Chapter III
Markov Chains: Introduction

1. Definitions

A *Markov process* $\{X_t\}$ is a stochastic process with the property that, given the value of X_t, the values of X_s for $s > t$ are not influenced by the values of X_u for $u < t$. In words, the probability of any particular future behavior of the process, when its current state is known exactly, is not altered by additional knowledge concerning its past behavior. A *discrete-time Markov chain* is a Markov process whose state space is a finite or countable set, and whose (time) index set is $T = (0, 1, 2, \ldots)$. In formal terms, the Markov property is that

$$\Pr\{X_{n+1} = j | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i\}$$

$$= \Pr\{X_{n+1} = j | X_n = i\} \tag{1.1}$$

for all time points n and all states $i_0, \ldots, i_{n-1}, i, j$.

It is frequently convenient to label the state space of the Markov chain by the nonnegative integers $\{0, 1, 2, \ldots\}$, which we will do unless the contrary is explicitly stated, and it is customary to speak of X_n as being in state i if $X_n = i$.

The probability of X_{n+1} being in state j given that X_n is in state i is called the *one-step transition probability* and is denoted by $P_{ij}^{n,n+1}$. That is,

$$P_{ij}^{n,n+1} = \Pr\{X_{n+1} = j | X_n = i\}. \tag{1.2}$$

The notation emphasizes that in general the transition probabilities are functions not only of the initial and final states, but also of the time of transition as well. When the one-step transition probabilities are independent of the time variable n, we say that the Markov chain has *stationary transition probabilities*. Since the vast majority of Markov chains that we shall encounter have stationary transition probabilities, we limit our discussion to this case. Then $P_{ij}^{n,n+1} = P_{ij}$ is independent of n, and P_{ij} is the conditional probability that the state value undergoes a transition from i to j in one trial. It is customary to arrange these numbers P_{ij} in a *matrix*, in the infinite square array

$$\mathbf{P} = \left\| \begin{array}{ccccc} P_{00} & P_{01} & P_{02} & P_{03} & \cdots \\ P_{10} & P_{11} & P_{12} & P_{13} & \cdots \\ P_{20} & P_{21} & P_{22} & P_{23} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & P_{i3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array} \right\|$$

and refer to $\mathbf{P} = |P_{ij}|$ as the Markov matrix or *transition probability matrix* of the process.

The ith row of $\mathbf{P}$, for $i = 0, 1, \ldots$, is the probability distribution of the values of X_{n+1} under the condition that $X_n = i$. If the number of states is finite, then $\mathbf{P}$ is a finite square matrix whose order (the number of rows) is equal to the number of states. Clearly, the quantities P_{ij} satisfy the conditions

$$P_{ij} \geq 0 \quad \text{for } i, j = 0, 1, 2, \ldots, \tag{1.3}$$

$$\sum_{j=0}^{\infty} P_{ij} = 1 \quad \text{for } i = 0, 1, 2, \ldots. \tag{1.4}$$

The condition (1.4) merely expresses the fact that some transition occurs at each trial. (For convenience, one says that a transition has occurred even if the state remains unchanged.)

A Markov process is completely defined once its transition probability matrix and initial state X_0 (or, more generally, the probability distribution of X_0) are specified. We shall now prove this fact.

Let $\Pr\{X_0 = i\} = p_i$. It is enough to show how to compute the quantities

$$\Pr\{X_0 = i_0, X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n\}, \qquad (1.5)$$

since any probability involving $X_{j_1}, \ldots, X_{j_k}$, for $j_1 < \ldots < j_k$, can be obtained, according to the axiom of total probability, by summing terms of the form (1.5).

By the definition of conditional probabilities we obtain

$$\Pr\{X_0 = i_0, X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n\}$$
$$= \Pr\{X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\} \qquad (1.6)$$
$$\times \Pr\{X_n = i_n | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\}.$$

Now, by the definition of a Markov process,

$$\Pr\{X_n = i_n | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\}$$
$$= \Pr\{X_n = i_n | X_{n-1} = i_{n-1}\} = P_{i_{n-1}, i_n}. \qquad (1.7)$$

Substituting (1.7) into (1.6) gives

$$\Pr\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\}$$
$$= \Pr\{X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}\} P_{i_{n-1}, i_n}.$$

Then, upon repeating the argument $n - 1$ additional times, (1.5) becomes

$$\Pr\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\}$$
$$= p_{i_0} P_{i_0, i_1} \cdots P_{i_{n-2}, i_{n-1}} P_{i_{n-1}, i_n}. \qquad (1.8)$$

This shows that all finite-dimensional probabilities are specified once the transition probabilities and initial distribution are given, and in this sense the process is defined by these quantities.

Related computations show that (1.1) is equivalent to the Markov property in the form

$$\Pr\{X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_0 = i_0, \ldots, X_n = i_n\}$$
$$= \Pr\{X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i_n\} \qquad (1.9)$$

for all time points n, m and all states $i_0, \ldots, i_n, j_1, \ldots, j_m$. In other words, once (1.9) is established for the value $m = 1$, it holds for all $m \geq 1$ as well.

Exercises

1.1.   A Markov chain $X_0, X_1, \ldots$ on states 0, 1, 2 has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\|\begin{array}{ccc} 0.1 & 0.2 & 0.7 \\ 0.9 & 0.1 & 0 \\ 0.1 & 0.8 & 0.1 \end{array}\right\| \end{array}
$$

and initial distribution $p_0 = \Pr\{X_0 = 0\} = 0.3$, $p_1 = \Pr\{X_0 = 1\} = 0.4$, and $p_2 = \Pr\{X_0 = 2\} = 0.3$. Determine $\Pr\{X_0 = 0, X_1 = 1, X_2 = 2\}$.

1.2.   A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\|\begin{array}{ccc} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{array}\right\| \end{array}.
$$

Determine the conditional probabilities

$$
\Pr\{X_2 = 1, X_3 = 1 | X_1 = 0\} \quad \text{and} \quad \Pr\{X_1 = 1, X_2 = 1 | X_0 = 0\}.
$$

1.3.   A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\|\begin{array}{ccc} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.4 & 0.1 & 0.5 \end{array}\right\| \end{array}.
$$

If it is known that the process starts in state $X_0 = 1$, determine the probability $\Pr\{X_0 = 1, X_1 = 0, X_2 = 2\}$.

1.4. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}
\begin{array}{ccc}
0 & 1 & 2 \\
\left\| \begin{array}{ccc}
0.1 & 0.1 & 0.8 \\
0.2 & 0.2 & 0.6 \\
0.3 & 0.3 & 0.4
\end{array} \right\|
\end{array}.
$$

Determine the conditional probabilities

$$\Pr\{X_1 = 1, X_2 = 1 | X_0 = 0\} \quad \text{and} \quad \Pr\{X_2 = 1, X_3 = 1 | X_1 = 0\}.$$

1.5. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}
\begin{array}{ccc}
0 & 1 & 2 \\
\left\| \begin{array}{ccc}
0.3 & 0.2 & 0.5 \\
0.5 & 0.1 & 0.4 \\
0.5 & 0.2 & 0.3
\end{array} \right\|
\end{array}
$$

and initial distribution $p_0 = 0.5$ and $p_1 = 0.5$. Determine the probabilities

$$\Pr\{X_0 = 1, X_1 = 1, X_2 = 0\} \quad \text{and} \quad \Pr\{X_1 = 1, X_2 = 1, X_3 = 0\}.$$

Problems

1.1. A simplified model for the spread of a disease goes this way: The total population size is $N = 5$, of which some are diseased and the remainder are healthy. During any single period of time, two people are selected at random from the population and assumed to interact. The selection is such that an encounter between any pair of individuals in the population is just as likely as between any other pair. If one of these persons is diseased and the other not, then with probability $\alpha = 0.1$ the disease is transmitted to the healthy person. Otherwise, no disease transmission takes place. Let X_n denote the number of diseased persons in the population at the end of the nth period. Specify the transition probability matrix.

1.2. Consider the problem of sending a binary message, 0 or 1, through a signal channel consisting of several stages, where transmission through each stage is subject to a fixed probability of error α. Suppose that $X_0 = 0$ is the signal that is sent and let X_n be the signal that is received at the nth stage. Assume that $\{X_n\}$ is a Markov chain with transition probabilities $P_{00} = P_{11} = 1 - \alpha$ and $P_{01} = P_{10} = \alpha$, where $0 < \alpha < 1$.

(a) Determine $\Pr\{X_0 = 0, X_1 = 0, X_2 = 0\}$, the probability that no error occurs up to stage $n = 2$.

(b) Determine the probability that a correct signal is received at stage 2.

Hint: This is $\Pr\{X_0 = 0, X_1 = 0, X_2 = 0\} + \Pr\{X_0 = 0, X_1 = 1, X_2 = 0\}$.

1.3. Consider a sequence of items from a production process, with each item being graded as good or defective. Suppose that a good item is followed by another good item with probability α and is followed by a defective item with probability $1 - \alpha$. Similarly, a defective item is followed by another defective item with probability β and is followed by a good item with probability $1 - \beta$. If the first item is good, what is the probability that the first defective item to appear is the fifth item?

1.4. The random variables $\xi_1, \xi_2, \ldots$ are independent and with the common probability mass function

$k =$	0	1	2	3
$\Pr\{\xi = k\} =$	0.1	0.3	0.2	0.4

Set $X_0 = 0$, and let $X_n = \max\{\xi_1, \ldots, \xi_n\}$ be the largest ξ observed to date. Determine the transition probability matrix for the Markov chain $\{X_n\}$.

2. Transition Probability Matrices of a Markov Chain

A Markov chain is completely defined by its one-step transition probability matrix and the specification of a probability distribution on the state of the process at time 0. The analysis of a Markov chain concerns mainly the calculation of the probabilities of the possible realizations of the process.

Central in these calculations are the n-step transition probability matrices $\mathbf{P}^{(n)} = \|P_{ij}^{(n)}\|$. Here $P_{ij}^{(n)}$ denotes the probability that the process goes from state i to state j in n transitions. Formally,

$$P_{ij}^{(n)} = \Pr\{X_{m+n} = j | X_m = i\}. \tag{2.1}$$

Observe that we are dealing only with temporally homogeneous processes having stationary transition probabilities, since otherwise the left side of (2.1) would also depend on m.

The Markov property allows us to express (2.1) in terms of $\|P_{ij}\|$ as stated in the following theorem.

Theorem 2.1 *The n-step transition probabilities of a Markov chain satisfy*

$$P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)}, \tag{2.2}$$

where we define

$$P_{ij}^{(0)} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

From the theory of matrices we recognize the relation (2.2) as the formula for matrix multiplication, so that $\mathbf{P}^{(n)} = \mathbf{P} \times \mathbf{P}^{(n-1)}$. By iterating this formula, we obtain

$$\mathbf{P}^{(n)} = \underbrace{\mathbf{P} \times \mathbf{P} \times \cdots \times \mathbf{P}}_{n \text{ factors}} = \mathbf{P}^n; \tag{2.3}$$

in other words, the n-step transition probabilities $P_{ij}^{(n)}$ are the entries in the matrix $\mathbf{P}^n$, the nth power of $\mathbf{P}$.

Proof The proof proceeds via a *first step analysis*, a breaking down, or analysis, of the possible transitions on the first step, followed by an application of the Markov property. The event of going from state i to state j in n transitions can be realized in the mutually exclusive ways of going to some intermediate state k ($k = 0, 1, \ldots$) in the first transition, and then going from state k to state j in the remaining $(n - 1)$ transitions. Because of the Markov property, the probability of the second transition is $P_{kj}^{(n-1)}$ and that of the first is clearly P_{ik}. If we use the law of total probability, then (2.2) follows. The steps are

$$P_{ij}^{(n)} = \Pr\{X_n = j|X_0 = i\} = \sum_{k=0}^{\infty} \Pr\{X_n = j, X_1 = k|X_0 = i\}$$

$$= \sum_{k=0}^{\infty} \Pr\{X_1 = k|X_0 = i\} \Pr\{X_n = j|X_0 = i, X_1 = k\}$$

$$= \sum_{k=0}^{\infty} P_{ik}P_{kj}^{(n-1)}.$$

If the probability of the process initially being in state j is p_j, i.e., the distribution law of X_0 is $\Pr\{X_0 = j\} = p_j$, then the probability of the process being in state k at time n is

$$p_k^{(n)} = \sum_{j=0}^{\infty} p_j P_{jk}^{(n)} = \Pr\{X_n = k\}. \tag{2.4}$$

Exercises

2.1. A Markov chain $\{X_n\}$ on the states 0, 1, 2 has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}
\begin{array}{ccc}
0 & 1 & 2 \\
\left\| \begin{array}{ccc}
0.1 & 0.2 & 0.7 \\
0.2 & 0.2 & 0.6 \\
0.6 & 0.1 & 0.3
\end{array} \right\|
\end{array}.
$$

(a) Compute the two-step transition matrix P^2.
(b) What is $\Pr\{X_3 = 1|X_1 = 0\}$?
(c) What is $\Pr\{X_3 = 1|X_0 = 0\}$?

2.2. A particle moves among the states 0, 1, 2 according to a Markov process whose transition probability matrix is

$$
\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}
\begin{array}{ccc}
0 & 1 & 2 \\
\left\| \begin{array}{ccc}
0 & \frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & 0 & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & 0
\end{array} \right\|
\end{array}.
$$

Let X_n denote the position of the particle at the nth move. Calculate $\Pr\{X_n = 0|X_0 = 0\}$ for $n = 0, 1, 2, 3, 4$.

2.3. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\| \begin{array}{ccc} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{array} \right\| \end{array}.
$$

Determine the conditional probabilities

$$
\Pr\{X_3 = 1 | X_0 = 0\} \quad \text{and} \quad \Pr\{X_4 = 1 | X_0 = 0\}.
$$

2.4. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\| \begin{array}{ccc} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.4 & 0.1 & 0.5 \end{array} \right\| \end{array}.
$$

If it is known that the process starts in state $X_0 = 1$, determine the probability $\Pr\{X_2 = 2\}$.

2.5. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\| \begin{array}{ccc} 0.1 & 0.1 & 0.8 \\ 0.2 & 0.2 & 0.6 \\ 0.3 & 0.3 & 0.4 \end{array} \right\| \end{array}.
$$

Determine the conditional probabilities

$$
\Pr\{X_3 = 1 | X_1 = 0\} \quad \text{and} \quad \Pr\{X_2 = 1 | X_0 = 0\}.
$$

2.6. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$
\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{array}{ccc} 0 & 1 & 2 \\ \left\| \begin{array}{ccc} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.1 & 0.4 \\ 0.5 & 0.2 & 0.3 \end{array} \right\| \end{array}.
$$

and initial distribution $p_0 = 0.5$ and $p_1 = 0.5$. Determine the probabilities $\Pr\{X_2 = 0\}$ and $\Pr\{X_3 = 0\}$.

Problems

2.1. Consider the Markov chain whose transition probability matrix is given by

$$
\mathbf{P} = \begin{array}{c}
 \\
0 \\
1 \\
2 \\
3
\end{array}
\begin{array}{cccc}
0 & 1 & 2 & 3 \\
\left\|\begin{array}{cccc}
0.4 & 0.3 & 0.2 & 0.1 \\
0.1 & 0.4 & 0.3 & 0.2 \\
0.3 & 0.2 & 0.1 & 0.4 \\
0.2 & 0.1 & 0.4 & 0.3
\end{array}\right\|
\end{array}.
$$

Suppose that the initial distribution is $p_i = \frac{1}{4}$ for $i = 0, 1, 2, 3$. Show that $\Pr\{X_n = k\} = \frac{1}{4}$, $k = 0, 1, 2, 3$, for all n. Can you deduce a general result from this example?

2.2. Consider the problem of sending a binary message, 0 or 1, through a signal channel consisting of several stages, where transmission through each stage is subject to a fixed probability of error α. Let X_0 be the signal that is sent and let X_n be the signal that is received at the nth stage. Suppose X_n is a Markov chain with transition probabilities $P_{00} = P_{11} = 1 - \alpha$ and $P_{01} = P_{10} = \alpha$, $(0 < \alpha < 1)$. Determine $\Pr\{X_5 = 0 | X_0 = 0\}$, the probability of correct transmission through five stages.

2.3. Let X_n denote the quality of the nth item produced by a production system with $X_n = 0$ meaning "good" and $X_n = 1$ meaning "defective." Suppose that X_n evolves as a Markov chain whose transition probability matrix is

$$
\mathbf{P} = \begin{array}{c}
 \\
0 \\
1
\end{array}
\begin{array}{cc}
0 & 1 \\
\left\|\begin{array}{cc}
0.99 & 0.01 \\
0.12 & 0.88
\end{array}\right\|
\end{array}.
$$

What is the probability that the fourth item is defective given that the first item is defective?

5.2. Let X be a Bernoulli random variable with parameter p. Compare $\Pr\{X \geq 1\}$ with the Markov inequality bound.

5.3. Let ξ be a random variable with mean μ and standard deviation σ. Let $X = (\xi - \mu)^2$. Apply Markov's inequality to X to deduce Chebyshev's inequality:

$$\Pr\{|\xi - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \qquad \text{for any } \varepsilon > 0.$$

Problems

5.1. Use the law of total probability for conditional expectations $E[E\{X|Y, Z\}|Z] = E[X|Z]$ to show

$$E[X_{n+2}|X_0, \ldots, X_n] = E[E\{X_{n+2}|X_0, \ldots, X_{n+1}\}|X_0, \ldots, X_n].$$

Conclude that when X_n is a martingale.

$$E[X_{n+2}|X_0, \ldots, X_n] = X_n.$$

5.2. Let $U_1, U_2, \ldots$ be independent random variables each uniformly distributed over the interval $(0, 1]$. Show that $X_0 = 1$ and $X_n = 2^n U_1 \cdots U_n$ for $n = 1, 2, \ldots$ defines a martingale.

5.4. Let $\xi_1, \xi_2, \ldots$ be independent Bernoulli random variables with parameter p, $0 < p < 1$. Show that $X_0 = 1$ and $X_n = p^{-n} \xi_1 \cdots \xi_n$, $n = 1$, $2, \ldots$, defines a nonnegative martingale. What is the limit of X_n as $n \to \infty$?

Problems

1.1. A simplified model for the spread of a disease goes this way: The total population size is $N = 5$, of which some are diseased and the remainder are healthy. During any single period of time, two people are selected at random from the population and assumed to interact. The selection is such that an encounter between any pair of individuals in the population is just as likely as between any other pair. If one of these persons is diseased and the other not, then with probability $\alpha = 0.1$ the disease is transmitted to the healthy person. Otherwise, no disease transmission takes place. Let X_n denote the number of diseased persons in the population at the end of the nth period. Specify the transition probability matrix.