

2.9 An Application to Markov Chains

Many natural phenomena progress through various stages and can be in a variety of states at each stage. For example, the weather in a given city progresses day by day and, on any given day, may be sunny or rainy. Here the states are “sun” and “rain,” and the weather progresses from one state to another in daily stages. Another example might be a football team: The stages of its evolution are the games it plays, and the possible states are “win,” “draw,” and “loss.”

The general setup is as follows: A real conceptual “system” is run generating a sequence of outcomes. The system evolves through a series of “stages,” and at any stage it can be in any one of a finite number of “states.” At any given stage, the state to which it will go at the next stage depends on the past and present history of the system—that is, on the sequence of states it has occupied to date.

Definition 2.15 Markov Chain

A **Markov chain** is such an evolving system wherein the state to which it will go next depends only on its present state and does not depend on the earlier history of the system.¹⁹

Even in the case of a Markov chain, the state the system will occupy at any stage is determined only in terms of probabilities. In other words, chance plays a role. For example, if a football team wins a particular game, we do not know whether it will win, draw, or lose the next game. On the other hand, we may know that the team tends to persist in winning streaks; for example, if it wins one game it may win the next game $\frac{1}{2}$ of the time, lose $\frac{4}{10}$ of the time, and draw $\frac{1}{10}$ of the time. These fractions are called the **probabilities** of these various possibilities. Similarly, if the team loses, it may lose the next game with probability $\frac{1}{2}$ (that is, half the time), win with probability $\frac{1}{4}$, and draw with probability $\frac{1}{4}$. The probabilities of the various outcomes after a drawn game will also be known.

We shall treat probabilities informally here: *The probability that a given event will occur is the long-run proportion of the time that the event does indeed occur.* Hence, all probabilities are numbers between 0 and 1. A probability of 0 means the event is impossible and never occurs; events with probability 1 are certain to occur.

If a Markov chain is in a particular state, the probabilities that it goes to the various states at the next stage of its evolution are called the **transition probabilities** for the chain, and they are assumed to be known quantities. To motivate the general conditions that follow, consider the following simple example. Here the system is a man, the stages are his successive lunches, and the states are the two restaurants he chooses.

Example 2.9.1

A man always eats lunch at one of two restaurants, *A* and *B*. He never eats at *A* twice in a row. However, if he eats at *B*, he is three times as likely to eat at *B* next time as at *A*. Initially, he is equally likely to eat at either restaurant.

- a. What is the probability that he eats at *A* on the third day after the initial one?

¹⁹The name honours Andrei Andreyevich Markov (1856–1922) who was a professor at the university in St. Petersburg, Russia.

b. What proportion of his lunches does he eat at A ?

Solution. The table of transition probabilities follows. The A column indicates that if he eats at A on one day, he never eats there again on the next day and so is certain to go to B .

		Present Lunch	
		A	B
Next Lunch	A	0	0.25
	B	1	0.75

The B column shows that, if he eats at B on one day, he will eat there on the next day $\frac{3}{4}$ of the time and switches to A only $\frac{1}{4}$ of the time.

The restaurant he visits on a given day is not determined. The most that we can expect is to know the probability that he will visit A or B on that day.

Let $\mathbf{s}_m = \begin{bmatrix} s_1^{(m)} \\ s_2^{(m)} \end{bmatrix}$ denote the *state vector* for day m . Here $s_1^{(m)}$ denotes the probability that he eats at A on day m , and $s_2^{(m)}$ is the probability that he eats at B on day m . It is convenient to let \mathbf{s}_0 correspond to the initial day. Because he is equally likely to eat at A or B on that initial day, $s_1^{(0)} = 0.5$ and $s_2^{(0)} = 0.5$, so $\mathbf{s}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$. Now let

$$P = \begin{bmatrix} 0 & 0.25 \\ 1 & 0.75 \end{bmatrix}$$

denote the *transition matrix*. We claim that the relationship

$$\mathbf{s}_{m+1} = P\mathbf{s}_m$$

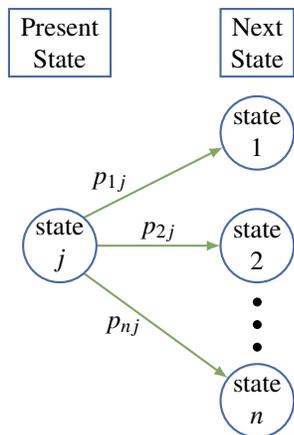
holds for all integers $m \geq 0$. This will be derived later; for now, we use it as follows to successively compute $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots$

$$\begin{aligned} \mathbf{s}_1 &= P\mathbf{s}_0 = \begin{bmatrix} 0 & 0.25 \\ 1 & 0.75 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.125 \\ 0.875 \end{bmatrix} \\ \mathbf{s}_2 &= P\mathbf{s}_1 = \begin{bmatrix} 0 & 0.25 \\ 1 & 0.75 \end{bmatrix} \begin{bmatrix} 0.125 \\ 0.875 \end{bmatrix} = \begin{bmatrix} 0.21875 \\ 0.78125 \end{bmatrix} \\ \mathbf{s}_3 &= P\mathbf{s}_2 = \begin{bmatrix} 0 & 0.25 \\ 1 & 0.75 \end{bmatrix} \begin{bmatrix} 0.21875 \\ 0.78125 \end{bmatrix} = \begin{bmatrix} 0.1953125 \\ 0.8046875 \end{bmatrix} \end{aligned}$$

Hence, the probability that his third lunch (after the initial one) is at A is approximately 0.195, whereas the probability that it is at B is 0.805. If we carry these calculations on, the next state vectors are (to five figures):

$$\begin{aligned} \mathbf{s}_4 &= \begin{bmatrix} 0.20117 \\ 0.79883 \end{bmatrix} & \mathbf{s}_5 &= \begin{bmatrix} 0.19971 \\ 0.80029 \end{bmatrix} \\ \mathbf{s}_6 &= \begin{bmatrix} 0.20007 \\ 0.79993 \end{bmatrix} & \mathbf{s}_7 &= \begin{bmatrix} 0.19998 \\ 0.80002 \end{bmatrix} \end{aligned}$$

Moreover, as m increases the entries of \mathbf{s}_m get closer and closer to the corresponding entries of $\begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$. Hence, in the long run, he eats 20% of his lunches at A and 80% at B .



Example 2.9.1 incorporates most of the essential features of all Markov chains. The general model is as follows: The system evolves through various stages and at each stage can be in exactly one of n distinct states. It progresses through a sequence of states as time goes on. If a Markov chain is in state j at a particular stage of its development, the probability p_{ij} that it goes to state i at the next stage is called the **transition probability**. The $n \times n$ matrix $P = [p_{ij}]$ is called the **transition matrix** for the Markov chain. The situation is depicted graphically in the diagram.

We make one important assumption about the transition matrix $P = [p_{ij}]$: It does *not* depend on which stage the process is in. This assumption means that the transition probabilities are *independent of time*—that is, they do not change as time goes on. It is this assumption that distinguishes Markov chains in the literature of this subject.

Example 2.9.2

Suppose the transition matrix of a three-state Markov chain is

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{array}{ccc} \text{Present state} & & \\ 1 & 2 & 3 \\ \begin{bmatrix} 0.3 & 0.1 & 0.6 \\ 0.5 & 0.9 & 0.2 \\ 0.2 & 0.0 & 0.2 \end{bmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \text{Next state} \end{array}$$

If, for example, the system is in state 2, then column 2 lists the probabilities of where it goes next. Thus, the probability is $p_{12} = 0.1$ that it goes from state 2 to state 1, and the probability is $p_{22} = 0.9$ that it goes from state 2 to state 2. The fact that $p_{32} = 0$ means that it is impossible for it to go from state 2 to state 3 at the next stage.

Consider the j th column of the transition matrix P .

$$\begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix}$$

If the system is in state j at some stage of its evolution, the transition probabilities $p_{1j}, p_{2j}, \dots, p_{nj}$ represent the fraction of the time that the system will move to state 1, state 2, \dots , state n , respectively, at the next stage. We assume that it has to go to *some* state at each transition, so the sum of these probabilities is 1:

$$p_{1j} + p_{2j} + \cdots + p_{nj} = 1 \quad \text{for each } j$$

Thus, the columns of P all sum to 1 and the entries of P lie between 0 and 1. Hence P is called a **stochastic matrix**.

As in Example 2.9.1, we introduce the following notation: Let $s_i^{(m)}$ denote the probability that the

system is in state i after m transitions. The $n \times 1$ matrices

$$\mathbf{s}_m = \begin{bmatrix} s_1^{(m)} \\ s_2^{(m)} \\ \vdots \\ s_n^{(m)} \end{bmatrix} \quad m = 0, 1, 2, \dots$$

are called the **state vectors** for the Markov chain. Note that the sum of the entries of \mathbf{s}_m must equal 1 because the system must be in *some* state after m transitions. The matrix \mathbf{s}_0 is called the **initial state vector** for the Markov chain and is given as part of the data of the particular chain. For example, if the chain has only two states, then an initial vector $\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ means that it started in state 1. If it started in state 2, the initial vector would be $\mathbf{s}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. If $\mathbf{s}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$, it is equally likely that the system started in state 1 or in state 2.

Theorem 2.9.1

Let P be the transition matrix for an n -state Markov chain. If \mathbf{s}_m is the state vector at stage m , then

$$\mathbf{s}_{m+1} = P\mathbf{s}_m$$

for each $m = 0, 1, 2, \dots$

Heuristic Proof. Suppose that the Markov chain has been run N times, each time starting with the same initial state vector. Recall that p_{ij} is the proportion of the time the system goes from state j at some stage to state i at the next stage, whereas $s_i^{(m)}$ is the proportion of the time it is in state i at stage m . Hence

$$s_i^{(m+1)}N$$

is (approximately) the number of times the system is in state i at stage $m+1$. We are going to calculate this number another way. The system got to state i at stage $m+1$ through *some* other state (say state j) at stage m . The number of times it was *in* state j at that stage is (approximately) $s_j^{(m)}N$, so the number of times it got to state i via state j is $p_{ij}(s_j^{(m)}N)$. Summing over j gives the number of times the system is in state i (at stage $m+1$). This is the number we calculated before, so

$$s_i^{(m+1)}N = p_{i1}s_1^{(m)}N + p_{i2}s_2^{(m)}N + \dots + p_{in}s_n^{(m)}N$$

Dividing by N gives $s_i^{(m+1)} = p_{i1}s_1^{(m)} + p_{i2}s_2^{(m)} + \dots + p_{in}s_n^{(m)}$ for each i , and this can be expressed as the matrix equation $\mathbf{s}_{m+1} = P\mathbf{s}_m$. \square

If the initial probability vector \mathbf{s}_0 and the transition matrix P are given, Theorem 2.9.1 gives $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \dots$, one after the other, as follows:

$$\begin{aligned} \mathbf{s}_1 &= P\mathbf{s}_0 \\ \mathbf{s}_2 &= P\mathbf{s}_1 \\ \mathbf{s}_3 &= P\mathbf{s}_2 \\ &\vdots \end{aligned}$$

Hence, the state vector \mathbf{s}_m is completely determined for each $m = 0, 1, 2, \dots$ by P and \mathbf{s}_0 .

Example 2.9.3

A wolf pack always hunts in one of three regions R_1 , R_2 , and R_3 . Its hunting habits are as follows:

1. If it hunts in some region one day, it is as likely as not to hunt there again the next day.
2. If it hunts in R_1 , it never hunts in R_2 the next day.
3. If it hunts in R_2 or R_3 , it is equally likely to hunt in each of the other regions the next day.

If the pack hunts in R_1 on Monday, find the probability that it hunts there on Thursday.

Solution. The stages of this process are the successive days; the states are the three regions. The transition matrix P is determined as follows (see the table): The first habit asserts that $p_{11} = p_{22} = p_{33} = \frac{1}{2}$. Now column 1 displays what happens when the pack starts in R_1 : It never goes to state 2, so $p_{21} = 0$ and, because the column must sum to 1, $p_{31} = \frac{1}{2}$. Column 2 describes what happens if it starts in R_2 : $p_{22} = \frac{1}{2}$ and p_{12} and p_{32} are equal (by habit 3), so $p_{12} = p_{32} = \frac{1}{2}$ because the column sum must equal 1. Column 3 is filled in a similar way.

	R_1	R_2	R_3
R_1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
R_2	0	$\frac{1}{2}$	$\frac{1}{4}$
R_3	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$

Now let Monday be the initial stage. Then $\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ because the pack hunts in R_1 on that day.

Then \mathbf{s}_1 , \mathbf{s}_2 , and \mathbf{s}_3 describe Tuesday, Wednesday, and Thursday, respectively, and we compute them using Theorem 2.9.1.

$$\mathbf{s}_1 = P\mathbf{s}_0 = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix} \quad \mathbf{s}_2 = P\mathbf{s}_1 = \begin{bmatrix} \frac{3}{8} \\ \frac{1}{8} \\ \frac{4}{8} \end{bmatrix} \quad \mathbf{s}_3 = P\mathbf{s}_2 = \begin{bmatrix} \frac{11}{32} \\ \frac{6}{32} \\ \frac{15}{32} \end{bmatrix}$$

Hence, the probability that the pack hunts in Region R_1 on Thursday is $\frac{11}{32}$.

Steady State Vector

Another phenomenon that was observed in Example 2.9.1 can be expressed in general terms. The state vectors $\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2, \dots$ were calculated in that example and were found to “approach” $\mathbf{s} = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$. This means that the first component of \mathbf{s}_m becomes and remains very close to 0.2 as m becomes large, whereas the second component gets close to 0.8 as m increases. When this is the case, we say that \mathbf{s}_m **converges** to \mathbf{s} . For large m , then, there is very little error in taking $\mathbf{s}_m = \mathbf{s}$, so the long-term probability that the system is in state 1 is 0.2, whereas the probability that it is in state 2 is 0.8. In Example 2.9.1, enough state vectors were computed for the limiting vector \mathbf{s} to be apparent. However, there is a better way to do this that works in most cases.

Suppose P is the transition matrix of a Markov chain, and assume that the state vectors \mathbf{s}_m converge to a limiting vector \mathbf{s} . Then \mathbf{s}_m is very close to \mathbf{s} for sufficiently large m , so \mathbf{s}_{m+1} is also very close to \mathbf{s} . Thus, the equation $\mathbf{s}_{m+1} = P\mathbf{s}_m$ from Theorem 2.9.1 is closely approximated by

$$\mathbf{s} = P\mathbf{s}$$

so it is not surprising that \mathbf{s} should be a solution to this matrix equation. Moreover, it is easily solved because it can be written as a system of homogeneous linear equations

$$(I - P)\mathbf{s} = \mathbf{0}$$

with the entries of \mathbf{s} as variables.

In Example 2.9.1, where $P = \begin{bmatrix} 0 & 0.25 \\ 1 & 0.75 \end{bmatrix}$, the general solution to $(I - P)\mathbf{s} = \mathbf{0}$ is $\mathbf{s} = \begin{bmatrix} t \\ 4t \end{bmatrix}$, where t is a parameter. But if we insist that the entries of \mathbf{s} sum to 1 (as must be true of all state vectors), we find $t = 0.2$ and so $\mathbf{s} = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$ as before.

All this is predicated on the existence of a limiting vector for the sequence of state vectors of the Markov chain, and such a vector may not always exist. However, it does exist in one commonly occurring situation. A stochastic matrix P is called **regular** if some power P^m of P has every entry greater than zero. The matrix $P = \begin{bmatrix} 0 & 0.25 \\ 1 & 0.75 \end{bmatrix}$ of Example 2.9.1 is regular (in this case, each entry of P^2 is positive), and the general theorem is as follows:

Theorem 2.9.2

Let P be the transition matrix of a Markov chain and assume that P is regular. Then there is a unique column matrix \mathbf{s} satisfying the following conditions:

1. $P\mathbf{s} = \mathbf{s}$.
2. The entries of \mathbf{s} are positive and sum to 1.

Moreover, condition 1 can be written as

$$(I - P)\mathbf{s} = \mathbf{0}$$

and so gives a homogeneous system of linear equations for \mathbf{s} . Finally, the sequence of state vectors $\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2, \dots$ converges to \mathbf{s} in the sense that if m is large enough, each entry of \mathbf{s}_m is closely approximated by the corresponding entry of \mathbf{s} .

This theorem will not be proved here.²⁰

If P is the regular transition matrix of a Markov chain, the column \mathbf{s} satisfying conditions 1 and 2 of Theorem 2.9.2 is called the **steady-state vector** for the Markov chain. The entries of \mathbf{s} are the long-term probabilities that the chain will be in each of the various states.

Example 2.9.4

A man eats one of three soups—beef, chicken, and vegetable—each day. He never eats the same soup two days in a row. If he eats beef soup on a certain day, he is equally likely to eat each of the others the next day; if he does not eat beef soup, he is twice as likely to eat it the next day as the alternative.

- If he has beef soup one day, what is the probability that he has it again two days later?
- What are the long-run probabilities that he eats each of the three soups?

Solution. The states here are $B, C,$ and V , the three soups. The transition matrix P is given in the table. (Recall that, for each state, the corresponding column lists the probabilities for the next state.)

	B	C	V
B	0	$\frac{2}{3}$	$\frac{2}{3}$
C	$\frac{1}{2}$	0	$\frac{1}{3}$
V	$\frac{1}{2}$	$\frac{1}{3}$	0

If he has beef soup initially, then the initial state vector is

$$\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Then two days later the state vector is \mathbf{s}_2 . If P is the transition matrix, then

$$\mathbf{s}_1 = P\mathbf{s}_0 = \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{s}_2 = P\mathbf{s}_1 = \frac{1}{6} \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$

so he eats beef soup two days later with probability $\frac{2}{3}$. This answers (a.) and also shows that he eats chicken and vegetable soup each with probability $\frac{1}{6}$.

²⁰The interested reader can find an elementary proof in J. Kemeny, H. Mirkil, J. Snell, and G. Thompson, *Finite Mathematical Structures* (Englewood Cliffs, N.J.: Prentice-Hall, 1958).

To find the long-run probabilities, we must find the steady-state vector \mathbf{s} . Theorem 2.9.2 applies because P is regular (P^2 has positive entries), so \mathbf{s} satisfies $P\mathbf{s} = \mathbf{s}$. That is, $(I - P)\mathbf{s} = \mathbf{0}$ where

$$I - P = \frac{1}{6} \begin{bmatrix} 6 & -4 & -4 \\ -3 & 6 & -2 \\ -3 & -2 & 6 \end{bmatrix}$$

The solution is $\mathbf{s} = \begin{bmatrix} 4t \\ 3t \\ 3t \end{bmatrix}$, where t is a parameter, and we use $\mathbf{s} = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.3 \end{bmatrix}$ because the entries of \mathbf{s} must sum to 1. Hence, in the long run, he eats beef soup 40% of the time and eats chicken soup and vegetable soup each 30% of the time.

Exercises for 2.9

Exercise 2.9.1 Which of the following stochastic matrices is regular?

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

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

Exercise 2.9.2 In each case find the steady-state vector and, assuming that it starts in state 1, find the probability that it is in state 2 after 3 transitions.

a. $\begin{bmatrix} 0.5 & 0.3 \\ 0.5 & 0.7 \end{bmatrix}$

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

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

d. $\begin{bmatrix} 0.4 & 0.1 & 0.5 \\ 0.2 & 0.6 & 0.2 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}$

e. $\begin{bmatrix} 0.8 & 0.0 & 0.2 \\ 0.1 & 0.6 & 0.1 \\ 0.1 & 0.4 & 0.7 \end{bmatrix}$

f. $\begin{bmatrix} 0.1 & 0.3 & 0.3 \\ 0.3 & 0.1 & 0.6 \\ 0.6 & 0.6 & 0.1 \end{bmatrix}$

Exercise 2.9.3 A fox hunts in three territories A , B , and C . He never hunts in the same territory on two successive days. If he hunts in A , then he hunts in C the next day. If he hunts in B or C , he is twice as likely to hunt in A the next day as in the other territory.

- What proportion of his time does he spend in A , in B , and in C ?
- If he hunts in A on Monday (C on Monday), what is the probability that he will hunt in B on Thursday?

Exercise 2.9.4 Assume that there are three social classes—upper, middle, and lower—and that social mobility behaves as follows:

- Of the children of upper-class parents, 70% remain upper-class, whereas 10% become middle-class and 20% become lower-class.
- Of the children of middle-class parents, 80% remain middle-class, whereas the others are evenly split between the upper class and the lower class.
- For the children of lower-class parents, 60% remain lower-class, whereas 30% become middle-class and 10% upper-class.

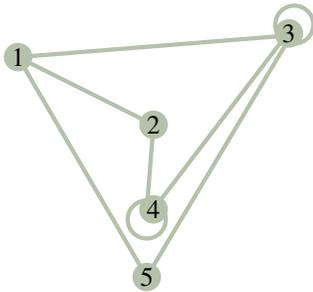
- Find the probability that the grandchild of lower-class parents becomes upper-class.
- Find the long-term breakdown of society into classes.

Exercise 2.9.5 The prime minister says she will call an election. This gossip is passed from person to person with a probability $p \neq 0$ that the information is passed incorrectly at any stage. Assume that when a person hears the gossip he or she passes it to one person who does not know. Find the long-term probability that a person will hear that there is going to be an election.

Exercise 2.9.6 John makes it to work on time one Monday out of four. On other work days his behaviour is as follows: If he is late one day, he is twice as likely to come to work on time the next day as to be late. If he is on time one day, he is as likely to be late as not the next day. Find the probability of his being late and that of his being on time Wednesdays.

Exercise 2.9.7 Suppose you have 1¢ and match coins with a friend. At each match you either win or lose 1¢ with equal probability. If you go broke or ever get 4¢ , you quit. Assume your friend never quits. If the states are $0, 1, 2, 3,$ and 4 representing your wealth, show that the corresponding transition matrix P is not regular. Find the probability that you will go broke after 3 matches.

Exercise 2.9.8 A mouse is put into a maze of compartments, as in the diagram. Assume that he always leaves any compartment he enters and that he is equally likely to take any tunnel entry.



- If he starts in compartment 1, find the probability that he is in compartment 1 again after 3 moves.
- Find the compartment in which he spends most of his time if he is left for a long time.

Exercise 2.9.9 If a stochastic matrix has a 1 on its main diagonal, show that it cannot be regular. Assume it is not 1×1 .

Exercise 2.9.10 If \mathbf{s}_m is the stage- m state vector for a Markov chain, show that $\mathbf{s}_{m+k} = P^k \mathbf{s}_m$ holds for all $m \geq 1$ and $k \geq 1$ (where P is the transition matrix).

Exercise 2.9.11 A stochastic matrix is **doubly stochastic** if all the row sums also equal 1. Find the steady-state vector for a doubly stochastic matrix.

Exercise 2.9.12 Consider the 2×2 stochastic matrix

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

where $0 < p < 1$ and $0 < q < 1$.

- Show that $\frac{1}{p+q} \begin{bmatrix} q \\ p \end{bmatrix}$ is the steady-state vector for P .
- Show that P^m converges to the matrix $\frac{1}{p+q} \begin{bmatrix} q & q \\ p & p \end{bmatrix}$ by first verifying inductively that $P^m = \frac{1}{p+q} \begin{bmatrix} q & q \\ p & p \end{bmatrix} + \frac{(1-p-q)^m}{p+q} \begin{bmatrix} p & -q \\ -p & q \end{bmatrix}$ for $m = 1, 2, \dots$ (It can be shown that the sequence of powers P, P^2, P^3, \dots of any regular transition matrix converges to the matrix each of whose columns equals the steady-state vector for P .)

Supplementary Exercises for Chapter 2

Exercise 2.1 Solve for the matrix X if:

a. $PXQ = R$; b. $XP = S$;

$$\text{where } P = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 0 & 3 \end{bmatrix}, Q = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 3 \end{bmatrix},$$

$$R = \begin{bmatrix} -1 & 1 & -4 \\ -4 & 0 & -6 \\ 6 & 6 & -6 \end{bmatrix}, S = \begin{bmatrix} 1 & 6 \\ 3 & 1 \end{bmatrix}$$

Exercise 2.2 Consider

$$p(X) = X^3 - 5X^2 + 11X - 4I.$$

- a. If $p(U) = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix}$ compute $p(U^T)$.
- b. If $p(U) = 0$ where U is $n \times n$, find U^{-1} in terms of U .

Exercise 2.3 Show that, if a (possibly nonhomogeneous) system of equations is consistent and has more variables than equations, then it must have infinitely many solutions. [Hint: Use Theorem 2.2.2 and Theorem 1.3.1.]

Exercise 2.4 Assume that a system $A\mathbf{x} = \mathbf{b}$ of linear equations has at least two distinct solutions \mathbf{y} and \mathbf{z} .

- a. Show that $\mathbf{x}_k = \mathbf{y} + k(\mathbf{y} - \mathbf{z})$ is a solution for every k .
- b. Show that $\mathbf{x}_k = \mathbf{x}_m$ implies $k = m$. [Hint: See Example 2.1.7.]
- c. Deduce that $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.

Exercise 2.5

- a. Let A be a 3×3 matrix with all entries on and below the main diagonal zero. Show that $A^3 = 0$.
- b. Generalize to the $n \times n$ case and prove your answer.

Exercise 2.6 Let I_{pq} denote the $n \times n$ matrix with (p, q) -entry equal to 1 and all other entries 0. Show that:

- a. $I_n = I_{11} + I_{22} + \cdots + I_{nn}$.
- b. $I_{pq}I_{rs} = \begin{cases} I_{ps} & \text{if } q = r \\ 0 & \text{if } q \neq r \end{cases}$.
- c. If $A = [a_{ij}]$ is $n \times n$, then $A = \sum_{i=1}^n \sum_{j=1}^n a_{ij}I_{ij}$.
- d. If $A = [a_{ij}]$, then $I_{pq}AI_{rs} = a_{qr}I_{ps}$ for all p, q, r , and s .

Exercise 2.7 A matrix of the form aI_n , where a is a number, is called an $n \times n$ **scalar matrix**.

- a. Show that each $n \times n$ scalar matrix commutes with every $n \times n$ matrix.
- b. Show that A is a scalar matrix if it commutes with every $n \times n$ matrix. [Hint: See part (d.) of Exercise 2.6.]

Exercise 2.8 Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, where A, B, C , and D are all $n \times n$ and each commutes with all the others. If $M^2 = 0$, show that $(A + D)^3 = 0$. [Hint: First show that $A^2 = -BC = D^2$ and that

$$B(A + D) = 0 = C(A + D).]$$

Exercise 2.9 If A is 2×2 , show that $A^{-1} = A^T$ if and only if $A = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$ for some θ or

$$A = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix} \text{ for some } \theta.$$

[Hint: If $a^2 + b^2 = 1$, then $a = \cos \theta$, $b = \sin \theta$ for some θ . Use

$$\cos(\theta - \phi) = \cos \theta \cos \phi + \sin \theta \sin \phi.]$$

Exercise 2.10

- a. If $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, show that $A^2 = I$.
- b. What is wrong with the following argument? If $A^2 = I$, then $A^2 - I = 0$, so $(A - I)(A + I) = 0$, whence $A = I$ or $A = -I$.