_1, ..., X_n = i_n\} = p(i_{n-1}, i_n)p(i_{n-2}, i_{n-1}) \ldots p(i_0, i_1)\Pi_0(i_0).$$

(ii) The distribution of X_n, $P\{X_n = j\}$, is given by the j^{th} component of the vector $\Pi_0 P^n$.

(iii) For every $n, m \geq 0$,

$$P\{X_n = j \,|\, X_0 = i\} = P\{X_{n+m} = j \,|\, X_m = i\} = p^n(i,j),$$

where $p^n(i,j)$ is the ij^{th} term of the matrix P^n.

Proof:

(i) Using the chain rule for conditional probability,

$$\begin{aligned}
&P\{X_0 = i_0, X_1 = i_1, ..., X_n = i_n\} \\
&= P\{X_n = i_n | X_{n-1} = i_{n-1}\} P\{X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}, \ldots, X_0 = i_0\} \\
&\qquad \ldots P\{X_1 = i_1 | X_0 = i_0\} P\{X_0 = i_0\} \\
&= P\{X_n = i_n | X_{n-1} = i_{n-1}\}\, P\{X_{n-1} = i_{n-1} | X_{n-2} = i_{n-2}\}, \ldots, \\
&\qquad \ldots P\{X_1 = i_1 | X_0 = i_0\} P\{X_0 = i_0\} \\
&= p(i_{n-1}, i_n)p(i_{n-2}, i_{n-1}) \ldots p(i_0, i_1)\Pi_0(i_0).
\end{aligned}$$

(ii) Let Y be a random variable with values in S and distribution $P\{Y = i\} = \lambda_i,\ i \in S$. Then using the chain rule for conditional probability,

$$\begin{aligned} P\{X_n = j\} &= \sum_{i_0 \in S} P\{Y = i_0,\ X_n = j\} \\ &= \sum_{i_0 \in S} \sum_{i_1 \in S} \cdots \sum_{i_{n-1} \in S} P\{Y = i_0,\ X_{i_1} = i_1,\ \ldots,\ X_{i_{n-1}} = i_{n-1},\ X_n = j\} \\ &= \sum_{i_0 \in S} \sum_{i_1 \in S} \cdots \sum_{i_{n-1} \in S} P\{Y = i_0\}\, P\{X_{i_1} = i_1 | X_{i_1 - 1} = i_1 - 1\} \ldots \\ &\qquad \ldots,\ P\{X_n = j | X_{i_{n-1}} = i_{n-1}\} \quad (2.1) \\ &= \sum_{i_0 \in S} \sum_{i_1 \in S} \cdots \sum_{i_{n-1} \in S} \lambda_i\, p(i_0, i_1) \cdots p(i_{n-1}.j) \quad (2.2) \end{aligned}$$

Thus for $Y = X_0$, we have

$$\begin{aligned} P\{X_n = j\} &= \sum_{i_0 \in S} \sum_{i_1 \in S} \cdots \sum_{i_{n-1} \in S} \Pi_0(i)\, p(i_0, i_1) \cdots p(i_{n-1}, j) \\ &= j^{th} \text{ element of the vector } \Pi_0\, P^n. \end{aligned}$$

(iii) Once again, using the markov property and the chain rule for conditional probability,

$$\begin{aligned} &P\{X_{n+m} = j \,|\, X_m = i\}\, P\{X_m = i\} \\ &\quad = P\{X_{n+m} = j,\ X_m = i\} \\ &\quad = \sum_{i_m \in S} \sum_{i_{m+1} \in S} \cdots \sum_{i_{m+n-1} \in S} P\{X_m = i_m,\ X_{m+1} = i_{m+1}, \\ &\qquad\qquad \ldots, X_{i_{n+m-1}} = i_{n+m-1},\ X_{n+m} = j\} \\ &\quad = \sum_{i_m \in S} \sum_{i_{m+1} \in S} \cdots \sum_{i_{m+n-1} \in S} P\{X_m = i\}\, P\{X_{m+1} = i_{m+1} | X_m = i\} \ldots \\ &\qquad\qquad \ldots, P\{X_{i_{n+m-1}} = i_{n+m-1} |, X_{n+m} = j\} \\ &\quad = \sum_{i_m \in S} \sum_{i_{m+1} \in S} \cdots \sum_{i_{m+n-1} \in S} P\{X_m = i\}\, p(i, i_{m+1}) \ldots\, p(i_{n+m-1}, j). \end{aligned}$$

Thus

$$P\{X_n = j \,|\, X_0 = i\} = P\{X_{n+m} = j \,|\, X_m = i\} = p^n(i,j),$$

where $p^n(i,j)$ is the ij^{th} term of the matrix P^n. ■

2.1.2 Definition:
Let $\{X_n\}_{n \geq 1}$ be a markov chain with initial vector Π_0, and transition probabilities matrix $P = (p(i,j)),\ i, j \in S$.

(i) For $n \geq 1$, and $j \in S$, $p_n(j) = P\{X_n = j\}$ is called the **distribution of** X_n.

(ii) For $n \geq 1$, $p_n(i,j)$ is called the n^{th} **stage transition probabilities**.

Above theorem gives us the probability of the system in a state at the n^{th} stage and the probability of the event that the system will move in n stages from a state i to a state j. And these can be computed if we know the initial distribution and powers of the transition matrix. Thus, it is important to compute the matrix P^n, P being the transition matrix. For large n, this is difficult to compute. Let us look at some examples.

2.1.3 Exercise:
Show that the joint distribution of $X_m, X_{m+1}, \ldots, X_{m+n}$ is given by

$$p(i_{n-1}, i_n) p(i_{n-2}, i_{n-1}) \cdots p(i_{m+1}, i_{m+2}) P\{X_{m+1} = i_{m+1}\}$$

Also write the joint distribution of any finite $X_{n_1}, X_{n_2}, ..., X_{n_r}$. for $n_1 < n_2, ... < n_r$.

2.1.4 Example:
Consider a markov chain $\{X_n\}_{n\geq 1}$ with the special situation where all the $X_{n's}$ are independent. Let us compute P^n, where P is the transition probability matrix. Because $X_n's$ are independent,

$$p(i,j) = P\{X_{n+1} = j | X_n = i\} = P\{X_{n+1} = j\}$$

for all j, i and for all n. Thus, each row of P is identical. By theorem 2.1.1(iii), for all i,

$$\begin{aligned} p^n(i,j) &= P\{X_{n+m} = j | X_m = i\} \\ &= P\{X_n = j | X_0 = i\} \\ &= P\{X_n = j\} = p(i,j). \end{aligned}$$

Therefore each $P^n(i,j) = p(i,j)$, i.e., $P^n = P$.

2.1.5 Example:
Let us consider the markov chain with two states $S = \{0, 1\}$ and transition matrix

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

Let $\Pi_0(0)$, $\Pi_0(1)$ be initial distributions. The knowledge of P and $\Pi_0(0), \Pi_0(1)$ helps us to answer various questions. For example, to compute the distribution of X_n, using the formula of conditional probability: $P(A|B)\,P(B) = P(A \cap B))$, we have for every $n \geq 0$,

$$\begin{aligned} P\{X_{n+1} = 0\} &= P\{X_{n+1} = 0, X_n = 0\} + P\{X_{n+1} = 0, X_n = 1\} \\ &= P\{X_{n+1} = 0 | X_n = 0\}\, P\{X_n = 0\} \\ &\qquad + P\{X_{n+1} = 0 | X_n = 1\}\, P\{X_n = 1\} \\ &= (1-p)P\{X_n = 0\} + qP\{X_n = 1\} \\ &= (1-p)P\{X_n = 0\} + q(1 - P\{X_n = 0\}) \\ &= (1-p-q)P\{X_n = 0\} + q. \end{aligned}$$

Thus, for $n = 0, 1, 2, ...,$

$$\begin{aligned} P\{X_1 = 0\} &= (1-p-q)\Pi_0(0) + q \\ P\{X_2 = 0\} &= (1-p-q)P\{X_1 = 0\} + q \\ &= (1-p-q)[q + (1-p-q)\Pi_0(0)] + q \\ &= (1-p-q)^2\Pi_0(0) + q(1-p-q) + q. \\ &= (1-p-q)^2\Pi_0(0) + q\sum_{j=0}^{1}(1-p-q)^j. \\ \cdots\ \cdots\ &\cdots\ \cdots\ \cdots \\ P\{X_n = 0\} &= (1-p-q)^n + q\sum_{j=0}^{n-1}(1-p-q)^j. \end{aligned}$$

$$\begin{aligned}
P^n(0,0) &= P\{X_n = 0|X_0 = 0\} \\
&= P\{X_n = 0\} \\
&= \left(\frac{q}{p+q}\right) + (1-p-q)^n \left[1 - \frac{q}{p+q}\right] \\
&= \left(\frac{q}{p+q}\right) + (1-p-q)^n \left(\frac{p}{p+q}\right).
\end{aligned}$$

Then, using the fact that $P\{X_0 = 0\} = 1$,

$$\begin{aligned}
P^n(0,1) &= P\{X_n = 1|X_0 = 0\} \\
&= P\{X_n = 1\} \\
&= \left(\frac{p}{p+q}\right) + (1-p-q)^n \left[0 - \frac{p}{p+q}\right] \\
&= \left(\frac{p}{p+q}\right) - (1-p-q)^n \left(\frac{p}{p+q}\right).
\end{aligned}$$

Then,

$$\begin{aligned}
P^n(1,0) &= P\{X_n = 0|X_0 = 1\} \\
&= P\{X_n = 0\} \\
&= \left(\frac{q}{p+q}\right) + (1-p-q)^n \left[\Pi_0(0) - \frac{q}{p+q}\right] \\
&= \left(\frac{q}{p+q}\right) + (1-p-q)^n \left[0 - \frac{q}{p+q}\right] \\
&= \left(\frac{q}{p+q}\right) - (1-p-q)^n \left(\frac{q}{p+q}\right).
\end{aligned}$$

And

$$P^n(1,1) = \left(\frac{1}{p+q}\right) + (1-p-q)^n \left(\frac{q}{p+q}\right)$$

Therefore,

$$P^n = \left(\frac{1}{p+q}\right) \begin{bmatrix} q & p \\ q & p \end{bmatrix} + \left(\frac{(1-p-q)^n}{p+q}\right) \begin{bmatrix} p & -p \\ -q & q \end{bmatrix}.$$

2.1.6 Exercise:
Consider the (random walk) markov chain as in example 1.1.10.

(i) If $p = q = 0$, what can be said about the machine?

(ii) If $p, q > 0$, show that

$$P\{X_n = 0\} = \frac{q}{p+q} + (1-p-q)^n \left[\Pi_0(0) - \frac{q}{p+q}\right]$$

and

$$P\{X_n = 1\} = \frac{p}{p+q} + (1-p-q)^n \left[\Pi_0(1) - \frac{p}{p+q}\right].$$

(iv) Find conditions on $\Pi_0(0)$ and $\Pi_0(1)$ such that distribution of X_n is independent of n.

(v) Compute the following:

$$P\{X_0 = 0, X_1 = 1, X_2 = 0\}$$

(vi) Can one compute joint distribution of X_{n+2}, X_{n+1}, X_n?

2.1.7 Note (In case P is diagonalizable:
As we observed earlier, it is not easy to compute P^n for a matrix P, even when it is finite. However, in the case P is diagonalized (see Appendix for more details), it is easy: let there exist an invertible matrix U such that $U\,P\,U^{-1} = D$, where D is a diagonal matrix. Then $P^n = U\,D^n\,U^{-1}$, and D^n is easy to compute.In this case, we can compute the elements of P^n. Let the state space has M elements and P be diagonalizable with diagonal elements of D be $\lambda_1, \lambda_2, \ldots, \lambda_M$, these are the eigenvalues of P. To find $p_n(i,j)$:

(i) Compute the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$, of P by solving the characteristic equation.

(ii) If all the eigenvalues are distinct, then for all n, p^n_{ij} has the form

$$p^n_{ij} = a_i \lambda_1^n + \ldots + a_M \lambda_M^n,$$

for some constants $a_i, \ldots, a_M$, depending upon i and j. These can be found by solving system of linear equations.

2.1.8 Example:
Let for a markov chain, the transition matrix is

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

and let us try to find a general formula for p^n_{11}. We first compute the eigenvalues of P by solving

$$det(P - \lambda I) = \begin{vmatrix} 0-\lambda & 1 & 0 \\ 0 & 1/2-\lambda & 1/2 \\ 1/2 & 0 & 1/2-\lambda \end{vmatrix} = 0.$$

This gives (complex) eigenvalues 1, $\pm(i/2)$. Thus, for some invertible matrix U,

$$P = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & i/2 & 0 \\ 0 & 0 & -i/2 \end{pmatrix} U^{-1},$$

and hence

$$P^n = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & (i/2)^n & 0 \\ 0 & 0 & (-i/2)^n \end{pmatrix} U^{-1}.$$

In fact U can be explicitly written in terms of the eigenvectors. In another way, above equation implies that for scalars a, b, c,

$$p^n_{11} = a + b(i/2)^n + c(-i/2)^n.$$

In order to have real solutions, we compare the real and imaginary parts of the above and have for all $n \geq 0$,

$$p^n_{11} = a + b(i/2)^n \cos(n\pi/2) + c(i/2)^n \sin(npi/2).$$

In particular for $n = 0, 1, 2$, we have

$$\begin{aligned} 1 &= p^0 11 = a + b \\ 0 &= p^1 11 = a + 1/2c \\ 0 &= p^2 11 = a - 1/4b. \end{aligned}$$

A solution of the above system is given by $a = 1/5, b = 4/5, c = -2/5$, and hence

$$p_{11}^n = 1/5 + (1/2)^n (4/5 \cos(n\pi/2) - 2/5(i/2)^n \sin(n\pi/2).$$

2.2. Kolmogorov-Chapman equation

We saw that given a markov chain $\{X_n\}_{n\geq 1}$ with state space S, initial distribution Π_0 and transition matrix P, we can calculate the distribution of X_n and other joint distributions. Thus, if we write Π_n for the distribution of X_n, i.e., if $\Pi_n(j) = P\{X_n = j\}$, then,

$$\Pi_n(j) = \sum_{k\in s} \Pi_0(k) p_{kj}^n.$$

or symbolically,

$$\Pi_n = \Pi_0 P^n.$$

Now we can write the joint distribution of $X_{n+1}, ...X_{m+n}$ as

$$P\{X_{m+t} = i_t,\ 0 \leq t \leq n\} = \Pi_{m+1}(i_1) p_{i_1, i_2}, ...p_{i_{n+1}, i_n}.$$

Entries of P^n are called the n^{th} **step transition probabilities**. Thus, the knowledge about the markov chain is contained in Π_0 and the matrix P. As noted earlier P is a matrix (may be an infinite) such that sum of each row is 1, i.e., a **stochastic matrix.** For consistency, we define $P^0 = Id$. The following is easy to show:

2.2.1 Theorem:
For $n, m \geq 0$ and $(i, j \in S$,

$$p^{n+m}(i,j) = \sum_{r\in S} p^n(i,r)\, p^m(r,j),$$

In matrix multiplication this is just

$$P^{n+m} = P^n\, P^m.$$

This is called the **Kolmogorov Chapman** equation.
Proof:
Using the property (v) conditional probability

$$\begin{aligned} p^{n+m}(i,j) &= P\{X_{n+m} = j \,|\, X_0 = i\} \\ &= \sum_{r\in S} P\{X_n = r, \,|\, X_0 = i\}\, P\{X_{n+m=j} \,|\, X_n = r, X_0 = i\} \\ &= \sum_{r\in S} p^n(i,r)\, P\{X_{n+m} = j \,|\, X_n = r, X_0 = i\} \\ &= \sum_{r\in S} p^n(i,r) p_m(r,j), \end{aligned}$$

The last equality follows from the fact that

$$P\{X_{n+m} = j \,|\, X_n = r, X_0 = i\} = P\{X_{n+m} = j \,|\, X_n = r\} = p^m(r,j),$$

as observed in theorem 1.5.5. ■

2.2.2 Example:
Consider the unrestricted random walk on the line, as in example 1.2.1, with probability p to move forward and $1 - p$ to come back. Then,

$$p^{2n+1}(0,0) = 0.$$

as only in even steps it can come back to starting point. And,

$$p^{2n}(0,0) = \begin{pmatrix} 2n \\ n \end{pmatrix} p^n (1-p)^{2n-n},$$

as there will be n moves to right and n back. Thus,

$$p^{2n}(0,0) = \begin{pmatrix} 2n \\ n \end{pmatrix} (pq)^n.$$

In fact,this is true for every diagonal entry. Other entries are difficult to compute. Note that

$$\sum_{n=1}^{\infty} p_{00}^{in} = \Sigma_{n=0}^{\infty} \begin{pmatrix} 2n \\ n \end{pmatrix} (pq)^n.$$

Using sterling's approximation,

$$n!\ \sqrt{2\pi n}^{n+1/2} e^{-n},$$

we have

$$P_{00}^{2n} = \Sigma_{n=0}^{\infty} \frac{(pq)^n 2^n}{\sqrt{n\pi}}$$

which is convergent if $pq < 1$, and divergent otherwise. Thus, 0 is transient if $p \neq q$, and recurrent if $p = q = 1/2$.

2.2.3 Example:
Consider the markov chain of exercise 1.3, with state space $S = \{1,2,3,4\}$, initial distribution $(1,0,0,0)$, and transition matrix

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

Then,

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

and

$$\Pi_0 P^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \end{pmatrix}.$$

Thus, if we want to find the probability that the walker will be in state 3 in two steps, then it is

$$\Pi_2(3) = (\Pi_0 P^2)(3) = 0.$$

Exercises

(2.1) Consider the markov chain of example 2.2.3. Show that

$$\Pi_n = \begin{cases} (0,\ 1/2,\ 0,\ 1/2) & \text{for n=1, 3, 5, } \ldots \\ \\ (1/2,0,1/2,0) & \text{for n= 2, 4, 6, } \ldots \end{cases}$$

(2.2) Let $\{X_n\}_{n\geq 0}$ be a markov chain with state space, initial probability distribution and transition matrix given by

$$S = \{1,2\},\ \Pi_0 = (1,0),\ P = \begin{pmatrix} 3/4 & 3/4 \\ 1/4 & 1/4 \end{pmatrix}.$$

Show that

$$\Pi_n = \left(\frac{1}{2}(1+2^{-n}), \frac{1}{2}(1+2^{-n})\right) \text{ for every } n.$$

(2.3) Consider the two state markov chain $\{X_n\}_{n\geq 0}$ with $\Pi_0 = (1,0)$, and transition matrix

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

Using the the facts that P is stochastic and the relation $P^{n+1} = P^n\, P$, deuce that

$$\begin{aligned} p^(n+1)(1,1) &= p^n(1,2)q + p^n(1,1)(1-p) \\ P^n(1,1) + p^n(1,2) &= 1, \end{aligned}$$

and hence,for all $n > 0$,

$$p^(n+1)(1,1) = p^n(1,1)(1-p-q) + q.$$

Show that this has a unique solution

$$p^n(1,1) = \begin{cases} \dfrac{q}{p+q} + \dfrac{p}{p+q}(1-p-q)^n & \text{for } p+q > 0 \\ \\ 1 & \text{for } p+q < 0. \end{cases}$$

Chapter 3

Classification of states

Let $\{X_n\}_{n\geq 0}$ be a Markov chain with state space S, initial distribution Π_0 and transition probability matrix P. We will denote the ij^{th} element of $p_n(i,j)$ also by p^n_{ij}. We start looking at the possibility of moving from one state to another.

3.1. Closed subsets and irreducible subsets

3.1.1 Definition

(i) We say a state j is **reachable** from a state i (or i **leads** to j or j is **approachable** from i,) if there exists some $n \geq 0$, such that $p^n_{ij} > 0$. We denote this by $i \to j$.
In other words, i leads to j in n steps with positive probability.

(ii) A subset C of the state space is said to be **closed** if no state from C leads to a state outside C.
Thus C is closed is same as for every $i \in C,\ j \notin C\ p^n_{ij} = 0\ \forall n \geq 0$. This means once the chain enters the set C it will never leave it.

(iii) A state j is called an **absorbing state** if the singleton set $\{j\}$ is a closed set.

3.1.2 Proposition:

(i) If $i \to j$ and $j \to k$, then $i \to k$.

(ii) A state j is reachable from a state i iff $p_{ii_1}p_{i_1i_2}\dots p_{i_{n-1}j} > 0$, for some $i_1, i_2, \dots, i_{n-1} \in S$.

(iii) $C \subset S$ is closed iff $\forall i \in C, j \notin C, p_{ij} = 0$.

(iv) The state space S is closed and for $i \in S$, the set $\{i\}$ is closed if $p_{ii} = 1$.

Proof:

(i) Follows from the fact that

$$p^{n+m}_{ik} = \sum_{r\in S} p^n_{ir}\, p^m_{rk} > p^n_{1j}p^m_{jk} > 0 \text{ for some } n, m > 0.$$

(ii) Follows from the equality

$$p^n_{ij} = \sum_{i_1,\dots i_{n-1}} p_{ii_1}p_{i_1i_2}\dots p_{i_{n-1}j}.$$

(iii) Clearly, $p^n_{ij} = 0 \;\forall\, n$ implies that $p_{ij} = 0$. Conversely, let for all $i \in C, j \notin C$, $p_{ij} = 0$. Then $p_{lk} = 0$ for $l \in C,\; k \notin C$, and $p_{k,l} = 0$ for $l \notin C, r \in C$. Thus, for all $r \in C$ and $k \notin C$,

$$p^2_{rk} = \sum_{l \in S} p_{rl}\, p_{lk} = \sum_{l \notin S} p_{rl} p_{lk} = 0.$$

Proceeding similarly, $p^n_{rk} = 0$ for all $n \geq 1$.

(iv) Proof of (iv) is obvious.

■

3.1.3 Definition:
A subset C of S is called **irreducible** if any two states in C lead to one another.

Let us look at some examples.

3.1.4 Example:
Consider a markov chain with transition matrix:

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}
\begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\
\left(\begin{array}{cccccc}
1 & 0 & 0 & 0 & 0 & 0 \\
1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\
0 & 1/5 & 2/5 & 1/5 & 0 & 1/5 \\
0 & 0 & 0 & 1/6 & 1/3 & 1/2 \\
0 & 0 & 0 & 1/2 & 0 & 1/2 \\
0 & 0 & 0 & 1/4 & 0 & 3/4
\end{array} \right) \end{array}.$$

We first look at which state leads to which state. Whenever, $i \to j$, we put a $*$ in the matrix entry. Note, $p_{ij} > 0$ will give a $*$ at ij^{th} entry, but $p_{ij} = 0$ need not give 0 in the matrix. For example, $p_{13} = 0$, but $1 \to 2 \to 3$, so p_{13} is replaced by $*$. For the above matrix, we have

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}
\begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\
\left(\begin{array}{cccccc}
* & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * \\
* & * & * & * & * & * \\
0 & 0 & 0 & * & * & * \\
0 & 0 & 0 & * & * & * \\
0 & 0 & 0 & * & * & *
\end{array} \right) \end{array}
\left| \begin{array}{lllllll}
1 \to & 2 \to & 3 \to & 4 \to & 5 & & \\
2 \to & 1 \to & 0, & 2 \to & 3 \to & 4 \to & 5 \\
3 \to & 4, & 3 \to & 5, & 3 \to & 3 & \\
4 \to & 3 \to & 4, & 4 \to & 5 & & \\
5 \to & 3 \to & 4, & 5 \to & 5 & &
\end{array} \right.$$

Clearly, every single state i is a closed set if $p_{ii} = 1$. For example in our case, $\{0\}$ is a closed . The set S is closed by definition for there is no state outside S. Thus, $\{0, 1, 2, 3, 4, 5\}$ is closed. A look at the matrix of communication tells us that the set $\{3, 4, 5\}$ is closed because none of $3, 4, 5$, lead to $0, 1, 2$. For example $\{1\}$ is not closed because $1 \to 2$. In fact, there is no other closed sets. The set $\{3, 4, 5\}$ is also irreducible.

3.1.5 Note (importance of closed irreducible sets):
Why one should bother about closed subsets of the state space? To find the answer, let us look at the above example again. Let us take a proper closed set, say $C = \{3, 4, 5\}$. Now if we remove the rows and columns corresponding to states 1 and 2 from the transition matrix, we get the sub-matrix

$$\begin{array}{c} \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1/6 & 1/3 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/4 & 0 & 3/4 \end{pmatrix} \end{array}$$

which has the property that sum of each row is 1. In fact, if we take P^2 and delete rows and columns not in C, and write it as $(P^2)_C$, then it is easy to check it is nothing by $(P_C)^2$. For in P^2 note for $i \in C$,

$$P^2_{ij} = 0 \text{ if } j \notin C.$$

Therefore,

$$1 = \sum_{j \in S} P^2_{ij} = \sum_{j \in C} p^2_{ij}.$$

Thus, $(P_C)^2$ is a stochastic matrix. Also, for $i, j \in C.$,

$$p^2_{ij} = \sum_{r \in S} p_{ir} p_{rj} = \sum_{r \in C} p_{ir} p_{rj} = \begin{cases} (ij)^{th} \text{entry of } P^2_C. \\ \\ 0 \text{ if } j \notin C. \end{cases}$$

because C is closed, and $p_{ir} = 0,$ for $r \notin C$. In general, $(P^n)_C = (P_C)^n$. Hence, one can consider the chain with state space C and analyze it. This reduces the number of states.

3.1.6 Definition:
Two states i and j are said to **communicate** if either is accessible from the other, i.e., $p^n_{ij} > 0$ and $p^m_{ji} > 0$ for some $m, n \geq 1$. In this case we write $i \leftrightarrow j$.

3.1.7 Proposition:

(i) For $i, j \in S$, let us say $i \sim j$ iff $i \leftrightarrow j$. Then $\sim$ is an equivalence relation on S

(ii) Each equivalence class, called **communicating class** *has no proper closed subsets.*

Proof:

(i) That $i \leftrightarrow i$ follows from the fact that $P^0 = Id$, and hence $p^0_{ii} = 1$. Obviously, it is symmetric, and transitivity follows from proposition 3.1.2(i).

(ii) Let C be an equivalence class. If A is a proper subset of C, let $j \in C \setminus A$. Let $i \in A$. Then $i \leftrightarrow j$ implying that $j \notin A$ is accessible from i A. Hence, A is not closed. ■

3.1.8 Note:
A communicating class need not be closed. It may be possible to start from one communicating class and enter another with positive probability. For example consider a markov chain with transition matrix

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

The communicating classes are $\{1,2,3\}, \{4\}, \{5,6\}$. Clearly, $3 \to 4$, but $4 \not\to 3$. Only $\{5,6\}$ is a closed subset.

3.1.9 Example:
Consider a markov chain with five states $\{1,2,3,4,5\}$ and with transition matrix

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

States 1 and 2 communicate with each other and with no other state. Similarly, states $3, 4, 5$ communicate among themselves only. Thus, the state space divides into two closed irreducible sets $\{1,2\}$ and $\{3,4,5\}$. For the sake of all practical purposes, analyzing the given markov chain is same as analyzing two smaller two chains with smaller state space, with transposition matrices

$$P_1 = \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 1/4 \end{pmatrix}, \; P_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

3.1.10 Theorem:
A set $C \subseteq S$ is irreducible if every state in C communicates with every other state in it.

Proof:
Suppose, C is irreducible. For $j \in C$, define

$$C_j = \{i \in C \,|\, p_{ij}^n = 0 \forall n \geq 0\}.$$

We claim that C_j is a closed set. To see this, let $k \notin C_j$. Then there exists some m such that $p_{kj}^m > 0$. Now if i is such that $p_{ik} > 0$, then

$$p_{ij}^{m+1} = \sum_{l \in S} p_{il} p_{lj}^m > p_{ik} p_{kj}^m > 0,$$

not possible if $i \in C_j$. Thus, $p_{ik} = 0$, for every $i \in C_j$ and $k \notin C_j$, implying that c_j is closed. In fact, C being irreducible, this implies that $C = C_j$, and hence any two states in C communicate with each other. Conversely, let $i \leftrightarrow j$ for all $i, j \in C$ and $A \subseteq C$ be a closed set. Then, for $j \in A$ and $i \in C$, since $i \leftrightarrow j$, we have $jinC$, and hence $A = C$, i.e., C is irreducible.

■

In view of note 3.1.5, one would like to partition the state space into irreducible subsets.

Exercises

(3.1) Let the transition matrix of a markov chain be given by

$$\begin{pmatrix} 1/2 & 0 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/3 & 0 & 1/6 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Write the transition graph and find all the disjoint closed subsets of the state space $S = \{1, 2, 3, 4, 5\}$.

(3.2) Consider the markov chain in example 1.2.2, Random walk with absorbing barriers. Show that the state space splits into three irreducible sets. Is it possible to go from one set to other?

(3.3) For the queuing markov chain in example in section 1.3, write the transition matrix and if $f(k) > 0$ for every k, deuce that S itself is irreducible.

(3.4) Let a markov chain have transition matrix

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

Show that it is an irreducible chain.

3.2. Periodic and aperiodic chains

Throughout this section $\{X_n\}_{n\geq 0}$ will be a markov chain with state space S, initial probability Π_0 and transition matrix P.

3.2.1 Definition:
A state j is said to have **period** d, if $p^n_{jj} > 0$ implies d divides n and d is the largest such integer.

In other words, 'period of j is the greatest common divisor of the numbers $\{n \geq 1 | p^n_{ij} > 0\}$.

A state j has period d, means that $p^n_{jj} = 0$ unless $n = md$ for some $m \geq 1$, and d is the greatest positive integer with this property. Thus, j has period d means the chain may come back to j at time points md only. But, it may never come back to the state j.

3.2.2 Example:
Consider a markov chain with transition matrix

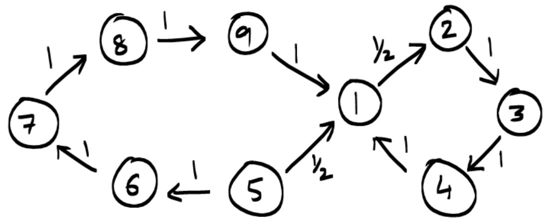
$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}.$$

Now $p_{jj} = 0 \ \forall\ j$. Therefore, period of each state is > 1. In fact, each state has period 2 for $p^2_{jj} > 0$ and $p^{(\text{odd})}_{jj} = 0$. But $\{3, 4\}$ form a closed set and once a particle goes to the set $\{3, 4\}$ (say from state 2,) it will never come out and return to 2.

3.2.3 Definition:
A state j is called **aperiodic state** if j has period 1. The chain is called **aperiodic chain** if every state in the chain has period 1.

In an aperiodic chain, if $i \leftrightarrow j$, then $p^n_{ij} > 0$ for all sufficiently large n, i.e., it is possible for the chain to come back to any state at any time.

3.2.4 Example:
Consider the transition graph of a markov chain with transition graph



Note that the starting in state 1, it can be revisited at stages $4, 6, 10, 8, ...$ Thus the state 1 has period 2.

3.2.5 Example (Birth and death chain):
Consider a markov chain on $S = \{0, 1, 2, \ldots\}$. Starting at i the chain can stay at i or move to $i-1$ or $i+1$ with probabilities

$$p(i,j) = \begin{cases} q_i & if j = i-1 \\ r_i & j = x, \\ p_i & j = i+1, \\ 0 & \text{otherwise.} \end{cases}$$

Saying that that it is an irreducible chain is same as saying that $p_i > 0$ for all $i \geq 0$, and $q_i > 0$ for all $i > 0$. It will be aperiodic if $r_i > 0$, see exercise (3.5) below. If $r_i = 0$ for all i, then the chain can return to i only after even number of steps. Thus the period of the chain can only be a multiple of 2. Since $p_{00}^2 = p_0 q_1 > 0$, every state has period 2.

3.2.6 Theorem:
If two states communicate with each other, then they have same periods.

Proof:
Let $d_i =$ period of i and $d_j =$ period j. It is enough to show that d_i divides r if $p_{jj}^r > 0$. $i \leftrightarrow j$ implies there exist n, m such that $p_{ij}^m > 0$ and $p_{ji}^n > 0$. By Kolmogorov-Chapman equations, for every $r \geq 0$,

$$p_{ii}^{m+r+n} \geq p_{ij}^m p_{jj}^r p_{ji}^n > 0.$$

This implies d_i divides $m+r+n$ for every $r \geq 0$, with $p_{jj}^r > 0$, In particular, with $r = 0$, as $p_{jj}^0 > 1$ implies that d_i divides $m+n$, and hence d_i divides $r = (m+r+n) - (m+n)$. Hence, $d_i \geq d_j$. Similarly,$d_i \leq d_j$. ■

Exercises

(3.5) Show that if a markov chain is irreducible and $p_{ii} > 0$ for some state i, then it is aperiodic.

(3.6) Show that the queuing chain of example 1.3 is aperiodic.

3.3. Visiting a state: transient and recurrent states

Let $i, j \in S$ be fixed. Let us consider the probability of the event that for some $n \geq 1$, the system will visit the state j given that it starts in the state i. Let

$$f_{ij}^n := P\{X_n = j, X_k \neq j, 1 \leq k \leq n-1 | X_0 = i\}\ , n \geq 1,$$

i.e., f_{ij}^n is the probability of first visit to state j starting at i in n steps. We are interested in computing

$$f_{ij} := \sum_{n-1}^{\infty} f_{ij}^n,$$

in terms of the transition probabilities. Let us first compute f_{ii}^n for any n. We define $f_{ii}^0 = 0$ for all i. It is the probability of eventual visit to state j starting from state i. Note that, $f_{ii}^1 = p_{ii}$ and f_{ij} is the probability that the system has a visit to j starting at i in some finite time.

3.3.1 Proposition:

(i) $f_{ij}^1 = p_{ij}$.

(ii) $f_{ij}^{n+1} = \sum_{r \neq j} p_{ir} f_{rj}^n$.

(iii) $p_{ij}^n = \sum_{k=0}^{n} p_{jj}^{n-k} f_{ij}^k$.

(iv) $p_{ii}^n = \sum_{k=1}^{n} {p_{ii}}^{n-k} {f_{ii}}^k$.

(v) $P\{$*system visits state* j *at least* 2 *times* $|X_0 = i\} = f_{ij} f_{jj}$.
More generally,

$$P\{\textit{system has } m \textit{ visits and at least to state } j\ |X_0 = i\} = f_{ij} f_{jj}^{(m-1)}.$$

Proof:

(i) Obvious.

(ii)

$$\begin{aligned} f_{ij}^{n+1} &= \sum_{r \neq j} P\{\text{from } \ i \text{ to } r \text{ in one step}\}\ P\{\text{first visit in } n^{th} \text{ step from } \ r \text{ to } j\} \\ &= \sum_{r \neq j} p_{ir} f_{rj}^n. \end{aligned}$$

(iii) Note that

$$\begin{aligned} p_{ij}^n &= \sum_{m=1}^{n} P\{\text{first visit to } j \text{ at } m^{th} \text{ step } |X_0 = i.\} P\{X_n = j | X_m = j\} \\ &= \sum_{m=1}^{n} f_{ij}^m p_{jj}^{n-m}. \end{aligned}$$

(iv) Follows from (iii)

(v)

$$
\begin{aligned}
&P\{\text{system visits state } j \text{ at least2 times } |X_0 = i\} \\
&= \sum_n \sum_k P\{\text{system has first visit to} j \text{ at } k|X_0 = i\} \times \\
&\qquad\qquad\qquad\qquad P\{\text{system has first visit at } n+k|X_k = j\} \\
&= \sum_n \sum_k f_{ij}^k f_{jj}^n = \left(\sum_n f_{ij}^k\right)\left(\sum_k f_{jj}^n\right) = f_{ij} f_{jj}.
\end{aligned}
$$

In the general case, similarly,

$$P\{\text{system has } m \text{ visits and at least to state j } |X_0 = i\} = f_{ij} f_{jj}^{(m-1)}.$$

■.

3.3.2 Definition:

(i) A state i is called **recurrent** if $f_{ii} = 1$, i.e., with probability 1, the system comes back to i.

(ii) A state i is called **transient** if $f_{ii} < 1$. Thus, the probability that the system starting at i does not come back to j, i.e., $(1 - f_{ii})$, is positive.

3.3.3 Theorem:

(i) The following statements are equivalent for a state j:
  (a) *The state is transient.*
  (b) $P\{$*system visits to j infinite number of times* $|X_0 = i\} = 0$.
  (c) $\sum_n p_{jj}^n < \infty$.

(ii) The following statements are equivalent for a state j:
  (a) *The state is recurrent.*
  (b) $P\{$*system visits to j infinite number of times* $|X_0 = i\} = 1$.
  (c) $\sum_n p_{jj}^n = \infty$.

Proof:

(i) Using (v) of theorem 3.3.1, we have

$$
\begin{aligned}
&P\{\text{system visits to } j \text{ infinite number of times } |X_0 = i\} \\
&\qquad = \lim_{m\to\infty} P\{\text{system has at least } m \text{ visits to state} j|X_0 = i\}. \\
&\qquad = \lim_{m\to\infty} (f_{ij} f_j j^{(m-1)}) \\
&\qquad = f_{ij}\left(\lim_{m\to\infty} (f_{jj})^{(m-1)}\right).
\end{aligned}
$$

Hence,

$$P\{\text{system visits to } j \text{ infinite number of times } |X_0 = i\} = 0 \text{ iff } f_{jj} < 1.$$

This shows that (b) holds iff (a) holds.

Next suppose (c) holds, i.e., $\sum_n p_{jj}^n < \infty$. Then by Borel-Cantelli lemma, (b) holds.

Conversely, let (a) holds, i.e., $f_{jj} < 1$. We shall show (c) holds. Using 3.3.2(ii), we have

$$\begin{aligned}\sum_{t=1}^{n} p_{jj}^{t} &= \sum_{t=1}^{n}\sum_{s=0}^{t-1} f_{jj}^{(t-s)} p_{jj}^{s} \\ &= \sum_{s=0}^{n-1} p_{jj}^{s} \sum_{t=s+1}^{n} f_{jj}^{(t-s)} \;\leq\; f_{jj} + \sum_{s=1}^{n} p_{jj}^{s} f_{jj}.\end{aligned}$$

Thus, $(1-f_{jj})\left(\sum_{t=1}^{n} p_{jj}^{t}\right) \leq f_{jj}$. Thus, for every $n \geq 1$

$$\sum_{t=1}^{n} p_{jj}^{t} \leq \frac{f_{jj}}{1-f_{jj}},$$

implying (c) as $f_{jj} < 1$. This completely proves (i).

Proof of (ii) follows from (i).

■

3.3.4 Example:

Consider the unrestricted random walk on the integers with probability p moving to right, probability q moving to left, and $p + q = 1$. It is clearly an irreducible chain.

Starting at 0 one can come back to 0 only in even number of steps. Thus,

$$p_{00}{}^{2n+1} = 0, \text{and } p_{00}^{2n} = \{X_{2n} = 0 \,|X_0 = 0\}.$$

Starting from 0 if it has to come back to 0 in $2n$ steps, then it can go to left in n steps and right by n steps. Thus,

$$p_{00}^{2n} = \binom{2n}{n} p^n q^n.$$

Therefore,

$$\sum_{m=0}^{\infty} p_{00}^{2n} = \sum_{n=0}^{\infty} p_{00}^{2n} = \sum_{n=0}^{\infty} \binom{2n}{n} p^n q^n.$$

To decide whether the state 0 is transient or not, one has to know whether this series is convergent or not. Note that,

$$\binom{2n}{n} = \frac{2n!}{n!.n!},$$

and by sterling's formula, $n! \sim (\sqrt{2\pi})n^{n+1/2}$, we have

$$\begin{aligned}\binom{2n}{n} &\sim \frac{(2n)^{2n+1/2}}{n^{n+1/2}.n^{n+1/2}} \\ &= 2^{2n}.2!\frac{n^{2n+1/2-2n-1}}{\sqrt{2\pi}} = \frac{2^{2n}}{\sqrt{n\pi}}\end{aligned}$$

Hence,

$$p_{00}^{2n} \sim \frac{(4pq)^n}{\sqrt{n\pi}}.$$

Since $p(1-p) = pq \leq 1/4$ and equality holds iff $p = q = 1/2$. Thus, for $\theta = 4pq$,

$$\sum_{n=0}^{\infty} p_{00}^{2n} \sim \begin{cases} \displaystyle\sum_{0}^{\infty} \frac{\theta^n}{\sqrt{n}}, \;\; \theta < 1 & \text{if } p \neq q \neq 1/2. \\ \displaystyle\sum_{0}^{\infty} \frac{1}{\sqrt{n}} & \text{if } p = q = 1/2. \end{cases}$$

One knows that for $\theta < 1$, $\sum_0^\infty \frac{\theta^n}{\sqrt{n}} < +\infty$ and is divergent if $\theta = 1$. Thus, 0 is a recurrent state iff $p = q = 1/2$. In fact same holds $\forall$ state j. If $p \neq q$, then intuitively particle will drift to $-\infty$ or $+\infty$. as 0 is the transit state and so in every state.

3.3.5 Theorem:
Let $i \to j$ and i be recurrent. Then,

(i) $f_{ji} = 1$, $j \to i$ *and* j *is recurrent.*

(ii) $f_{ij} = 1$

Proof:

(i) Since $i \to j$, there exists $n \geq 1$ such that $p_{ij}^n > 0$. Let n_0 be the smallest positive integer such that $p_{ij}^{n_0} > 0$. Then, $p_{ij}^m = 0$ for $1 \leq m < n$. Since $p_{ij}^{n_0} > 0$, there exists states $i_1, i_2, ... i_{n_0-1}$, none equal to j such that

$$P\{X_{n_0} = j, X_{n_0-1} = i_{n_0-1}, ... X_1 = i_1 | X_0 = i\} \quad > \quad 0. \tag{3.1}$$

Suppose $f_{ji} < 1$. Then $(1 - f_{ji}) > 0$, i.e.,

$$P\{\text{system starts at } j \text{ but never visits } i\} \quad > \quad 0. \tag{3.2}$$

Therefore,

$$\begin{aligned}
\alpha : \quad = \quad & P\{X_1 = i_1, ..., X_{n_0-1} = i_{n_0-1}, X_{n_0} = j, \\
& \qquad X_n \neq i \text{ for } n > n_0 | X_0 = i\} \\
= \quad & P\{X_n \neq i \text{ for } n \geq n_0 + 1 | X_{n_0} = j, X_{n_0-1} = i_{n_0-1}, \ldots \\
& \qquad ..., X_0 = i\} \times P\{X_{n_0} = j, X_{n_0-1} = i_{n_0-1}, \ldots \\
& \qquad \qquad ..., X_{i_1} = i_1 | X_0 = i\} \\
= \quad & P\{X_n \neq i \text{ for } n \geq n_0 + 1 | X_{n_0} = j\} \times \\
& \qquad P\{X_{n_0} = j, ..., X_1 = i_1 | x_0 = i\}, \\
> \quad & 0,
\end{aligned}$$

using equations (3.1) and (3.2). Thus

$$P\{X_n \neq i \text{ for every } n | X_0 = i\} > \alpha > 0 \text{ for all } n,$$

i.e., the system starts at i and never comes back to i, i.e., i cannot be a recurrent state. Hence, if i is recurrent then our assumption that $f_{ji} < 1$ is not true.

Thus, i recurrent implies $f_{ji} = 1$. But then,

$$f_{ji} = \sum_{m \geq 1} f_{ji}^m = 1,$$

and hence for some m, $f_{ji}^m > 0$, i.e., with positive probability there is a first visit to i starting from j. Hence $p_{ji}^m \geq f_{ji}^m > 0$, i.e., $j \to i$. Thus, we have shown $i \to j$ and i recurrent implies $f_{ji} = 1$. and hence $j \to i$. Further,

$$p_{jj}^{m+n+n_0} \quad = \quad \sum_{r,k} p_{jr}^m p_{rk}^n p_{kj}^{n_0}. \ > \ p_{ji}^m p_{ik}^n p_{ij}^{n_0}.$$

Using this,

$$\begin{aligned}\sum_{n\geq 1} p_{jj}^n &\geq \sum_{n=m+1+n_0} p_{jj}^n.\\ &= \sum_{n\geq 1} p_{jj}^{m+n+n_0}\\ &> \sum_{n\geq 1} p_{ji}^m p_{ii}^n p_{ij}^{n_0} = p_{ji}^m \left(\sum_{n\geq 0} p_{ii}^n\right) p_{ij}^{n_0} = \infty,\end{aligned}$$

because

$$p_{ji}^m > 0,\ p_{ij}^{n_0} > 0, \text{ and } \sum p_{ii}^n = +\infty.$$

Thus, j is recurrent, proving (i).

(ii) Apply (i) to i and j, interchange. ■

3.3.6 Corollary:
If $i \to j$ and $j \to i$, then, either both are transient or both are recurrent.

Proof:
If i is recurrent, and $i \to j$ then, j is recurrent by above theorem. Let i be transient and j be recurrent. But as $j \to i$, and hence by above theorem i is recurrent, not possible. Hence, i transient implies j transient. ■

3.3.7 Corollary:

Let $C \subset S$ be an irreducible set. Then, either all states in C are recurrent or all are transient. Further, if C is a communicating class and all its states are recurrent, then C is closed.

Proof:
Since all states in C communicate with each other,by corollary 3.3.6, all states in C are either transient or recurrent. Next suppose C is a communicating class and $j \notin C$. Let $i \to j$ for some $i \in C$. Then by above theorem above, $j \to i$, and hence $j \in C$, not true. Hence C is closed.

■

Hence we know how to characterize irreducible markov chains.

3.3.8 Exercise:
Show that if a state j is transient, then $\sum_{n=1}^{\infty} p_{ij}^n < \infty$ for all i.

3.3.9 Theorem:
Let $\{X_n\}_{n\geq 1}$ be an irreducible markov chain with state space S and transition probability matrix P.

(i) Either all states are transient in which case

$$\sum_{n\geq 0} p_{ij}^n < +\infty \ \ \textit{for all}\, i, j.$$

and

$$P\{X_n = j \textit{ infinite}\, n's | X_0 = i\} = 0$$

(ii) All states are recurrent in which case

$$\sum_{n\geq 0} p_{ii}^n = +\infty \textit{ for all}\, i.$$

3.3.10 Corollary:
If S is finite then it has at least one recurrent state.

Proof:
Suppose all states are transient. Then,

$$\sum_{n\geq 0} p_{ij}^n < +\infty \text{ for all } i, j.$$

Thus, $\lim_{n\to\infty} p_{ij}^n = 0$. Hence, as S is finite and P is a stochastic matrix,

$$0 = \lim_{n\to\infty} \sum_{j\in S} p_{ij}^n = 1$$

a contradiction. ■

3.3.11 Corollary:
In a finite irreducible chain, all states are recurrent.

3.3.12 Examples:
The two states markov chain with transition matrix

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

is irreducible, finite and hence all states are recurrent.

3.3.13 Example :
Consider the chain discussed in example 3.1.3 with transition matrix

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 1/5 & 2/5 & 1/5 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1/4 & 0 & 3/4 \end{pmatrix} \end{array}.$$

Let us find its transient, recurrent states.

(i) 0 is an absorbing state as $p_{00} = 1$ and hence is recurrent.

(ii) As observed earlier $\{3, 4, 5\}$ is a finite ,closed, irreducible set, hence by corollary 3.3.11, all states are recurrent.

(iii) Now if 2 was a recurrent state, since $2 \to 0$, and by theorem 3.3.5, we should have $0 \to 2$, but that is not true. Hence 2 is not recurrent and hence must be transient. Similarly, 1 is transient.

Thus we can write the state space as $S = \{1, 2\} \cup \{3, 4, 5\}$, where first set consists of transient states and second is irreducible set of recurrent states.

3.3.14 Example :
Let us find transient/recurrent state for chains with transition matrices:

$$P = \begin{array}{c} \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} \end{array}, \quad Q = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{array},$$

$$R = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 0 & 1/2 \end{pmatrix} \end{array}$$

Chain with transition matrix P is finite irreducible and thus recurrent and finite. The chain with transition matrix Q is also finite irreducible and hence recurrent. For the chain with transition matrix R, $\{1,2\}$ and $\{3,4\}$ are irreducible sets and hence are recurrent. Since, $5 \to 1$ but $1 \not\to 5$ so 5 cannot be recurrent. Therefore, 5 is transient. Once again, we have the decomposition $S = \{5\} \cup \{1,2\} \cup \{3,4\}$, where first set consists of transient state and second and third sets are irreducible sets of recurrent states.

We had saw in above example, that the state space S could be written as $S_T \cup C, \cup ...$ When S_T consists of all transient states, $C_1, C_2, ...$ are closed irreducible sets containing of recurrent states. We show this is possible in general.

3.3.15 Proposition:
For every recurrent state i there exists a closed subset $C(i)$ such that the following holds:

(i) Each $C(i) \neq \emptyset$, is closed and irreducible.

(ii) Either $C(i_1) \cap C(i_2) = \emptyset$ or $C(i_1) = C(i_2)$.

(iii) $\cup_i C(i) = S_R$, set of all recurrent states

Proof:
For $i \in S_R$, define

$$C(i) = \{j \in S | i \to j\}.$$

We prove that the sets $C(i)$ has the required properties.

(i) $i \in C(i)$ for $p^0_{ii} = 1$ and hence $C(i) \neq \emptyset$.

If $j \in C(i)$ then j is recurrent and $j \to i$. Hence $i \leftrightarrow j$.

Thus, any two states in C communicate with each other, i.e., C is irreducible. If $k \notin C(i)$, then $i \not\to k$, for otherwise $k \to i$ implying $k \in C$.

Also for $j \notin C$, $i \leftrightarrow j$ and hence $j \not\to k$ for if $j \to k$ then $i \to k$. Therefore, $C(i)$ is closed.

(ii) If $i \in C(i_1) \cap C(i_2)$, then for $j \in C(i_1)$,

$$j \leftrightarrow i_1 \leftrightarrow i \leftrightarrow i_2$$

implying $C(i_1) \subseteq C(i_2)$.

Similarly, $C(i_2) \subseteq C(i_1)$.

(iii) is obvious. ■

3.3.16 Theorem (Decomposition of state space):
$S = S_T \cup S_R$, where S_T consists of all transient states, S_R consists of all recurrent states, such that $S_R = C_1 \cup C_2 \cup ...$, consisting of closed irreducible disjoint sets C_i.

Proof:
Clearly $S = S_T \cup S_R$ is possible by definition. Required decomposition follows from proposition 3.3.15. ■

3.3.17 Note:
Thus, we can write the state space as

$$S = S_T \cup C_1 \cup C_2 \cup ...$$

where S_T consists of transient states, each C_i is irreducible recurrent. On each C_i chain can be analyzed (irreducible.) If S_T is also irreducible closed, we can analyze chain on it separately also. In general, locating a recurrent state in a chain may not be easy.

3.3.18 Some questions:

(i) If chain starts in S_T, what is the probability that it will stay always in S_T?

(ii) Given $i \in S$ what is the probability that it will hit a closed irreducible set C of recurrent states and stay in it for ever? Clearly, this probability is

$$p_c(i) = \begin{cases} 1 & \text{if } i \in C. \\ \\ 0 & \text{if } i \notin C \text{ but } i \text{ is recurrent.} \end{cases}$$

So case of interest is for $i \in S_T$, what is $p_c(i)$? (iii) Can we have an alternative criterion for a state to be transient or recurrent ?

We shall answer some of these in the next section.

3.4. Absorption probability

Let C be an irreducible closed set of recurrent states. For $i \in S$,

$$p_c(i) = \begin{cases} P\{\text{system hits } C \text{ eventually} | x_0 = i\} \\ \\ \bigcup_{n \geq 0} P\{X_n \in c\, | x_0 = i\} \end{cases}$$

Note that if $i \in C$, then, $p_c(i) = 1$. If $i \notin C$, but i is recurrent, then $p_c(i) = 0$. So, the problem is to compute $p_c(i)$ when i is in S_T? The answer is given by

3.4.1 Theorem: *$p_c(i)$ satisfy the system of equations for $i \in S_T$ to go from i to C, we can go either from i to $j \in C$ in one step or i to $j \in S_T$ in one step and then from j to C.*

$$p_c(i) = \sum_{j \in C} p_{ij} + \sum_{j \in S_T} p_{ij}\, p_c(j).$$

Thus, to find $p_c(i)$ one has to solve these equations.When S_T is infinite, this is not known how to solve these equations. Moreover, solutions need not be unique in that case. When S_T is finite , one can show a unique solution exists. We give an example to illustrate this.

3.4.2 Example
Let

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 1/5 & 2/5 & 1/5 & 0 & 1/5 \\ 0 & 0 & 0 & 1/6 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1/4 & 0 & 3/4 \end{pmatrix} \end{array}.$$

Then as observed in example 3.3.13, $C = \{0\}$ is a closed irreducible set and $S_T = \{1, 2\}$. Let us find $p_c(1),\ p_c(2)$ for $C = \{0\}$. We have to solve

$$\begin{aligned} p_c(1) &= \sum_{j\in c} p_{ij} + \sum_{j\in S_T} p_{ij}p_c(j) \\ &= p_{i\,0} + p_{1\,1}p_c(1) + p_{1\,2}p_c(2). \\ p_c(1) &= \frac{1}{4} + \frac{1}{2}p_c(1) + p_{\frac{1}{4}}p_c(2) \qquad (3.3) \\ p_c(2) &= p_{2\,0} + p_{2\,1}p_c(1) + p_{2\,2}p_c(2) \\ &= 0 + \frac{1}{5}p_c(1) + \frac{2}{5}p_c(2) \qquad (3.4) \end{aligned}$$

One can solve (3.4) and (3.5) to get $p_c(1) = \frac{3}{5},\ p_c(2) = \frac{1}{5}$.

3.4.3 Definition:
A Markov chain is called an **absorbing chain** if

(i) It has at least one absorbing state; and

(ii) For every state in the chain, the probability of reaching an absorbing state in a finite number of steps is nonzero.

Suppose an absorbing markov chain has r absorbing states and a set of transient states grouped as set S_T : we can write

$$S = S_T \cup C_i \cup C_2 \cup \ldots \cup C_r,$$

where each C_i is a singleton set corresponding to each absorbing state. Thus, if need be we can renumber the states and assume that the transition matrix has the form

$$P = \begin{array}{c} \text{absorbing states } 1 \leftrightarrow r \qquad \leftarrow S_T \rightarrow \\ \begin{array}{c} \uparrow \\ 1 \le i \le r \\ \downarrow \\ \uparrow \\ j \in S_T \\ \downarrow \end{array} \left(\begin{array}{c|c} \text{I} & \text{O} \\ \hline \text{R} & \text{Q} \end{array}\right) \end{array},$$

where, R is the rectangular submatrix giving transition probabilities from non-absorbing to absorbing states, Q is the square submatrix giving these probabilities from non- absorbing to non-absorbing states, I is an identity matrix, and 0 is a rectangular matrix of zeros. Note that for every n

$$P^n = \left(\begin{array}{c|c} I & 0 \\ \hline (I + Q + Q^2 + \ldots + Q^{n-1})R & Q^n \end{array}\right).$$

Thus, if $Q^n = (q^n_{ij})$, then q^n_{1j} represents the probability of going from the non-absorbing state i to the non-absorbing state j in n steps. Since the absorption probabilities satisfy:

$$p_{C(i)}(j) = p_{ji} + \sum_{k\in S_T} p_{C(i)}(k),\ i = 1, ...r, \text{and } j \in S_T,$$

we have

$$B = R + QB$$

where, B is the matrix with ij^{th} entry being $p_{C(i)}(j)$. Thus,

$$B = (I - Q)^{-1}R = NR,$$

where $N := (I - Q)^{-1}$, if it exists. Hence, to calculate the absorption probabilities, one has to to show that N exists and calculate $(I - Q)^{-1}$. The matrix $N = (I - Q)^{-1}$ is called the **fundamental matrix** of the absorbing chain.

3.4.4 Theorem:
For every absorbing chain the following holds:

(i) If q^n_{ij} denote the denote the entries of Q^n, then the mean absorption tine for a state i is

$$\mu_i := \sum_{j\in S}\sum_{m=0}^{\infty} q^n_{ij}.$$

(ii) $Q^n \to 0$, as $n \to \infty$.

(iii) $N := (I - Q)^{-1}$ exists.

(iv) If $B = (I - Q)^{-1}R = NR = [b_{ij}]_{i\times j}$, Then b_{ij} is the probability that the chain will be absorbed in state j starting from state i.

(v) If $N = [n_{ij}]_{i\times j}$, then n_{ij} is the expected number of steps needed to go from $i \in S_T$ to $j \in S_T$.

Proof:

(i) The mean absorption tine for a state i is

$$\begin{aligned}
\mu_i &= \sum_{k=1}^{\infty} P\{\text{Starting in state } i \text{ chain is absorbed at } k^{th} \text{ step}\} \\
&= \sum_{m=0}^{k-1} P\{\text{Starting in state } i \text{ chain is absorbed at } k^{th} \text{ step}\} \\
&= \sum_{m=0}^{\infty}\sum_{k=m+1}^{\infty} P\{\text{Starting in state } i \text{ chain is absorbed at } k^{th} \text{ step}\} \\
&= \sum_{m=0}^{\infty} P\{\text{Starting in state } i \text{ chain is absorbed after } m \text{ step}\} \\
&= \sum_{m=0}^{\infty} P\{\text{Starting in state } i \text{ chain is not absorbed by } m^{th} \text{ step}\} \\
&= \sum_{m=0}^{\infty}\left(\sum_{j\in S} q^m_{ij}\right) \\
&= \sum_{j\in S}\sum_{m=0}^{\infty} q^m_{ij}.
\end{aligned}$$

(ii) Note that $q^n_{ij} = p^n_{ij}$ for transient states i, j and for i a transient states i, j , $\sum_0^\infty p^n_{ij} < \infty$. Hence, $q^n_{ij} \longrightarrow 0$, as $n \to \infty$.

(iii) Define

$$N_n := I + Q + Q^2 + \ldots + Q^n = \sum_{k+0}^{n} Q^n, n \geq 1.$$

It is easy to check that

$$N_n(I - Q) = (I - q)N_n = I - Q^{n+1}, \text{ for all } n \geq 1. \tag{3.5}$$

Since, $\sum_0^\infty Q_{ij}^n < \infty$ by (ii), and $q_{ij}^n \longrightarrow 0$, we have

$$(I - Q)^{-1} = N := \lim_{n\to\infty} N_n$$

exists.

(iv) Claim is obious.

(v) For $i, j \in S_T$, let

$$X^{(k)} = \begin{cases} 1 & \text{chain is in state } j \text{ after } k \text{ steps, starting at } i. \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} P(X^{(k)} = 1) &= q_{ij}^k \\ P(X^{(k)} = 0) &= 1 - q_{ij}^k \end{aligned}$$

Thus, $E(X^{(k)}) = q_{ij}^k$. Hence the expected number of times the chain is in state j in first n steps ,starting in i, is

$$E(X^{()} + X^{()} + \ldots + X^{()}) = \sum_{k=0}^{n} q_{ij}^k.$$

Thus, using Fubini's theorem, the expected number of times the chain is in state j, starting in i, is

$$\begin{aligned} E\left(\sum_k^\infty X^{(k)}\right) &= E\left(\lim_{n\to\infty} \sum_k^n X^{(k)}\right) \\ &= \lim_{n\to\infty} E\left(\sum_k^n X^{(k)}\right) \\ &= \lim_{n\to\infty}\left(\sum_k^n q_{ij}^k\right) = \sum_0^\infty q_{ij}^n. \end{aligned}$$

■

The matrix N also helps us to compute t_i, the mean(average) number of steps (time) for which the chain will be in transient states starting from the state $i \in S_T$. This is given by $t_i = \sum_{j\in S_T} n_{ij}$. This is also the absorbtion time starting at i. We apply these to our case of random walk with absorbing barriers (with $n+1$ states.)

3.4.5 Example:

$$P = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ \\ \vdots \\ \text{n} \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & \ldots & \ldots & \ldots & n \end{array} \\ \left(\begin{array}{ccccccc} 1 & 0 & 0 & \ldots & \ldots & \ldots & 0 \\ q & 0 & p & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & & \ldots \\ 0 & 0 & \ldots & \ldots & q & 0 & p \\ 0 & 0 & 0 & & \ldots & \ldots & 1 \end{array}\right) \end{array}$$

We write this as (interchange 2^{nd} and n^{th} row and 2^{nd} and n^{th} column.)

$$P = \begin{array}{c} \\ 0 \\ \text{n} \\ \vdots \\ \\ \vdots \\ \text{n-1} \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & n & 1 & \ldots & \ldots & \ldots n-1 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & \ldots & \ldots & 0 \\ q & 0 & 0 & p & \ldots & \ldots & \\ \ldots & \ldots & q & \ldots & \ldots & \ldots & \\ \ldots & 0 & \ldots & q & 0 & p & \\ 0 & 0 & \ldots & 0 & q & 0 & \end{pmatrix} \end{array}$$

$$R = \begin{pmatrix} q & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & p \end{pmatrix}, Q = \begin{pmatrix} 0 & p & 0 & \ldots & 0 & 0 \\ q & 0 & p & 0 & \ldots & 0 \\ 0 & q & 0 & p & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \\ \ldots & \ldots & q & 0 & p & 0 \\ 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & q & 0 \end{pmatrix}$$

$$I - Q = \begin{pmatrix} 1 & -p & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ -q & 1 & -p & \ldots & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots 0 & 0 & \ldots & 1 & -p \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & q & 1 \end{pmatrix}$$

The inverse of $I - Q$ is given by $N = (n_{ij})$,

(i) If $p + q,\ p/q = c$,then

$$n_{ij} = \frac{1}{(p-q)(r^n - 1)} \times \begin{cases} (r^j - 1)(r^{n-i} - 1) \text{ for } j \leq i \\ (r^i - 1)(r^{n-i} - r^{j-1}) \text{ for } \ j \geq i. \end{cases}$$

(ii) If p=q=1/2, then

$$n_{ij} = \frac{2}{n} \begin{cases} j(n-i) \text{ for } j \leq i \\ i(n-j) \text{ for } j \geq i. \end{cases}$$

And the time to stay in the transient state is

$$t_i = \sum_{j=1}^{n-1} n_{ij} = \begin{cases} \frac{1}{p-q}[n\left(\frac{r^n - r^{n-i}}{r^n - 1}\right) - i] \text{ if } p \neq 1/2 \\ i(n-i) \text{ if } p = 1/2. \end{cases}$$

From this we can draw the following:

Conclusions:

- The time t_i of stay in transient state starting from i, or equivalently to the time to get out of transient state depends upon i, even when $p = 1/2$. Note $t_i = i(n-i)$ will be maximum when i is in the middle namely $i = n/2$ (if n is even). Therefore, $t_{\max} = (n/2)^2$. Thus, when both players have equal probability of win, thee time of ruin is the product of the fortunes of the two players namely $i(n-i)$. And the game will last maximum time when both have equal amount to start with. But if $p \neq 1/2$, one can show that

$$i_{\max} = log_r((r-1)n)$$

and for $r > 1$,

$$t_{\max} \ \frac{n - i_{\max}}{p-q},$$

which is of lower order of magnitude compared to the case $p = 1/2$ case i.e., the game will finish much faster in this case. Next we calculate $B = NR$.

Since N is $(n-1) \times (n-1)$ and R is $(n-1) \times 2$ matrix, $B = (n-1) \times 2$ matix.First row gives probability for absorption in n. But $p_{i0} = 1 - p_{in}$ $\forall i$. Thus enough to calculate any one one of them. We have

$$b_{in} = \sum_{k=1}^{n-1} n_{ik}.r_{kn} = \begin{cases} \dfrac{r^n - r^{n-i}}{r^n - 1} & \text{if } p \neq 1/2 \ (p/q = r) \\ i/n & \text{if } p = 1/2. \end{cases}$$

(i) If $p = 1/2$, probability that A is ruined is proportional to the ratio of fortune A starts with to total fortune n

(ii) If $p > 1/2$, i.e., A has an advantage over B, then his chance of ruining his opponent is $\dfrac{r^n - r^{n-i}}{r^n - 1}$. Suppose $r = 2, i = 1$ then this is

$$\frac{2^n - 2^{n-1}}{2^n - 1} = \frac{2^{n-1}}{2^n - 1} = 1/2 \text{ (as } n \to \infty) = 1 - \frac{2^{n-1}}{2^{n-1}},$$

which is quite good (for example if $n = 2$, i.e., opponent also has 1 rupee, then this is 2/3 but even for n large this means A can have a good chance of ruining B even if B has large capital and A starts only with rupee 1.

Imagine A is a gambling house and B is the player. Gambling house fixes odd states $r > 1$ and make sure i is large for them. Then always stays in business is approximately

$$\lim_n \to \infty(b_{in}) \simeq 1 - (1/r)^i \simeq 1.$$

Note that

$$\lim_n \to \infty(b_{in}) = \lim_n \to \infty \left(\frac{r^n - r^{n-i}}{r^n - 1} = 1 - (1/r)^i \right) \simeq 1.$$

Therefore,

$$\text{probability that the player wins all the money } = \simeq 1 - (\lim_n \to \infty b_{in}).$$

Thus, gambling houses stay in business no matter how much is bet at the tables. Let us see the absorbtion of the gambling house winning all the money for r near 1, (i.e., odds are favorable to the gambling house but not very much :$t_i \simeq i(n-i)$). Thus, if the gambling house can cover $10,000 = 10^4$ bets while all gambler put together can provide 10^6 bets, then $i(n-i) = 10^10$ units of time, which is very large. So it will take very long time to win all the money and in the mean time more new gamblers would have born.

3.5. More on recurrence/transience

3.5.1 Another way of deciding a state is recurrent/transient

LetS_T denote the collection of all transient states of a system with transition probability matrix P. From P remove the rows and columns for the states not in S_T. Let Q be the sub matrix obtained. (Q in general will not be a stochastic matrix.) It will only be non-stochastic matrix. Let

$$Q = (q_{ij})ij \in S_T.$$

Consider a system of linear equations in variables $x_1, x_2, ...,\ S_T = \{1, 2, ...\}$:

$$x_i = \sum_{k \in T} q_{ik}\, x_k,\ 0 \leq x_i \leq 1,\ i \in S_T.$$

(i) The maximal solution of the above system are the probabilities that if a system starts at $i \in S_T$, then it will stay in that. Thus, maximal

$$x_i = P\{X_k \in S_T | x_0 = i.\} \; \forall k.$$

(ii) From a transient state, what is the probability it will go in a recurrent state and then stay in there only? Let C be a closed set of recurrent state. The probability $y_C(i)$ that the system starting at a state i will reach state C and then forever remains in it. Clearly, if C is irreducible, then

$$y_C(i) = 1 \text{ if } i \in C.$$
$$y_C(i) = 0 \text{ if } i \notin C, \text{ but } i \text{ is recurrent.}$$

The case of interest is when i is transient. In that case, $y_C(i)$ is the minimum non negative solution of the equation.

$$y_C(i) = \sum_{j \in C} p(ij) + \sum_{j \in T} p(ij)\; y_C(i).$$

Let $i_0 \in S$. Consider C_{i_0} is the closure of the set $\{i_0.\}$

Let $C_{i_0} = \{j_1, j_2, ...\}$ Then, i_0 will be transient iff the system of equations

$$x_{j_i} = \textstyle\sum_k p_{j_i} j_k x_{jk} \; 0 \leq j_i \leq 1 \;\; \}$$

have a non trivial solution. Note $x_{ji} = 0 \; \forall i$ is the only solution. Let us apply this criterion. We give some examples.

3.5.2 Example:
Consider the following queueing model.

$$X_n = \text{Number of customers at the counter.}$$
$$\xi_n = \text{Number of new customers that arrive in } n^{th} \text{ minute.}$$

Each $\{\xi_n\}$ can take only three values $\{0, 1, 2\}$ with probability $\{\alpha_0, \alpha_1, \alpha_2\}$, i.e., the distribution of ξ_n is

$$\begin{array}{c} p\{\xi_n = 0\} = \alpha_0 \\ p\{\xi_n = 1\} = \alpha_1 \\ p\{\xi_n = 2\} = \alpha_2 \end{array} \quad ,$$

$\alpha_0 + \alpha_1 + \alpha_2 = 1$. Then, $S = \{0, 1, 2, ...\}$ and transition probability matrix.

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \\ \end{array}
\begin{array}{c} \quad 0 \quad\; 1 \quad\; 2 \quad\; 3 \qquad\qquad \\ \begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & 0 & ... & \cdots \\ \alpha_0 & \alpha_1 & \alpha_2 & 0 & ... & \cdots \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & 0 & \cdots \\ 0 & 0 & \alpha_0 & \alpha_1 & \alpha_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \end{array},$$

$$p_{00} = \alpha_0 = \text{Number of customers at } n \text{, no new comes}$$
$$p_{01} = \alpha_1 = \text{Number of customers at } n \text{, one customer comes}$$
$$p_{02} = \alpha_2 = \text{Number of customers at } n \text{, two customer comes}$$

If we assume $\alpha_0, \alpha_2 \neq 0$, then, the chain is irreducible. We want to know whether it is recurrent or transient. Let us look at state 0. We have to see whether we can find a non-trivial solution of

$$x_i{=}\textstyle\sum_{j \neq 0} p_{ij} x_j, \;\; 0{\leq}\, x_1 \leq 1, \;\; i \neq 0.\; 0 \leq x_1 \leq 1.$$

In our case, equations are

$$\begin{aligned} x_1 &= \alpha_1 x_1 + \alpha_2 x_2. \\ x_k &= \alpha_0 x_{k-1} + \alpha_1 x_k + \alpha_2 x_{k+1},\ k \geq 2. \end{aligned}$$

One can show that a solution is given by (see Billiysley page 126 :)

x_k=

$$\begin{cases} B[(\frac{\alpha_0}{\alpha_2})^k - 1], & \text{if } \alpha_0 \neq \alpha_2, \\ B_k & \text{if } \alpha_0 = \alpha_2, \end{cases}$$

for some constant B. Thus, if $\alpha_0 \geq \alpha_2$, then $(\frac{\alpha_0}{\alpha_2})^k \to \infty$ and $B_k k \to \infty$. as $k \to \infty$. Hence, non trivial solution exists if $\alpha_0 < \alpha_2$ in which case chain will be transient. But transient means with probability 1, the chain must go away from state j. Hence with probability 1 the chain queue size will go to ∞. Not in this case, $(\alpha_2 - \alpha_0) > 0$ which is the expected increase in queue length. Queue goes to ∞ iff this is > 0.If $\alpha_0 \geq \alpha_2$, then chain is recurrent i.e., every state will be visited infinitely often.

Since

$$S = S_T \cup C_1 \cup C_2 \cup ...$$

We ask the question: Given system starts in S_T, what is probability that it will stay in it? The answer is as follows:

3.5.3 Theorem:
Let $U \subset S_T$ and $i \in U$. Then,

$$\sum_i = P\{X_n \in U \forall N \geq 1 | x_0 = i\},\ i \in U$$

are the maximal solutions of the system

$$x_i = \textstyle\sum_{j \in U}\ p_{ij}\ x_j | i \in U\ \ 0 \leq x_i \leq 1.$$

Let us look at an example.

3.5.4 Example(staying in transient states):
Consider the unrestricted random walk with transition matrix

$$\begin{pmatrix} q & 0 & p & 0 & 0 & \dots \\ 0 & q & 0 & p & 0 & \dots \\ 0 & 0 & q & 0 & p & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

We know all the states are transient. Consider $U = \{0, 1, 2, ...\} \subset S$. We want to know what is the probability that staying at some $i \in U$. It will stay in U. This is given by maximal solution of

$$x_i = \textstyle\sum_{j \in U}\ p_{ij}\ x_j | i \in U\ \ 0 \leq x_i \leq 1.$$

In our case, these are

$$\begin{aligned} x_0 &= p_{00}x_0 + p_{01}x_1 = px_1. \\ x_i &= p_{ii-1}x_{i-1} + p_{ii+1}x_{i+1}. \\ (p+q)x_i &= qx_{i-1} + px_{i+1}. \\ p(x_{i+1} - x_i) &= q(x_i - x_{i-1}). \\ x_{i+1} - x_i &= \frac{q}{p}(x_i - x_{i-1}). \end{aligned}$$

The only bounded solution for $q \geq p$ is $x_0 = 0 = x_1$, implying $x_{i+1} = 0\ \forall i$. Therefore, $x_i = 0$. In this case, probability of staying on non negative side is zero. If $q < p$, then maximal solution can be found as $x_1 = \dfrac{x_0}{p}$.

$$
\begin{aligned}
x_2 - x_1 &= \frac{q}{p}(x_1 - x_0) = \left(\frac{q}{p}\right)^2 x_0 \\
x_2 &= x_1 + \left(\frac{q}{p}\right)^2 x_0 \\
&= \frac{x_0}{p} + \left(\frac{q}{p}\right)^2 x_0 \\
&= x_0 \left[\frac{1}{p} + \left(\frac{q}{p}\right)^2\right].
\end{aligned}
$$

Thus, for general $n \geq 1$,

$$x_n = 1 - \left(\frac{q}{p}\right)^n.$$

Therefore,

$$x_n = P\{\text{system stay in } (0,1,2)|X_0 = n\} = 1 - \left(\frac{q}{p}\right)^n$$

As n becomes large, this probability goes to 1.

Chapter 4

Stationary distribution for a markov chain

4.1. Introduction

Let $\{X_n\}n \geq 0$ be a markov chain and P be it's transient probability matrix. Let $\Pi_0(i)$ be its initial distribution. In this chapter, we want to analyze the asymptotic (long run) behavior of the chain. Suppose there exist $\{\mu_i\}_{i\in S}$, such that $\mu_i \geq 0$ for all $i \in S$, $\sum_{i\in S} \mu_i = 1$ and

$$\sum_{i\in S} \mu_i P_{ij} = \mu_j,\ j \in S. \tag{4.1}$$

Then (4.1) implies that

$$\sum_{i\in S} \mu_i p_{ij}^2 = \sum_{i\in S} \mu_i \left(\sum_{l\in S} p_{il} p_{lj}\right) = \sum_{l\in S} \mu_l P_{lj}^n = \mu_j.$$

Using induction, for all $n \geq 0$,

$$\sum_{i\in S} \mu_i p_{ij}^n = \mu_j,\ j \in S.$$

In case, $\mu_i = \Pi_0(i)$ for every i, we have

$$P\{X_n = j\} = \sum_{i\in S} \Pi_0(i) p_{ij}^n = \Pi_0(j),\ j \in S.$$

Thus all $X_n's$ have the same distribution. Thus, in some sense, the chain is very stable.

4.1.1 Definition:
A markov chain $\{X_n\}n \geq 0$ with P as it's transient probability matrix and $\Pi_0(i)$ as its initial distribution is said have a **stationary distribution** or an **invariant distribution** if there exist $\{\mu_i\}_{i\in S}$, such that μ_i for all $i \geq 0$, $\sum \mu_i = 1$, and

$$\sum_{i\in S} \mu_i P_{ij}^n = \mu_j,\ j \in S.$$

Given a markov chain, one would like to answer the following questions:

(i) When does the markov chain has a stationary distribution?

(ii) How to get the stationary distribution if it exists?

(iii) Is stationary distribution unique?

(iv) What are the consequences of having a stationary distribution?

In the next section we look at the concept of stopping times, needed to answer the above questions.

4.1.2 Example:
Consider a markov chain with

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

Intuitively, chain spends one third of the time in state 1, one third of the time in state 2, one third of the time in state 3. In fact, if we take $\Pi_0 = (1/3, 1/3, 1/3)$, then $\Pi_0 = \Pi_0 P$, i.e., Π_0 is a stationary distribution. The chain is irreducible with period 3.

$$p_{ii}^n = \begin{cases} 0 & \text{if } n \text{ is not a multiple of 3,} \\ 1 & \text{if } n \text{ is a multiple of 3.} \end{cases}$$

Thus, $\{p_{ii}^n\}_{n\geq 1}$ is not convergent.

4.1.3 Example:
On a highway, three out of four trucks on the road are followed by a car, while only one out of every five cars are followed by a truck. What fraction of vehicles on the road are trucks?
To answer this question, we construct a markov chain as follows: Consider sitting on the side of the road and observe vehicles go by. Then, observation at time n, is

$$X_n = \begin{cases} 0 \\ 1, \end{cases}$$

where 0 signifies the appearance of a truck and 1 signifies appearance of a car. Thus the state space is $S = \{0, 1\}$ with transition matrix

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1/4 & 3/4 \\ 1/5 & 4/5 \end{pmatrix} \end{array}$$

If we want each X_n to have same distribution Π_0, then $\Pi_0 = \Pi_0 P$. Let $\Pi_0 = (p_0, p_1)$. Then

$$(p_0, p_1) = (p_0, p_1) \begin{pmatrix} 1/4 & 3/4 \\ 1/5 & 4/5 \end{pmatrix}.$$

Therefore,

$$\begin{aligned} p_0 &= \frac{p_0}{4} + \frac{p_1}{5} \\ p_1 &= \frac{3p_0}{4} + \frac{4p_1}{5} \\ 1 &= p_0 + p_1. \end{aligned}$$

This implies $p_0 = 4/19$, $p_1 = 15/19$. So, in the long run the fraction of trucks will be $4/19$.

4.2. Stopping times and strong markov property

Given a markov chain $\{X_n\}_{n\geq 1}$, let $\mathcal{A}_n$ denote the σ algebra determined by the random variables $X_0, X_1, ..., X_n$.

4.2.1 Definition:
A random variable $T : \Omega \to \mathbb{N} \cup \infty$ is called a **stopping time** if $\{T = N\} \in \mathcal{A}_n$ for all n.

4.2.2 Examples:

(i) For $i \in S$, let

$$S_i = \begin{cases} \inf\{n \geq 0 | X_n = i\} & \text{if such an } n \text{ exists.} \\ +\infty & \text{otherwise} \end{cases}$$

It is a stopping time called the **first passage time**.

(ii) For $i \in S$, let

$$T_i = \begin{cases} \inf\{n \geq 1 | X_n = i\} & \text{if such an } n \text{ exists.} \\ +\infty & \text{otherwise} \end{cases}$$

It is a stopping time called the **first return to state** i.

(iii) For $A \subseteq S$, let

$$T_A = \begin{cases} \inf\{n \geq 1 | X_n \in A\} & \text{if such an } n \text{ exists.} \\ +\infty & \text{otherwise} \end{cases}$$

T_A is a stopping time called the **first visit to set** A.

4.2.3 Note:
The event $\{T_j = n | X_0 = i\}$ is starting at i and first visit to j is at time n. In our earlier notations (see section 3.3,) $P\{T_j = n | X_0 = i\} = f_{ij}^n$. Thus,

$$f_{ii} = \sum_{n=0}^{\infty} f_{ii}^n = P\{T_j < +\infty | X_0 = i\}$$

Thus, a state i is recurrent iff $P\{T_i < +\infty | X_0 = i\} = 1$ and a state i is transient iff $P\{T_i < +\infty | X_0 = i\} < 1$.

Let $\{X_n\}_{n \geq 1}$ be a markov chain and T be a stopping time. Let

$$\mathcal{A}_T = \{B \in \mathcal{A} | B \cap \{T = n\} \in \mathcal{A}_n \ \forall n\}$$

It is called the **stopping tine σ- algebra** or the σ- algebra determined by the shopping time.

4.2.4 Proposition (Strong Markov Property):
For every $A \in \mathcal{A}_T, m > 0, i_1, i_2, ... i_m \in S$,

$$\begin{aligned} &P\{A \cap \{X_{T+1} = i_1, X_{T+2} = i_2, ..., X_{T+m} = i_m\} | X_T = i, T < +\infty\} \\ &\qquad = P\{A | X_T = i, T < +\infty\} P\{X_1 = i_1, ..., X_m = i_m | X_0 = i\} \end{aligned} \tag{4.2}$$

Proof:
It is enough to prove (4.2) for events $A \in \mathcal{A}$ of the type $A \cap \{T = m\} \in \cap \mathcal{A}_T$. For such an event A, (4.2) is

$$\begin{aligned} &P\{A \cap \{X_{n+1} = i_1, ..., X_{n+m} = i_m\} | X_n = i\} \\ &\quad = \quad P\{A | X_n = i\} P\{X_1 = i_1, ..., X_m = i_m | X_0 = i\} \end{aligned} \tag{4.3}$$

Now note that by the the markov property,(4.3) holds when A is a simple event in $\mathcal{A}_n$, determined by and the general event is a countable disjoint union of such events. ■

4.3. Existence and uniqueness:

From now onwards, we shall write

$$\begin{aligned} P_i(A) &= P\{A|X_0 = i\} \text{ for all events } A. \\ E_i(f) &= Ef|X_0 = i \text{ for every random variables } f. \end{aligned}$$

4.3.1 Theorem:
Let $\{X_n\}_{n\geq 1}$ be an irreducible recurrent chain with transition matrix P. Then there exist numbers $\{r_i^k\}_k, i \in S$ with the following properties:

(i) $r_k^k = 1$ for all $k \in S$.

(ii) $r_i^k = \sum_i \in S r_i^k p_{ij}\ j \in S$.

(iii) $0 < r_i^k < +\infty$ for all$i \in S$.

In other words, the chain has an stationary (invariant) measure.

Proof:
Let for $k, i \in S$,

$$r_i^k := \sum_{n-1}^{\infty} E_k(\mathcal{X}_{X_n = i, T_k \geq n})$$

this represents the total expected number of visits to state i between (any) two visits to state k. Note $\{X_n = i, T_k \geq n\}$ represent the event that the chain is in state at time n and not yet visited state k by n.
Thus if (k is recurrent),for $i = k$, the chain visits X_n only once till next visit to k. Thus,

$$r_k^k = 1 \forall k \in S.$$

This proves (i)
Next, for $k, j \in S, k \neq j$

$$\begin{aligned} r_j^k &= \sum_{n=1}^{\infty} E_k(\mathcal{X}_{X_n = j, T_k \geq n}) \\ &= \sum_{n=1}^{\infty} P_k(X_n = j, T_k \geq n) \\ &= P_k(X_1 = j, T_k \geq 1) + \sum_{n=2}^{\infty} P_k(X_n = j, T_k \geq n) \\ &= p_{kj} + \sum_{n=2}^{\infty} \sum_{\substack{i \in S \\ i \neq k}} P_k(X_n = j, X_{n-1} = i, T_k \geq n) \\ &= p_{kj} + \sum_{\substack{i \in S \\ i \neq k}} \sum_{n=2}^{\infty} P_k(X_n = j, X_{n-1} = i, T_k \geq n) \\ &= p_{kj} + \sum_{\substack{i \in S \\ k \neq j}} \sum_{n=2}^{\infty} P_k(X_{n-1} = i, T_k \geq n) p_{ij}. \end{aligned}$$

$$= p_{kj} + \sum_{\substack{i \in S \\ k \neq j}} \sum_{n=2}^{\infty} P_k(X_{n-1} = i, T_k \geq n-1) p_{ij} (\text{as} X_{n-1} = i)$$

$$= p_{kj} r_k^k + \sum_{\substack{i \in S \\ k \neq j}} r_i^k p_{ij}$$

$$= \sum_i \in S r_i^k p_{ij}.$$

This proves (ii).
Finally, since $j \leftrightarrow k$ for every j and $k \in S$ and chain is irreducible (recurrent), thus there exists m such that $p_{ki}^m > 0$. Thus,

$$r_i^k > r_k^k p_{ki}^m = p_{ki}^m > 0.$$

(because $r_k^k = 1$) Also,

$$1 = r_k^k = r_i^k p_{ik}^m \forall i.$$

This implies $r_i^k < +\infty$. ■

4.3.2 Theorem (uniqueness):
Let $\{X_n\}_{n \geq 1}$ be an irreducible chain and λ is an invariant distribution for P with $\lambda_k = 1$. Then, $\lambda \geq \mathbf{r^k}$, where r^k is as defined in theorem above. If in addition P is recurrent, then $\lambda_\mathbf{1} = \mathbf{r^k}$.

Proof: Using invariance of $\lambda, \forall i$

$$\begin{aligned}
\lambda_i &= \sum_{i_1 \in S} \lambda_{i_1} p_{i_1} i. \\
&= \sum_{\substack{i_1 \in S \\ i_1 \neq k}} \lambda_{i_1} p_{i_1} i. + p_{ki} (\text{because} \lambda_k = 1) \\
&= \sum_{\substack{i_1 \in S \\ i_1 \neq k}} (\sum_{i_2} \in S \lambda_{i_2} p_{i_2} i_1) + p_{ki} \\
&= \sum_{\substack{i_1 \in S \\ i_1 \neq k}} (\sum_{\substack{i_2 \in S \\ i_2 \neq k}} \lambda_{i_2} p_{i_2} i_1 + p_{k_{i_1}}) p_{i_1} i + p_{ki} \\
&= \sum_{\substack{i_1, i_2 \in S \\ i_1 \neq k, i_2 \neq k}} \lambda_{i_2} p_{i_2} i_1 P_{i_1} i + (\sum_{\substack{i_1 \in S \\ i_1 \neq k}} p_{ki_1} p_{i_1}, i + p_{ki}) \\
&= - - - - \\
&\geq \sum_{n=0}^{m} (\sum_{i_1, i_2, \ldots, i_n \neq k} p_{ki_1} p_{i_n} p_{i_{n-1} i_{n-2}} \ldots p_{i,i}) \\
&= \sum_{n=0}^{m} P_k(X_n = i, T_k \geq n)
\end{aligned}$$

Since this holds $\forall n$,

$$\lambda_i \geq \sum_{n=0}^{\infty} p_k\{X_n = j, T_k \geq n\} = r_i^k.$$

In case, the chain is recurrent, we can select n such that $p_{jk}^n > 0$. Let $\mu = \lambda - r^k$. Then μ is also an invariant measure and $\mu_k = \lambda_k - r_k^k = 0$. Thus,

$$0 = \mu_k = \sum_{i \in S} \mu_i p_{ik}^n \geq \mu_j p_{jk}^n \geq 0.$$

implies

$$\mu_j = 0 \;\forall\, j, \qquad (\text{as } \; p_{jk}^n > 0).$$

To go from the invariant measure to a distribution, we need $\sum_{i=0}^{\infty} r_i^k < +\infty$. For this we make the following definition:

4.3.3 Definition:
Let $\{X_n\}_{n \geq 1}$ be a chain.

(i) Let $m_i = E_i(T_i < +\infty)$, be the **expected return time** for state i.

(ii) We say a recurrent state i is **positive recurrent** if $m_i < +\infty$. Otherwise, we call it **null recurrent.** Note that $m_i = \sum_{j \in S} r_j^i$.

We have the following theorem.

4.3.4 Theorem:
Let $\{X_n\}_{n \geq 1}$ be an irreducible chain. Then the following are equivalent.

(i) All the states are positive recurrent.

(ii) There exists a state that is positive recurrent.

(iii) There exists an invariant distribution π with the property $\pi_i = \dfrac{1}{m_i}$ $\forall i$.

Proof:
$(i) \Rightarrow (ii)$ is obvious.
$(ii) \Rightarrow (iii)$: If k is a positive recurrent state, consider $r_j^k, j \in S$ as constructed in first theorem. Since $m_k = \sum_{j \in S} r_j^k$ is finite, define

$$\pi_i = \frac{r_i^k}{m_k}.$$

Then, $\pi_{i\,i \in S}$ is an invariant distribution.
$(iii) \Rightarrow (i)$: Take any $k \in S$ as fixed. Since P is irreducible and $\sum_{i \in S} \pi_i = 1$,

$$\pi_k = \sum_{i \in S} \pi_{i_k}^n > 0 \text{ for some } n$$

Hence $\pi_k > 0 \;\forall k$. Define

$$\lambda_i = \frac{\pi_i}{\pi_k},\; k \geq 0.$$

Then $\lambda_i$is an invariant measure and $\lambda_k = 1$. Thus, by theorem 4.3.2, $\lambda \geq r_k$. Hence,

$$m_k = \sum_{i \in S} r_i^k \leq \sum_{i \in S} \lambda_i = \frac{\Sigma \pi_i}{\pi_k} = \frac{1}{\pi_k} < \infty. \tag{4.4}$$

Thus, k is **positive recurrent.** In fact, P recurrent implies (theorem 4.3.2) that $\lambda = r^k$. Hence, equation(4.4) says

$$m_k = \frac{1}{\lambda_k} \;\forall k.$$

■

4.3.5 Example (Random walk on the line):
Recall $i-1 \leftarrow^{q} i \rightarrow^{p} i+1$

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

(i) The walk is **transient** if $4pq < 1$.

(ii) For $p = q = \frac{1}{2}$: It is called the symmetric random walk, it is recurrent. Consider the measure $\pi_i = 1 \; \forall i \in S$. Then,

$$\pi_i = \frac{1}{2}\pi_{i-1} + \frac{1}{2}\pi_{i+1}.$$

Hence,$\Pi = (1,1)$is an invariant measure. In case, an invariant distribution exists it must be a scalar multiple of π, but $\Sigma\pi_i = +\infty$. Hence, there does not exist any stationary distribution. It is **null recurrent.**

4.3.6 Example (Asymmetric random walk):
Let

$$p_{i,i-1} = q < p = p_{i,i+1}.$$

Though each state is transient and theorem does not apply, let us try to find invariant distribution Π.. For this,

$$\begin{aligned} \Pi \text{ invariant} \quad & \Leftrightarrow \Pi P = \Pi \\ & \Leftrightarrow \pi_{i-1}p + \pi i + 1q \\ & = \pi_i \end{aligned}$$

This gives a recurrence relation, and a general solution can be found:

$$\pi_i = A + B\left(\frac{p}{q}\right)^i,$$

where A,B are arbitrary constants. This shows that invariant measure need not be unique.

4.3.7 Example(Simple symmetric Random walk on $\mathbb{Z}^2$):

$$\begin{array}{ccccccccc} & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \bullet & \cdots & \bullet & \cdots & \bullet & \cdots \\ & \vdots & 1/4 & \uparrow & 1/4 & & \\ \cdots & \bullet & \longleftarrow & \bullet & \longrightarrow & \bullet & \cdots \\ & \vdots & 1/4 & \downarrow & 1/4 & & \\ \cdots & \bullet & \cdots & \bullet & \cdots & \bullet & \cdots \\ & \vdots & & \vdots & & \vdots & \end{array}$$

$$P_{ij} = \begin{cases} 1/4 \text{ if } |i-j| = 1. \\ \\ 0 \text{ otherwise} \end{cases}$$

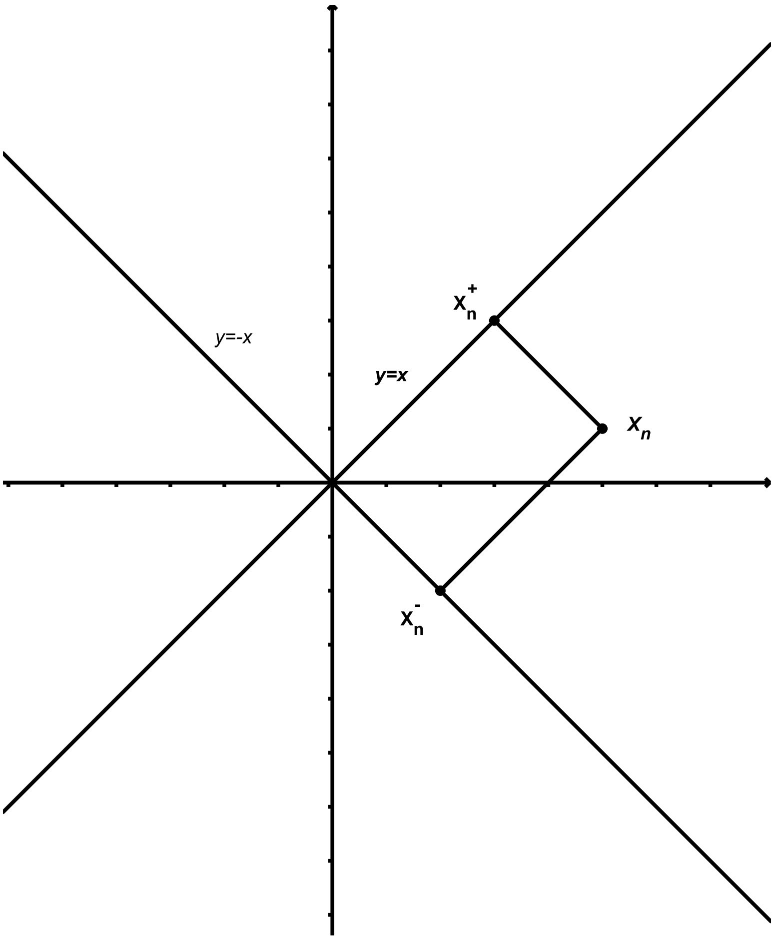
Then

$$P_{00}^{2n} = \left[\begin{pmatrix} 2n \\ n \end{pmatrix}\left(\frac{1}{2}\right)^{2n}\right]^2.$$

An intuitive way of seeing this is as follows: Consider X_n^+ as orthogonal projection of X_n onto $y = x$ and X_n^- as orthogonal projection onto $y = -x$. Then X_n^+ and X_n^-

are symmetric independent random variable on $\sqrt{2}\,\mathbb{Z}$ and

$$X_n = 0 \text{ iff } X_n^+ = 0 = X_n^- .$$



Now Sterling formula gives

$$p_{00}^{2n} \sim \frac{2}{A^2 n}$$

Hence, $\sum_{n=0}^{\infty} p_{00}^n = +\infty$, that is the symmetric random walk is recurrent.

4.3.8 Remark:

(i) Similar analysis is possible for random walks in $\mathbb{R}^3$.

(ii) For random walks on the line/plane:

$$p_{ij}^n \to 0; \text{as } n \to \infty \ \forall i, j.$$

4.3.9 Theorem (Existence for finite state space):
Let S be finite and

$$p_{ij}^n \to \pi_j \text{ as } n \to \infty \forall i$$

Then $(\pi_j)_{j\in S}$ is an invariant distribution.
Proof:
Note that

$$\begin{aligned}\Sigma_{j\in S}\pi_j &= \sum_{j\in S}(\lim_{n\to\infty}p_{ij}^n\ \forall i\\ &= \lim_{n\to\infty}(\sum_{j\in S}p_{ij}^n) = 1,\end{aligned}$$

because P is stochastic. And

$$\begin{aligned}\pi_j &= \lim_{n\to\infty}(p_{ij}^n) = \lim_{n\to\infty}(p_{ij}^{n+1}\\ &= \lim_{n\to\infty}(\sum_{k\in S}p_{ik}^n p_{kj})\\ &= \sum_{i\in S}(\lim_{n\to\infty}p_{ik}^n)p_{kj}\\ &= \sum_{i\in S}\ \pi_k\ p_{kj}.\end{aligned}$$

■

Question:
When can we generalize above theorem? Some answers are given in the next section. For more details see Billingsley [4]

4.4. Asymptotic behavior

4.4.1 Theorem:
Let $(X_n)_{n\geq 1}$ be an irreducible aperiodic chain for which a stationary distribution π exists. Then the chain is **persistent** *with*

$$\lim_{n\to\infty}p_{ij}^n = \pi_j\ \forall\ i,j.$$

Further, all $\pi_j > 0$ and the stationary distribution is unique.

4.4.2 Theorem:
Let $X_n)_{n\geq 1}$ be irreducible aperiodic chain for which no stationary distribution exists. Then

$$lim_{n\to\infty}p_{ij}^n = 0\ \forall\ i,j.$$

4.4.3 Classification of irreducible aperiodic chains:

(i) Transient. $\sum_n p_{ij}^n < +\infty$. This implies the general fact that $\lim_{n\to\infty}p_{ij}^n = 0$.

(ii) Recurrent. $\sum_n p_{ij}^n = \infty$. No stationary distribution and by theorem 4, $\lim_{n\to\infty}p_{ij}^n = 0$.

(iii) $\exists$ stationary distribution. Hence, positive recurrent. This implies that

$$p_{ij}^n \to \pi_j > 0.$$

Diagonalization of matrices

Let A be a $n \times n$ matrix with entries from $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

A.1.1 Definition: A matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, i.e., there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

We would like to know when is a given matrix A diagonalizable and if so, how to find P such that $P^{-1}AP$ is diagonal? Next theorem answers this question.

A.1.2 Theorem:
Let A be a $n \times n$ matrix. If A is diagonalizable, then there exist scalars

$$\lambda_1, \lambda_2, \ldots, \lambda_n \;\; \text{in} \;\; \mathbb{F}$$

and vectors

$$\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_n \;\; \text{in} \;\; \mathbb{F}^n$$

be such that the following holds:

(i) $A\boldsymbol{C}_i = \lambda_i \boldsymbol{C}_i$ *for all* $1 \leq i \leq n$. *That is A has n-eigenvalues.*

(ii) The set $\{\boldsymbol{C}_1, \ldots, \boldsymbol{C}_n\}$ *is linearly independent, and hence is a basis of* $\mathbb{F}^n$.

Theorem A.1.2 says that if A is diagonalizable then not only A has n eigenvalues, it has a basis consisting of eigenvectors. In fact, the converse of this is also true.

A.1.3 Theorem:

(i) Let A be a $n \times n$ matrix. If A has n linearly independent eigenvectors, then A is diagonalizable.

(ii) Let A be a $n \times n$ matrix. If A has n distinct eigenvalues, then A is diagonalizable .

(iii) If A is real symmetric, then there exists an orthogonal matric P such that PAP^{-1} is diagonal.

A.1.4 Note:
Theorem 10.1.3 not only tells us when is A diagonalizable, it also gives us a matrix P which will diagonalize A, i.e., $P^{-1}AP = D$, and gives us the resulting diagonal matrix also. The columns vectors of the matrix P are the n eigenvectors of A and the diagonal matrix D the diagonal entries as the eigenvalues of A corresponding to these n-eigenvectors.

For more details refer ”From Geometry to Algebra- An Introduction to Linear Algebra” by Inder K. Rana, Ane Books, New Delhi, 2010.

References

Markov chains:

[1] 'Finite micro chains'
- Kemeny and Sneel
Springer-verlag

[2] 'Introduction to stochastic processes'
- Hoel,Port and Stone
Houghton Mifflin Company

[3] 'A first course in stationary process'
- Karlin and Taylor
Academic Press

Probability and Measure:

[4] 'Introduction to probability and measure'
- P.Billingsley.

[5] 'Introduction to probability and measure'
- K.R. Parthasaathy.

Measure theory:

[6] 'An Introduction to Measure and Integration'
- Inder K. Rana.
Narosa Publishers

Linear Algebra

[7] 'From Geometry to Algebra- An Introduction to Linear Algebra'
- Inder K. Rana,
Ane Books, New Delhi, 2010.

Index