Chapter 12

Markov processes

It is no exaggeration to say that Markov processes are the most (mathematically) important category of stochastic processes. They are also the easiest to understand for the following reason. A Markov process is one for which the *future* state of a system depends only on its *present* state and not on its *past*.

Consider a discrete variable X that changes randomly in time (at discrete intervals). The sequence of values of X at times $t = 0, 1, 2, 3 \ldots$ are denoted by $X_0, X_1, X_2, X_3 \ldots$ Such a sequence is often called a *time series* because the ordering is defined by the time at which the event occurs. An example of such a sequence could be a series of coin tosses where $X = 1$ is HEADS and $X = 0$ is used to denote the outcome TAILS. Then a typical sequence might look like:

$$00101110101110001 \cdots \qquad .$$

In this example, each outcome in the sequence is independent of all other outcomes. Each toss of the coin is not affected by the *past* outcomes or correlated in any way with *future* outcomes.

A formal definition of a *Markov chain* is that, given a sequence of random variables:

$$X_0, X_1, \cdots, X_{n-1}, X_n \tag{12.1}$$

then

$$\boxed{P(X_n = x_n | X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1}) \qquad .} \tag{12.2}$$

So if X_{n-1} represents the *present*, and X_n is the (uncertain) *future* , then we can consider the given sequence as the *history* or *past* of the chain. We say that a Markov chain has no *memory* of its past, and therefor no dependence on what happened in the past.

The simple random walk is an example of such a Markov chain. Suppose the walk starts at the point $x = a$ at $t = 0$, and then takes n steps. Let us denote the sequence of steps by:

$$X_1, X_2, \cdots, X_n \qquad . \tag{12.3}$$

where each step can have the value, $X_i \in \{-1, 0, +1\}$, corresponding to a step to the left, no step, or a step to the right. Then, given the walker starts at $S_0 = a$, the position of the walker after i steps will be

$$S_i = a + X_1 + X_2 + \cdots + X_i \qquad . \tag{12.4}$$

Suppose the walker has taken n steps, then the position on the next step (which may be random) only depends on where the walker is now. The future position has no dependence on how (the sequence of left/right steps) the walker arrived at the present position. In mathematical terms:

$$P(S_{n+1} = s_{n+1} | S_0 = a, S_1 = s_1, \ldots, S_n = s_n) = P(S_{n+1} = s_{n+1} | S_n = s_n) \qquad . \tag{12.5}$$

12.1 Transition matrix

A Markov process is said to be *homogeneous* (or *stationary*) if the probability of a transition does not depend on the time at which it occurs. In mathematical terms:

$$\boxed{P(X_m = x_k|X_{m-1} = x_j) = P(X_{m+n} = x_k|X_{m+n-1} = x_j)} \quad \text{for all} \quad n \geq 0 \quad . \tag{12.6}$$

In this formula n represents the time-gap between an earlier and later time in the sequence. For the present, we consider only *homogeneous* Markov processes.

In a Markov process we call the value of the random variable (X) the *state* of the system. For a discrete system, with a finite number of states, the states could then be labelled: $x_1, x_2, \ldots, x_n$.

So any step (*transition*) between two states is random, but has an associated probability. The conditional probability relating the one-step (present-to-future) process is called the *transition matrix* for the process. So at any time t_{m-1}, the probability of the system hopping from a given state $X_m = x_j$ to another state x_k is:

$$\boxed{p_{jk} = P(X_m = x_k|X_{m-1} = x_j)} \tag{12.7}$$

where the label j (row index) denotes the present *state* of the system. and the label k (column index) denotes the future (uncertain) state of the system. For N possible states, this will be an $N \times N$ matrix.

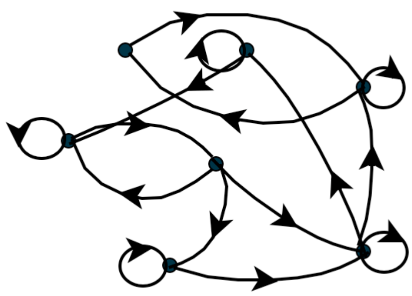

Figure 12.1: A transition graph representation of a Markov chain. The dark circles are the *nodes* of the graph (states of the system): in this case $N = 7$. The arrowed lines, called *edges* indicate the *transitions*, that is, the possible one-step jumps. Any node i which connects towards node j (directed arrow) means that a possible one-step jump $i \rightarrow j$ can occur. The graph can be represented by an equivalent *transition matrix* P, in which the arrows are equivalent to non-zero matrix elements in row i and column j. So each *edge* in the diagram will have an associated transition probability.

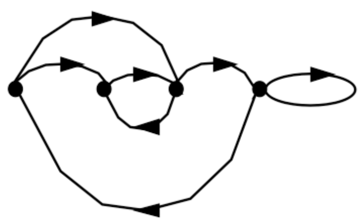

Figure 12.2: An example of a four-state Markov chain.

Now, at this point, we state that the N-states form a *partition of the event space*. That is, it is impossible for the system to be in two states simultaneously. The state of the system at any time is random, but it is unique.

This means that although the future state is uncertain, one of the states will certainly be occupied. That

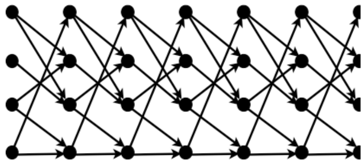

Figure 12.3: The probability tree of the Markov chain shown in figure (12.2). The vertical axis is the state label/value. The horizontal axis is time, with the arrows indicating the possible transitions that arise from the process (12.2). This diagram illustrates the homogeneity of the Markov process, time invariance. As we move along the diagram to the right or left, increase or decrease time, the picture remains the same. This is what is meant by time homogeneity, expressed by the relation (12.6).

is, the probabilities must add to 1. In terms of the transition matrix this translates to:

$$\sum_{k=1}^{N} p_{jk} = 1 \qquad \text{for all} \quad j \qquad . \tag{12.8}$$

That is, summing along every row of the *transition matrix* gives a total of 1. In the theory of linear algebra, a matrix with this property is also called a *stochastic matrix*.

EXAMPLE

Consider a coin which can be either HEADS or TAILS. There are two possible states which we will denote by $X = 1$ (HEADS) and $X = 0$ (TAILS).

For the coin, the probability of HEADS is $0 \leq p \leq 1$ on each and every toss, and the probability of TAILS is, $q = 1 - p$. Then each toss of the coin allows the state to change according to:

$$P(X_n = 0|X_{n-1} = 0) = p_{00} = q \qquad , \qquad P(X_n = 0|X_{n-1} = 1) = p_{10} = q \qquad . \tag{12.9}$$

if the toss is TAILS.

Similarly, we have:

$$P(X_n = 1|X_{n-1} = 1) = p_{11} = p \qquad , \qquad P(X_n = 1|X_{n-1} = 0) = p_{01} = p \qquad . \tag{12.10}$$

if HEADS turns up.

These transition probabilities (the changes between two successive tosses) have no dependence upon *when* they occur in the sequence of the chain. That is they have no dependence on n. So this is a homogeneous Markov chain for which the *transition matrix* can be written as:

$$\mathsf{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} q & p \\ q & p \end{pmatrix} \tag{12.11}$$

where the first row corresponds to starting in state 0, and the second row starting in state 1.

EXAMPLE

Consider a Markov chain, with states $X = 0$ and $X = 1$ defined by the transition matrix:

$$\mathsf{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \tag{12.12}$$

with, $0 \leq a, b \leq 1$. The transition matrix can be represented by a *transition graph* or *transition diagram*. The states are represented by the *nodes* (or states) and the *transitions* are indicated by *directed arrowed lines* called *edges*. An example of a transition diagram for a 7-state system is shown in figure 12.1. For the transition matrix (12.12) the corresponding graph is shown in figure (12.4). Note that the Markov chain is fully described by either its transition matrix or its transition graph.

EXAMPLE

QUESTION

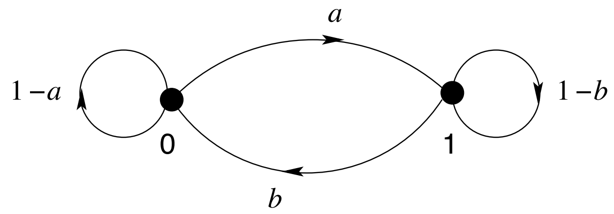

Figure 12.4: The transition graph corresponding to matrix (12.12). The dark circles are the nodes of the graph (states of the system). The directed and labelled lines show the transitions and the corresponding probabilities, respectively.

A simple symmetric random walk takes place between absorbing barriers at $x = 0$ and $x = 3$. Write down the transition matrix for this process, and draw the corresponding transition diagram.

SOLUTION

The position of the walker is the state of the system. Thus there are states corresponding to $X = 0, 1, 2, 3$, and thus the transition matrix will be 4×4.

For the absorbing barriers at $x = 0$ and $x = 3$, we know that, once the state reaches these points it never leaves. Thus the transition matrix for the first row will be:

$$p_{00} = 1 \quad , \quad p_{01} = p_{02} = p_{03} = 0 \quad .$$

The system is certain to be at $x = 0$ on the next step.

Similarly, if the system starts at $x = 3$ it is certain to remain there. Thus, for the last row of the matrix we have:

$$p_{33} = 1 \quad , \quad p_{30} = p_{31} = p_{32} = 0 \quad .$$

Starting at $x = 1$ (that is the second row of the matrix) we can either go left to $x = 0$ with probability $\frac{1}{2}$, or right to $x = 2$ with probability $\frac{1}{2}$. it is impossible to get to $x = 3$ in one step, however, Thus $p_{13} = 0$. Also the walker must step away, thus the probability of remaining at $x = 1$ on the next step is also zero.

Continuing in this manner we arrive at the answer:

$$\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad . \tag{12.13}$$

Note that each row sums to 1, as it should.

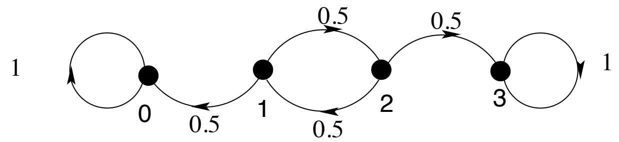

Figure 12.5: The transition graph corresponding to the simple random walk. The corresponding transition matrix is given by (12.13). The dark circles are the nodes of the graph (states of the system). The directed and labelled lines show the transition directions and the corresponding probabilities, respectively.

12.2 The n-step transition probability

Consider the probability of a two-step transition from state i to state j. That is: $X_0 = i \rightarrow X_2 = j$ (see figure 12.6). First we note that the states of the system form a partition of the event space. This can be calculated using the partition rule in the form:

$$P(X_2 = j|X_0 = i) = \sum_k P(X_2 = j|X_0 = i, X_1 = k)P(X_1 = k|X_0 = i) \quad . \tag{12.14}$$

That is, we have conditioned on the first step $i \to k$. Each state of the system is disjoint, and the sets of states in complete (covers all possibilities).

Given that the process is *Markovian* then we can simplify some of the terms. Specifically:

$$P(X_2 = j|X_0 = i, X_1 = k) = P(X_2 = j|X_1 = k) \qquad . \tag{12.15}$$

That is, the past (X_0) has no bearing on the transition probability for the next step. Furthermore, since the process is homogeneous, we have:

$$P(X_2 = j|X_1 = k) = P(X_1 = j|X_0 = k) = p_{kj} \qquad . \tag{12.16}$$

Similarly:

$$P(X_1 = k|X_0 = i) = p_{ik} \qquad , \tag{12.17}$$

so (12.14) simplifies to:

$$P(X_2 = j|X_0 = i) = \sum_k p_{kj}p_{ik} = \sum_k p_{ik}p_{kj} \qquad . \tag{12.18}$$

Then, we can easily show that the two-step matrix is also *stochastic*, as it should be. That is, starting in state i, after two-steps one should end up in one of the states j. The total of the transition probabilities should always sum to 1. So summing over j, the final states, for the two-step matrix gives us:

$$\sum_j P(X_2 = j|X_0 = i) = \sum_{j,k} p_{kj}p_{ik} = \sum_k p_{ik} = 1 \qquad . \tag{12.19}$$

Recall that, given a matrix A and a matrix B, then the product AB is defined as:

$$(\mathsf{AB})_{ij} \equiv \sum_{k=1}^{n} \mathsf{A}_{ik}\mathsf{B}_{kj}$$

which exists if the number of columns of A equals the number of rows of B. Examination of (12.18), shows that it can be written as such a matrix product:

$$P(X_2 = j|X_0 = i) = (\mathsf{P}.\mathsf{P})_{ij} = \left(\mathsf{P}^2\right)_{ij} \qquad , \tag{12.20}$$

where the superscript indicates the square of the matrix P. That is, a two-step transition matrix is the product of two single-step matrices.

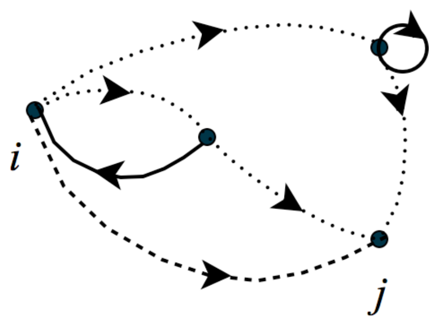

Figure 12.6: The one-step transition $i \to j$ is indicated by the dashed line. The two-step transitions $i \to k \to j$ are shown by dotted lines.

Then, in general, any n-step process is the result of n one-step processes:

$$\boxed{P(X_n = j|X_0 = i) = \left(\mathsf{P}^n\right)_{ij}} \qquad . \tag{12.21}$$

This is the most important result in this chapter, and is now proved.

This can be proved by induction as follows. Clearly, for $n = 1$:

$$P(X_1 = j|X_0 = i) = (\mathsf{P}^n)_{ij} \quad , \tag{12.22}$$

by definition. Now we show that, assuming (12.21) is valid for n this *implies* that the relation is also valid for $n+1$, and thus $n = 2$ and so on.

Consider, $P(X_{n+1} = j|X_0 = i)$. By conditioning on the first step we have that:

$$P(X_{n+1} = j|X_0 = i) = \sum_k P(X_{n+1} = j|X_0 = i, X_1 = k)P(X_1 = k|X_0 = i) \quad , \tag{12.23}$$

Now using the Markovianity property (12.2):

$$P(X_{n+1} = j|X_0 = i, X_1 = k) = P(X_{n+1} = j|X_1 = k) \quad , \tag{12.24}$$

next using the homogeneity property (12.6), we have that

$$P(X_{n+1} = j|X_1 = k) = P(X_n = j|X_0 = k) \quad , \tag{12.25}$$

Now according to our assertion/assumption (12.21) we can write this as:

$$P(X_n = j|X_0 = k) = (\mathsf{P}^n)_{kj} \quad , \tag{12.26}$$

Then, it follows that (12.23) can be written as:

$$P(X_{n+1} = j|X_0 = i) = \sum_k (\mathsf{P}^n)_{kj} (\mathsf{P})_{ik} \quad , \tag{12.27}$$

Then clearly, according to the rules of matrix multiplication:

$$P(X_{n+1} = j|X_0 = i) = \left(\mathsf{P}^{n+1}\right)_{ij} \quad , \tag{12.28}$$

which proves the assertion by induction.

12.3 Chapman-Kolmogorov relation

Since, for matrix products, we have the following rule of indices:

$$\mathsf{P}^{m+n} = \mathsf{P}^m\mathsf{P}^n = \mathsf{P}^n\mathsf{P}^m \tag{12.29}$$

this gives the *Chapman-Kolmogorov relation*

$$\boxed{\left(\mathsf{P}^{m+n}\right)_{ij} = \sum_k (\mathsf{P}^m)_{ik} (\mathsf{P}^n)_{kj}} \quad . \tag{12.30}$$

or in equivalent form:

$$\boxed{P(X_{m+n} = j|X_0 = i) = \sum_k P(X_m = k|X_0 = i)P(X_n = j|X_0 = k)} \quad . \tag{12.31}$$

That is a transition probability over $m+n$ steps can be divided into two stages: m steps (to an intermediate point,k) followed by n steps to the final point j. Then we sum over all possible intermediate points (partitions) that are possible.

12.4 Examples

- Consider a Markov chain defined between the states $X = 1, 2, 3, 4$ defined by the transition matrix:

$$\mathsf{P} = \begin{pmatrix} 0.9 & 0.0 & 0.1 & 0.0 \\ 0.5 & 0.2 & 0.0 & 0.3 \\ 0.2 & 0.2 & 0.2 & 0.4 \\ 0.4 & 0.4 & 0.2 & 0.0 \end{pmatrix} \quad . \tag{12.32}$$

Then, suppose we wish to calculate the two-step transition probabilities, for example; $P(X_2 = 3|X_0 = 4)$. We can do this by calculating the matrix P^2 and looking at the element in row 4 (initial state) and column 3 (final state).

$$\mathsf{P}^2 = \begin{pmatrix} 0.8300 & 0.0200 & 0.1100 & 0.0400 \\ 0.6700 & 0.1600 & 0.1100 & 0.0600 \\ 0.4800 & 0.2400 & 0.1400 & 0.1400 \\ 0.6000 & 0.1200 & 0.0800 & 0.2000 \end{pmatrix} \quad . \tag{12.33}$$

So, from this matrix we can read off the transition probabilities, for example:

$$P(X_2 = 3|X_0 = 4) = 0.0800 \quad , \quad P(X_2 = 2|X_0 = 2) = 0.1600 \quad . \tag{12.34}$$

Notice that P^2 is also *stochastic* as it should be, and as proven above (12.19). That is summing along the rows gives 1 in each case.

Suppose we are interested in the 8-step processes [1], then we calculate P^8, and so on:

$$\mathsf{P}^8 = \begin{pmatrix} 0.7676 & 0.0588 & 0.1115 & 0.0622 \\ 0.7668 & 0.0593 & 0.1115 & 0.0624 \\ 0.7655 & 0.0602 & 0.1113 & 0.0630 \\ 0.7663 & 0.0596 & 0.1114 & 0.0628 \end{pmatrix} \quad . \tag{12.35}$$

An interesting feature of this last result is that the rows of this matrix seem to be almost identical. This is a special feature of this system, indicating that a unique *equilibrium state* exists We discuss this aspect in much more detail in the next few chapters.

- Let's return to the simple random walk represented by figure (12.5). In this case the states are $X = 0, 1, 2, 3$ and the transition matrix is (12.13):

$$\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad . \tag{12.36}$$

Suppose the present state of the system is $X = 2$, and we are interested in what might happen after two steps. In the context of the gambler's ruin analogy, we start with $k = 2$ and play two games of coin tossing.

According to the Chapman-Kolmogorov formula we can use the matrix P^2 to answer this question.

$$\mathsf{P}^2 = \begin{pmatrix} 1.00 & 0 & 0 & 0 \\ 0.50 & 0.25 & 0 & 0.25 \\ 0.25 & 0 & 0.25 & 0.50 \\ 0 & 0 & 0 & 1.00 \end{pmatrix} \quad . \tag{12.37}$$

Reading along row 1- we see that, if we start in $X_0 = 0$, we are certain to stay there. This makes sense since there are no transition leaving $X = 0$ (see figure **??**). Similarly, if one starts in $X_0 = 3$, one never leaves that state.

[1] Computations carried out with GNU Octave. GNU Octave is a high-level language, primarily intended for numerical computations. It is free software under the terms of the GNU General Public License (GPL) and available under Linux/Windows/MacOSX.

Reading along row 3 (corresponding to starting in $X = 2$) we see that, from the first column, $P(X_2 = 0|X_0 = 2) = 0.25$, and so forth.

Let's consider the long-run, namely, for any starting point, $X_0 = i$, where would one end up:

$$P(X_\infty = j|X_0 = i) = \mathsf{P}^\infty)_{\mathsf{ij}} \quad .(12.38)$$

We find that after 6 steps

$$\mathsf{P}^6 = \begin{pmatrix} 1.0000 & 0 & 0 & 0 \\ 0.6562 & 0.0156 & 0 & 0.3281 \\ 0.3281 & 0 & 0.0156 & 0.6562 \\ 0 & 0 & 0 & 1.0000 \end{pmatrix} \quad . \tag{12.39}$$

As before (rows 1 and 4) starting in one of the ends, one is trapped there. The middle two rows (rows 2 and 3) show how the random process evolves of walkers starting at $X_0 = 1$ and $X_0 = 2$.

Now consider the pattern that emerges in the long-run, say after 100 steps.

$$\mathsf{P}^{100} = \begin{pmatrix} 1.0000 & 0 & 0 & 0 \\ 0.6667 & 0 & 0 & 0.3333 \\ 0.3333 & 0 & 0 & 0.6667 \\ 0 & 0 & 0 & 1.0000 \end{pmatrix} \quad . \tag{12.40}$$

Row 2 tells us the eventual fate of a simple random walk that starts from $X_0 = 1$, and so on.

Since this problem has already been solved in earlier chapters, we can check that the result is consistent with the formula previously obtained. The problem corresponds to gambler's ruin with $p = q = 0.5$ In that case, we found, starting at $X_0 = k$ the probability of eventually ending up at $X_\infty = 0$, given $N = 3$ was

$$P(X_\infty = 0|X_0 = k) = p_k = 1 - \frac{k}{N} = 1 - \frac{k}{3} \quad , \tag{12.41}$$

The numbers given by this formula agree with those for our transition matrix (12.40) reading down the first column corresponding to, $k = 0, 1, 2$ and 3, respectively.

We also note that, in the long run, we end up in either $X = 0$ or $X = 3$. That is, eventually, the walker/gambler is absorbed at one end or the other. This is a consequence of the law of large numbers, but we arrive at the same conclusion based on numerical results by noting the two columns of zeros running down the centre of the matrix (12.40), which tells us there is zero probability of ending up somewhere in the middle.