11.2.1 Introduction

Consider a discrete-time random process $\{X_m, m = 0, 1, 2, \ldots\}$. In the very simple case where the X_m's are independent, the analysis of this process is relatively straightforward. In this case, there is no "memory" in the system, so each X_m can be looked at independently from previous ones.

However, for many real-life processes, the independence assumption is not valid. For example, if X_m is the stock price of a company at time $m \in \{0, 1, 2, ..\}$, then it is reasonable to assume that the X_m's are dependent. Therefore, we need to develop models where the value of X_m depends on the previous values. In a Markov chain, X_{m+1} depends on X_m, but given X_m, it does not depend on the other previous values X_0, X_1, $\cdots$, X_{m-1}. That is, conditioned on X_m, the random variable X_{m+1} is independent of the random variables X_0, X_1, $\cdots$, X_{m-1}.

Markov chains are usually used to model the evolution of "states" in probabilistic systems. More specifically, consider a system with the set of possible states $S = \{s_1, s_2, \ldots\}$. Without loss of generality, the states are usually chosen to be 0, 1, 2, $\cdots$, or 1, 2, 3, $\cdots$, depending on which one is more convenient for a particular problem. If $X_n = i$, we say that the system is in state i at time n. The idea behind Markov chains is usually summarized as follows: "conditioned on the current state, the past and the future states are independent."

For example, suppose that we are modeling a queue at a bank. The number of people in the queue is a non-negative integer. Here, the state of the system can be defined as the number of people in the queue. More specifically, if X_n shows the number of people in the queue at time n, then $X_n \in S = \{0, 1, 2, \ldots\}$. The set S is called the *state space* of the Markov chain. Let us now provide a formal definition for discrete-time Markov chains.

Discrete-Time Markov Chains

Consider the random process $\{X_n, n = 0, 1, 2, \cdots\}$, where $R_{X_i} = S \subset \{0, 1, 2, \cdots\}$. We say that this process is a **Markov chain** if

$$\begin{aligned} P(X_{m+1} = j | X_m = i, X_{m-1} = i_{m-1}, \cdots, X_0 = i_0) \\ = P(X_{m+1} = j | X_m = i), \end{aligned}$$

for all $m, j, i, i_0, i_1, \cdots i_{m-1}$.

If the number of states is finite, e.g., $S = \{0, 1, 2, \cdots, r\}$, we call it a **finite** Markov chain.

If $X_n = j$, we say that the process is in state j. The numbers $P(X_{m+1} = j | X_m = i)$ are called the **transition probabilities**. We assume that the transition probabilities do not depend on time. That is, $P(X_{m+1} = j | X_m = i)$ does not depend on m. Thus, we can define

$$p_{ij} = P(X_{m+1} = j | X_m = i).$$

In particular, we have

$$\begin{aligned} p_{ij} &= P(X_1 = j | X_0 = i) \\ &= P(X_2 = j | X_1 = i) \\ &= P(X_3 = j | X_2 = i) = \cdots. \end{aligned}$$

In other words, if the process is in state i, it will next make a transition to state j with probability p_{ij}.