	Kingdom of Saudi Arabia **AL-Imam Muhammad bin Saud Islamic University** **Faculty of science**	**المملكة العربية السعودية** **جامعة الامام محمد بن سعود الإسلامية** **كلية العلوم**	

MARKOV CHAINS AND MATRICES

A research project in mathematics (Mat 499)

First semester 1444

Written by

Majid Alharbi

Supervised by

Prof. Abdullah Aljouiee

CONTENTS

Introduction

Markov chains have several advantages as mathematical models. They are general enough to provide useful models for many situations. They have many results that are known, and they are easy to use. The diversity of applications can be illustrated with a few examples. Some of which are in Psychology learning, Demography, Biology, Ecological systems and more. The ease of use is a consequence of using matrix algebra.

CHAPTER 1

STATES AND TRANSITION MATRICES

Stochastic process deal with number of successive steps or stages, each stage has different states and. In general, the probabilities of the results at one stage depend on the results of preceding stages and this dependence can take many forms. For example, tree diagrams for three experiments each consisting of repetitions of a sub-experiment with results labeled X and Y are shown in figure1. Notice that in each stage the sum of the probabilities to the right of the stage is 1.

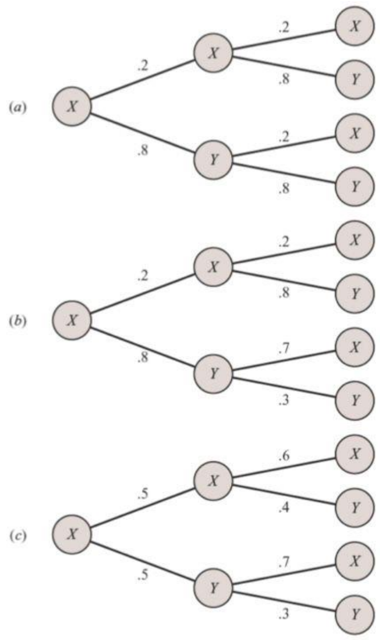

Figure1

Definition 1.1

A Markov chain is a stochastic process which satisfies the condition:

1. At each stage the result is one of a fixed number of states.
2. The conditional probability of a transition from any given state to any other state depends on only the two states satisfy the following conditions:

(i) At each stage the result is one of a fixed number of states

(ii) The conditional probability of a transition from any given state to any other state depends on only the two states.

Example 1.2

A freelance computer network consultant is employed only when she has a contract for work, and each of her contracts is for 1 week of work. Each week she is either employed (E) or unemployed (U) and her records support the following assumptions about the conditional probabilities

(a) If she is employed this week then next week, she will be employed with probability 0.8 and unemployed with probability 0.2.

(b) If she is unemployed this week then next week, she will be employed with probability 0.6 and unemployed with probability 0.4. (see Figure2)

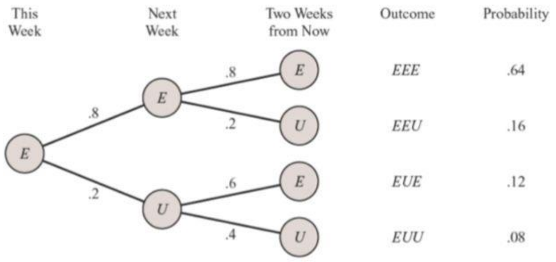

Figure2

If she is employed this week, what is the probability that she will be employed 2 weeks from now? A tree diagram for this situation is shown in Figure2. Since we know that she is employed this week the Begin box of the tree diagram is replaced by E. From the tree diagram we conclude that the probability that she is employed 2 weeks from now is

$$0.64+0.12 = 0.76$$

However, if we had asked for the probability that she will be employed in 5 weeks or in 10 weeks or in the long run, then we need to develop techniques which will be more effective in such problems. We need to have new terminology and notation and will be concerned with systems which can be in any one of N possible states.

In Figure2, we have indicated states E and U and on arrows connecting the states, the probabilities of being in successive states on successive observations. For instance, the 0.2 on the arrow directed from E to U means that if the system is in state E on one observation, then it is in state U on the next observation with probability 0.2. As we noted in the definition of Markov chains, the fundamental property which distinguishes a Markov chain from other sequential probabilistic processes can be described as follows.

Markov Property

If a system is in state on one observation, then the conditional probability that it is in state on the next observation depends on only and not on what happened before the system reached state or on the stage of the experiment. This probability will be denoted by P_{ij} , the probability of making a direct transition from i to j.

One way that this property can be viewed intuitively is to think of Markov chains as mathematical descriptions for systems without memories. That is the probability that the system makes a transition from one state (say, state i) to another state (or even back into state) depends on the two states and not on the number of transitions or the states occupied before the system reached state i. The transition probabilities P_{ij} are the numbers on the arrows of the transition diagram. Thus, in Example 1.2, if we label state E by 1 and state U by 2, then P_{11}= 0.8 and P_{12}= 0.2.

Transition Matrix

Consider a Markov chain with N states. Let P_{ij} be the probability of making a direct transition from state i to state j. So, we have $1 \leq i \leq N$, $1 \leq j \leq N$. Consider the matrix $P = \left[P_{ij}\right]$. This matrix is called the transition matrix for the Markov chain.

Example 1.3

Find the transition matrix for the process described in Example 1.2.

To solve this problem, let E and U be labeled as states 1 and 2, respectively. The transition probabilities are given on the transition diagram in Figure3.

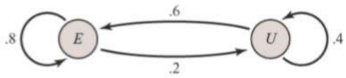

Figure3

And the matrix is

$$P = \begin{bmatrix} .8 & .2 \\ .6 & .4 \end{bmatrix}.$$

Note that if we had labelled the states differently, then we would have obtained a different transition matrix P. However, providing the identification of states with rows and columns is consistent, then all probabilities computed for the process will be the same, regardless of the transition matrix used.

Example 1.4.

Consider the tree diagram in Figure4. Is this a diagram of a Markov chain?

The states are labelled A and B in Figure4. Note that the probability of making a transition from A to A at stage I is $\frac{1}{3}$, the probability of making a transition from A to A at the second stage is $\frac{1}{3}$, and the probability of making a transition from A to A at the third stage is $\frac{1}{5}$. Thus, the probability of making a transition from A to A does depend on the stage, and consequently this is not the tree diagram of a Markov chain.

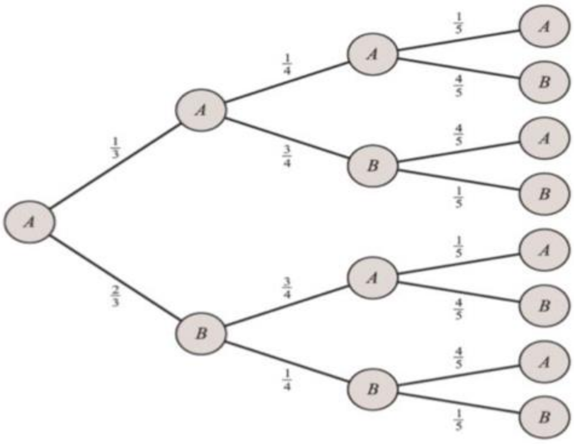

Figure4

Example 1.5.

The dispatcher in the office of the White Wheel Taxi Company frequently contacts a driver who is away from the office by radio with the name and address of the next customer. For reasons of efficiency the dispatcher attempts to contact a driver who is, or who will be, in the same area as the person requesting a taxi. Of course, it may not be possible to do so. One of the drivers keeps a record of radio dispatches for a week, and the data are as in table1.

TABLE 8.1

Current Location of Driver	Location of Next Rider	Percentage of Messages
East	East	50
	Central	40
	West	10
Central	East	10
	Central	60
	West	30
West	East	30
	Central	60
	West	10

Table1

To formulate this as a Markov chain, we must identify the states of the system. We suppose that the service area can be divided into three districts: East, Central and West. If the driver is in the East district, then we say that the system is in state 1. Similarly, the system is in states 2 or 3 when the driver is in the Central or West district, respectively. Transitions are the moves of the driver that result from calls by the dispatchers who provide locations of new customers. Using the data contained in Table1, we have the transition diagram shown in Figure5, and the transition matrix P.

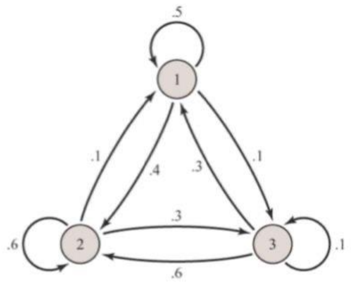

Figure5

$$P = \begin{bmatrix} .5 & .4 & .1 \\ .1 & .6 & .3 \\ .3 & .6 & .1 \end{bmatrix}$$

Decision problems of resource allocation and scheduling such as the dispatching of taxis in Example 8.4 arise in many different settings and they can be studied using a variety of techniques The approach used normally depends on the specific goals and constraints of the study (resources available, customer service expectations, cost, etc.) For instance, rather than depending only on driver location the dispatcher may consider the priority of the customer the type of vehicle each driver has, the experience of the driver whether the driver s shift is about to end and so forth.

CHAPTER 2

BASIC PROPERTIES OF MARKOV CHAINS

We have seen that information about transitions in a Markov chain whose transitions are possible, and their probabilities can be given in a transition diagram or a transition matrix. For computational purposes, the transition matrix is often the most useful. The transition matrix for a Markov chain with N states, introduced in chapter 1, is an N × N matrix whose (i, j) entry is the probability of a transition from state i to state j in one step. There are corresponding probabilities for transitions from one state to another in k steps; these are usually called k-step transition probabilities. The conditional probability of making a transition from state i to state j in exactly k steps is denoted by $P_{ij}(k)$. The matrix whose (i , j) entry is $P_{ij}(k)$ is denoted by P(k) and will be called the k-step transition matrix for a Markov chain.

Example 2.1.

Use tree diagrams to find the two-step transition matrix P (2) for the transition matrix

$$P = \begin{bmatrix} .8 & .2 \\ .6 & .4 \end{bmatrix}.$$

To determine the first row of P (2), we use a tree diagram which represents a two-stage experiment in which the system is initially in state 1. Such a tree diagram is shown in Figure6a.

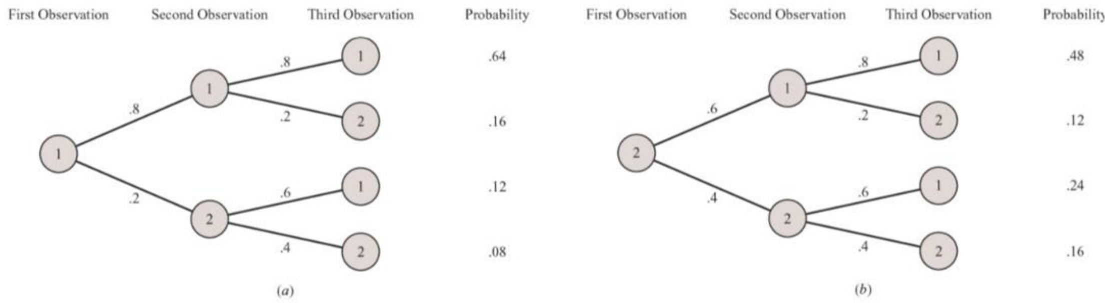

Figure6

From this tree diagram we see that a transition from state 1 to state 1 in two steps can occur in either of two ways, and the probability p_{11} (2) is the sum

$$P_{11}(2) = .64+12= .76$$

Similarly

$$P_{12}\ (2) =.16+.08=.24$$

To determine the second row of P(2), we use a tree diagram in which the system is initially in state 2, Figure6b in this case. We have

$$P_{21}(2) =.48+.24=.72$$
$$P_{22}(2) =.12+.16=.28$$

Therefore

$$P\ (2) = \begin{bmatrix} P_{11}(2) & P_{12}(2) \\ P_{21}(2) & P_{22}(2) \end{bmatrix} = \begin{bmatrix} .76 & .24 \\ .72 & .28 \end{bmatrix}$$

The technique illustrated in Example 8.5 can be used to construct P (2) for any Markov chain for which the transition matrix P can be determined. However, for large matrices the process is cumbersome, and one of the very useful properties of Markov chains is that there is a simple method of finding P (2) from P without using tree diagrams. The idea behind the method can be seen by looking more carefully at the calculation of $P_{12}(2)$ in Example 2.1. Using the tree diagram, we found

P_{12} (2) = .8(.2)+.2(.4)= $P_{11}P_{12} + P_{12}P_{22}$

This last expression is exactly the (1,2) entry in the matrix product PP.

Now let us look at a similar argument in the case of a Markov chain with N states and transition matrix P. We find an expression for $P_{ij}(2)$: the probability that if the system is in state i on one observation, then it is in state j on the second subsequent observation. The system can move from state i to state j in two steps by moving from i to 1 to j. This happens with probability $P_{i1}\ P_{1i}$. Recall that the probability that the system makes a transition from state 1 to state j in one step is independent of the states it occupied before state 1. Likewise, the system can move from state i to state j through any of states 2, 3,..., N. These events happen with probabilities $P_{i2}\ P_{2j}\ , P_{i3}\ P_{3j}\ , \dots , P_{iN}\ P_{Nj}$, respectively. Since the system must move from state i to state j through exactly one intermediate state, we have

$$P_{ij}\ (2) = P_{i1}P_{1j} + P_{i2}P_{2j} + P_{i3}P_{3j} + \cdots + P_{iN}P_{Nj}$$

This expression for $P_{ij}(2)$ is exactly the (i, j) entry in the matrix product PP. We have the result

$$P(2) = PP = P^2$$

This is a special case of the following more general result. And we come up with the following statement.

Theorem 2.2.

Let P be the (one-step) transition matrix for a Markov chain. Then the matrix P(k) of k-step transition probabilities is

$$P(k) = P^k.$$

Example 2.3.

To compute P (2) using Theorem 2.2, we just multiply P by itself to get

$$P(2) = \begin{bmatrix} .8 & .2 \\ .6 & .4 \end{bmatrix}^2 = \begin{bmatrix} .8 & .2 \\ .6 & .4 \end{bmatrix} . \begin{bmatrix} .8 & .2 \\ .6 & .4 \end{bmatrix} = \begin{bmatrix} .76 & .24 \\ .72 & .28 \end{bmatrix}.$$

Notice that this is exactly what we have got earlier.

Example 2.4.

A Markov chain has transition matrix

$$P = \begin{bmatrix} .3 & .3 & .4 \\ .5 & .5 & .0 \\ 1 & 0 & 0 \end{bmatrix}$$

Find the two-step transition matrix P(2). And If the system is initially observed in state 1, what is the probability that it is in state 1 two observations later?

We apply Theorem2.2 to get

$$P(2) = P^2 = \begin{bmatrix} .3 & .3 & .4 \\ .5 & .5 & 0 \\ 1 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} .64 & .24 & .12 \\ .4 & .4 & .2 \\ .3 & .3 & .4 \end{bmatrix}$$

The second question is merely $P_{11}(2)$, that is 0.64.

Example 2.5.

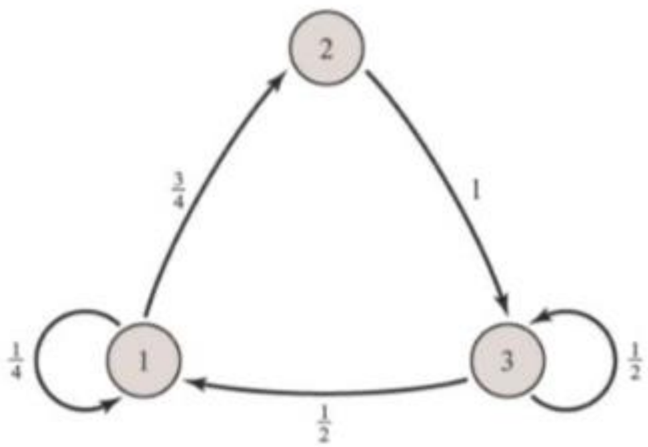

Figure7

This transition diagram for a Markov chain has a transition matrix

Problem

Find the matrix of three-step transition probabilities for this Markov chain

$$P = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

And the three-state transition matrix is

$$P(3) = P^3 = PP^2 = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{16} & \frac{3}{16} & \frac{3}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{3}{8} & \frac{3}{8} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{25}{64} & \frac{3}{64} & \frac{36}{64} \\ \frac{3}{8} & \frac{3}{8} & \frac{1}{4} \\ \frac{14}{64} & \frac{18}{64} & \frac{1}{2} \end{bmatrix}.$$

A probability vector is a vector with nonnegative coordinates for which the sum of the coordinates is 1 and apparently each row of a transition matrix P of a Markov chain is a probability vector. Likewise, for each k the rows of matrix P(k) are probability vectors. Indeed, the entries of the ith row of P(k) are the probabilities $P_{i1}(k), P_{i2}(k), \ldots.., P_{iN}(k)$, and the system must move from state i to some state (perhaps i itself) in k steps.

Theorem 2.6.

If P is a matrix whose rows are probability vectors, then P^k has probability vectors.

CHAPTER 3

REGULAR MARKOV CHAINS

In the preceding chapters, we introduced the basic property of Markov chains, and we developed the formula $P(k) = P^k$, which shows that the k-step transition matrix equals the kth power of the 1-step transition matrix. In this chapter and the next, we continue by considering briefly two special types of Markov chains which are especially useful in applications, and we give such an application at the end of the chapter. Since Markov chains are stochastic processes, we do not ordinarily know what will happen at each stage, and we must describe the system in terms of probabilities.

State Vector

Consider a Markov chain with N states. A state vector for the Markov chain is a probability N vector $X = [x_1 \quad x_2 \quad \dots \quad x_N]$. The ith coordinate x_i of the state vector X is to be interpreted as the probability that the system is in state i. We write a state vector as a row vector.

Theorem 3.1

If X_k and X_{k+1} denote the state vectors which describe a Markov chain after k and $k+1$ transitions, respectively, then $X_{k+1} = X_k P$,where P is the transition matrix of the chain. That is, the state vector X_k, which describes the system after k transitions is the product of the initial state vector and the kth power of the transition matrix.

Example 3.2

A Markov chain has the transition matrix $P = \begin{bmatrix} .5 & .5 \\ .8 & .2 \end{bmatrix}$. If the system begins in state 2, find the state vector after two transitions.

The initial state vector is $X_0 = [0 \quad 1]$. According to the Theorem 3.1, the state vector after two transitions is

$$X_2 = X_0 P^2 = [\,1 \quad 0\,] \begin{bmatrix} .5 & .5 \\ .8 & .2 \end{bmatrix} \begin{bmatrix} .5 & .5 \\ .8 & .2 \end{bmatrix} = [.56 \quad .44\,]$$

There are various ways of classifying Markov chains, and we choose one which distinguishes among chains based on their long-run behaviour, i.e., on the behaviour of the state vector after many transitions. As we will see in our examples, the long-run behaviour of the state vector may provide important information in applications. Also, after X_k has been determined for some value of *k*, in general this does not provide much information about X_{k+1} or X_{k+2} without further computation. Thus, if you are interested in studying a stochastic process over many transitions, then it is appropriate to develop some tools for determining its long-run behaviour.

Regular Markov Chain

Definition 3.3

A Markov chain with transition matrix P is regular if there is a positive integer k such that P^k has all positive entries.

Example 3.4

Markov chains associated with $P_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$ is regular since

$$P_1^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

And apparently, all its entries are positive. While Markov chains associated with

$$P_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is not regular since

$$P_2^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I\ .$$

And this means P_2^n contains a zero entry for every *n*.

The definition of a regular chain (although stated in terms of the powers of P) has the following important consequence. For each j and for k sufficiently large, each of the transition probabilities $P_{1j}(k), P_{2j}(k), \dots, P_{Nj}(k)$ is close to the same number, call it q_j. That is, each of the entries in the jth column of the k-step transition matrix $P(k)$ is close to P_j. Another way of saying this is that for large values of *k*, the *k*-step transition matrix

$$P(k) = \begin{bmatrix} P_{11}(k) & P_{12}(k) & P_{1N}(k) \\ P_{21}(k) & P_{22}(k) & P_{2N}(k) \\ . & . & . \\ . & . & . \\ P_{N1}(k) & P_{N2}(k) & P_{NN}(k) \end{bmatrix}$$

is very close to a matrix that has all rows identical

$$\begin{bmatrix} W \\ W \\ . \\ . \\ . \\ W \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & \dots w_N \\ w_1 & w_2 & \dots\ w_N \\ . & . & . \\ . & . & . \\ . & . & . \\ w_1 & w_2 & w_N \end{bmatrix}$$

where $W = [w_1 \quad w_2 \quad w_N\ ]$

Example 3.6

Consider the following transition matrix

$$P = \begin{bmatrix} .5 & .4 & .1 \\ .1 & .6 & .3 \\ .3 & .6 & .1 \end{bmatrix}$$

A straightforward computation of the *k*-step transition matrices (best carried out on a computer) gives

$$P(2) = \begin{bmatrix} .32 & .50 & .18 \\ .20 & .58 & .22 \\ .24 & .54 & .22 \end{bmatrix}$$

$$P(4) = \begin{bmatrix} .2456 & .5472 & .2072 \\ .2328 & .5552 & .2120 \\ .2376 & .5520 & .2104 \end{bmatrix}$$

$$P(8) = \begin{bmatrix} .2369 & .5526 & .2105 \\ .2368 & .5526 & .2105 \\ .2369 & .5526 & .2015 \end{bmatrix}$$

where the entries have been rounded off to the four decimal places shown. It is now clear that the rows of P(8) are essentially equal. This illustrates the assertion that as *k* increases, the *k*-step transition matrix $P(k)$ becomes closer and closer to a matrix all of whose rows are equal to the same vector W.

The rows in a transition matrix P, and those of its powers P^k, are all probability vectors. For regular chains the rows in P^k all become closer and closer to the same probability vector as *k* increases. This special probability vector is determined by P and is called a stable vector for P. The following theorem summarises these concepts.

Theorem 3.7

Let P be the transition matrix for a regular Markov chain. There is a unique probability vector $W = [w_1 \quad w_2 \quad ... \; w_N]$ such that for each state *j* the difference $\left|P_{ij}(k) - w_j\right|$ can be made as small as we choose by selecting *k* sufficiently large. The vector W is known as a stable vector, and its coordinates are known as stable probabilities for the Markov chain.

In a regular Markov chain, the probabilities $P_{ij}(k)$ are for all large values of k nearly equal to the stable probabilities w_j. This assertion holds for each initial state $i, i = 1,2, \dots, N$. The stable probabilities W_j can be obtained from the vector W, which is closely approximated by any row of $P(k)$ for large values of k. However, obtaining W from $P(k)$ usually requires computing P^k for several large values of k, a method that may be impractical. Fortunately, there is an alternative method of obtaining the stable probabilities.

Theorem 3.8

Let P be the transition matrix of a regular Markov chain. Then there is a unique probability vector W which satisfies $WP = W$ The coordinates of this vector are the stable probabilities for the Markov chain.

This theorem provides a direct method of obtaining the stable probabilities. Indeed, we need only solve a system of linear equations

Example 3.9

The matrix $P = \begin{bmatrix} .25 & .75 \\ .60 & .40 \end{bmatrix}$ is the transition matrix of a regular Markov chain. To determine the vector W of stable probabilities for this Markov chain. we make use of the theorem quoted above. That is, we find the probability vector W which satisfies the system of equations.

If $W = [w_1 \quad w_2 \ ]$, then the condition that W be a probability vector requires that $w_1 + w_2 = 1$, and the system $W(P - I\ ) = 0$ or

$$[w_1 \ \ w_2 \ ]\begin{bmatrix} -.75 & .75 \\ .60 & -.60 \end{bmatrix} = [0 \quad 0\,]$$

$$-.75\, w_1 + .60\, w_2 = 0$$

$$.75\, w_1 - .60\, w_2 = 0$$

Notice that the last two equations are equivalent, and thus we are left with the system of equations

$$w_1 + w_2 = 1$$

$$.75\, w_1 - .60\, w_2 = 0$$

Whose solution is $w_1 = \frac{4}{9}$ and $w_2 = \frac{5}{9}$.

So, the required vector of stable probabilities for the Markov chain is

$$W = \begin{bmatrix} \frac{4}{9} & \frac{5}{9} \end{bmatrix}.$$

Example 3.10

To find the stable probabilities for the Markov chain whose transition matrix is

$$P = \begin{bmatrix} .5 & .5 & 0 \\ 0 & .5 & .5 \\ .75 & .25 & 0 \end{bmatrix}$$

We solve the system

$$[w_1 \;\; w_2 \;\; w_3] \begin{bmatrix} -.5 & .5 & 0 \\ 0 & -.5 & .5 \\ .75 & .25 & -1 \end{bmatrix} = [0 \;\; 0 \;\; 0]$$

In addition to the condition that W be a probability vector which is

$$w_1 + w_2 + w_3 = 1$$

A regular method will lead to the solution $W = \begin{bmatrix} \frac{3}{9} & \frac{4}{9} & \frac{2}{9} \end{bmatrix}$.

We now have a means of computing the vector of stable probabilities for any regular Markov chain. To show that a Markov chain is regular, we must be able to show that some power of the transition matrix has all positive entries. It is important to note that we do not need to know the actual entries of the power of the matrix. We only need to know that the entries are all positive.

Definition 3.11

The ith state of a Markov chain is said to be an absorbing state if $P_{ii} = 1 \; and \; P_{ij} = 0 \; for \; j \neq i$. That is, state i is absorbing if the ith row of the transition matrix is the ith unit vector.

Example 3.12

In the Markov chain whose transition matrix is shown below, the second state is absorbing. Note, however, that the fifth state is not absorbing. Even though the fifth row contains a single 1, it is not in the fifth column, so the fifth row is not the ith unit vector.

$$\begin{bmatrix} \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Absorbing chains must have absorbing states, but that is not enough. It must also be possible to go from non-absorbing states to absorbing states

Definition 3.13

Absorbing Markov chain is a Markov chain is said to be absorbing if
(a) There is at least one absorbing state, and

(b) For every non-absorbing state i there is some absorbing state j and a positive integer k such that the probability of a transition from state i to state j in k steps is positive.

Example 3.14

The matrix P is the transition matrix of a Markov chain

$$P = \begin{bmatrix} \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Here, state 2 is an absorbing state, so condition (a) of the definition of an absorbing Markov chain is satisfied. Also $P_{32} > 0 \ and \ P_{12}(2) > 0$, so condition (b) is satisfied. Therefore, P is the transition matrix of an absorbing Markov chain.

To write the transition matrix in canonical form, we relabel the states so that absorbing states are listed first. In this example the only absorbing state is state 2. We relabel the states so that the original state 2 is relabeled as state 1. The labels assigned to the remaining states are not important. We relabel the states as follows:

Old label	1	2	3
New label	2	1	3

After the states are relabeled, the transition matrix changes to reflect the new labels. For instance, the old (1, 3) entry $P_{13} = \frac{2}{3}$ becomes the new (2, 3) entry; the old (2, 2) entry $P_{22} = 1$ becomes the new (1, 1) entry, and so on. The transition matrix written in canonical form is

$$P' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

This allows us to write the canonical form of P in the following form which is not unique.

$$P' = \begin{bmatrix} I & O \\ R & Q \end{bmatrix}$$

Here, I is a $k \times k$ identity matrix, O is a matrix with all zeros, and R and Q consist of transition probabilities which correspond to transitions which lead directly to absorption, R, and transitions which do not lead directly to absorption, Q. In Example 3.14, matrix P' is in the canonical form with a 1×1 identity matrix $I = [1]$ and with

$$R = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix} \text{ and } Q = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Definition 3.15

In the canonical form $P' = \begin{bmatrix} I & O \\ R & Q \end{bmatrix}$, we call the matrix $F := (I - Q\ )^{-1}$ the fundamental Matrix.

Remark 3.15

It can be shown that the matrix $(I - Q\ )$ is always invertible.

Example 3.16

An absorbing Markov chain has the transition matrix which is written in a canonical form

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{bmatrix}.$$

We have

$$Q = \begin{bmatrix} \frac{2}{3} & 0 \\ \frac{2}{5} & \frac{2}{5} \end{bmatrix}, \qquad I - Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} & 0 \\ \frac{2}{5} & \frac{2}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 \\ -\frac{2}{5} & \frac{3}{5} \end{bmatrix}$$

So,

$$F = (I - Q)^{-1} = \begin{bmatrix} 3 & 0 \\ 2 & \frac{5}{3} \end{bmatrix}.$$

Theorem 3.17

The (i, j) entry in the fundamental matrix F gives the expected number of times that a system which begins in the ith non-absorbing state will be in the *j*th non-absorbing state before it reaches an absorbing state. The sum of the entries in the ith row of F gives the expected number of transitions of a system which begins in the jth nonabsorbing state and continues until it first reaches an absorbing state.

The fundamental matrix can also be used to obtain other types of information about the system. For example, if the system has a single absorbing state, then the system will eventually reach that state, but if there is more than one absorbing state, then it may in general be absorbed in any one of them. Given the state in which the system begins, the likelihood that it will be absorbed *i* various absorbing states can be computed by using the fundamental matrix.

CHAPTER 4

An application

In many areas in temperate climates there is a natural progression as time passes: from open meadow grasslands through brush of vegetation to young forests and eventually to mature forests. Even in the absence of interference by humans, there are events which significantly alter, and sometimes reverse the progression. Such an event, and a very important one, is fire. Fires arise naturally through lightning strikes and are an important contributor to the perpetuation of grasslands. The natural progression of vegetation, and especially the occurrence of fires, is influenced by random events and therefore can be modelled by using stochastic processes. In addition to fires, other random events influencing vegetation include the introduction of seeds of plants not currently represented, the amount and timing of rainfall, the feeding habits of wildlife, and similar natural events. Under certain circumstances, Markov chains are appropriate as models for such situations. To construct a Markov chain model for plant succession, we focus on a single area which we suppose small enough that it can be classified into exactly one of four states. The state is determined by the dominant vegetation form: grassland (G), brush and shrubs (B), young forest (YF), and mature forest (MF). We suppose that the area is observed every decade and that the state or character of the area is noted at each observation. For the moment, we consider the progression from state to state in the absence of fire. We assume that the period between observations is such that progression proceeds at most one step between successive observations. That is, if the area is grassland at one observation, then at the next observation it is either grassland or brush and shrubs, and so on for other plant types. We also assume that once an area becomes a mature forest, it remains so throughout the time of observation. Continuing to consider the situation in the absence of fires, suppose the data support the assumption that one-step (10-year) transition probabilities are as follows:

$$\Pr[G|G] = .7 \; and \; \Pr[B|G] = .3$$

$$\Pr[B|B\ ] = .8\ and\ \Pr[YF|B] = .2$$

$$\Pr[YF|YF] = .5\ and\ \Pr[MF|YF] = .5$$

$$\Pr[MF|MF] = 1$$

Next, we turn to what happens when there is a fire. We consider only fires which are severe enough to cause an area to revert to grassland. Fires which have no effect on the state of the system or which cause the system to revert to a state other than grassland are not considered in this model. Suppose that in any decade a fire which reverts the area to grassland occurs with probability .1. We note that if the occurrence of such fires can be viewed as a Bernoulli process, then this assumption leads to the conclusion that the expected number of such fires is one per century. It follows that, in each decade, with probability .9 the area does not have a fire which affects the state of the system. Combining our assumptions about the occurrence of fires and what happens in the two situations (no fire and fire), we find that between two successive observations we have the following:

Probability of transition from grassland to grassland

$$\Pr[G|G\ and\ no\ fire]\ .\Pr[no\ fire\ ] + \Pr[G|G\ and\ fire]\ .\Pr[fire]$$
$$= .7(.9) + 1(.1) = .73$$

This approach also yields transition probabilities for the other possible transitions. For example,

Probability of transition from grassland to brush

$$\Pr[B|G\ and\ no\ fire].Pr[no\ fire] + Pr[B|G\ and\ fire]\ .Pr[fire]$$
$$= .3(.9) + 0(.1) = .27$$

Probability of transition from brush to grassland

$$\Pr[G|B\ and\ no\ fire].\Pr[no\ fire] + \Pr[G|B\ and\ fire\ ].\Pr[fire]$$
$$= 0(.9) + 1(.1) = .1$$

Probability of transition from brush to brush

$$\Pr[B|B\ and\ no\ fire].\Pr[no\ fire] + \Pr[B|B\ and\ fire].\Pr[fire]$$
$$= .8(.9) + 0(.1) = .72$$

Probability of transition from brush to young forest

$$\Pr[YF|B \ and \ no \ fire].\Pr[no \ fire] + \Pr[YF|B \ and \ fire].\Pr[fire]$$
$$= .2(.9) + 0(.1) = .18$$

We now have the entries in the first two rows of the transition matrix for this Markov chain. Entries in the last two rows can be determined by using the same approach. We collect all this information in a transition matrix which describes the plant succession, including the possibility of fire. The transition matrix P, with states as the types of vegetation which dominate the area, is

$$\begin{matrix} & G & B & YF & MF \end{matrix}$$
$$P = \begin{bmatrix} .73 & .27 & 0 & 0 \\ .1 & .72 & .18 & 0 \\ .1 & 0 & .45 & .45 \\ .1 & 0 & 0 & .9 \end{bmatrix}.$$

Suppose that the succession of vegetation is described by this model and that the area is observed over many decades. What is the probability that over the long run it is grassland? Since a power of the matrix P has all positive entries, this is a regular Markov chain, and we can answer the question by determining the stable vector for the chain. That is, we find the unique probability vector W which satisfies the equation $WP = W$. The system of equations $WP = W$

$$[w_1 \ \ w_2 \ \ w_3 \ \ w_4] \begin{bmatrix} .73 & .27 & 0 & 0 \\ .1 & .72 & .18 & 0 \\ .1 & 0 & .45 & .45 \\ .1 & 0 & 0 & .9 \end{bmatrix} = [w_1 \ \ w_2 \ \ w_3 \ \ w_4]$$

Or

$$.27w_1 - .1w_2 - .1w_3 - .1w_4 = 0$$
$$.27w_1 - .28w_2 = 0$$
$$.18w_2 - .55w_3 = 0$$
$$.45w_3 - .1w_4 = 0$$

The solution of this system which satisfies the additional condition $w_1 + w_2 + w_3 + w_4 = 1$ is

$$W = [w_1 \ \ w_2 \ \ w_3 \ \ w_4] = [.270 \ \ .261 \ \ .085 \ \ .384]$$

From this we conclude that over the long term we would expect the area to be grassland about 27 percent of the time, brush about 26.1 percent of the time, young forest about 8.5 percent of the time, and mature forest about 38.4 percent of the time.

Next suppose that through intervention it is possible to control fires if a decision is made to do so. Also suppose it is social policy to control fires in mature forests, and therefore once the area reaches a mature forest it remains in that state. In this situation the transition matrix becomes

$$\begin{array}{cccc} G & B & YF & MF \end{array}$$

$$\begin{bmatrix} .73 & .27 & 0 & 0 \\ .1 & .72 & .18 & 0 \\ .1 & 0 & .45 & .45 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This is the transition matrix of an absorbing Markov chain. Suppose that the area is initially a grassland. How many years before it becomes a mature forest?

To answer the question, we determine the expected number of transitions required for the system to first reach state MF, given that it began in state G. Labelling non-absorbing states in the order G, B, YF, the matrices Q and F (as defined in the third chapter) are

$$Q = \begin{bmatrix} .73 & .27 & 0 \\ .1 & .27 & .18 \\ .1 & 0 & .45 \\ \end{bmatrix}$$

$$F = \begin{bmatrix} 7.04 & 6.79 & 2.22 \\ 3.34 & 6.79 & 2.22 \\ 1.28 & 1.23 & 2.22 \end{bmatrix}$$

From this we conclude that if the system is initially in state 1 (grassland), then the expected number of transitions until the system first reaches state 4 (mature forest) is $7.04 + 6.79 + 2.22 = 16.05$, and consequently the expected number of years is about 160.

References

1- Ching,W and others, ***Markov Chains***, ***Models, Algorithms and Applications***, Springer (2013).

2- Maki, D and others, ***Finite Mathematics***, 6th Ed, Mc Graw Hill (2017).

3- Modica,G & Poggiolini, L, **A First Course in Probability and Markov Chains,** Wiley, (2013).

4- Privault, N, ***Understanding Markov Chains, Examples and Applications***, Springer (2013).