

## 11.3: Ergodic Markov Chains\*\*

A second important kind of Markov chain we shall study in detail is an ergodic Markov chain, defined as follows.

### Definition: ergodic chain

A Markov chain is called an ergodic Markov chain if it is possible to go from every state to every state (not necessarily in one move).

In many books, ergodic Markov chains are called *irreducible*.

### Definition: Markov chain

A Markov chain is called an ergodic chain if some power of the transition matrix has only positive elements.

In other words, for some  $n$ , it is possible to go from any state to any state in exactly  $n$  steps. It is clear from this definition that every regular chain is ergodic. On the other hand, an ergodic chain is not necessarily regular, as the following examples show.

### ✓ Example 11.3.1

Let the transition matrix of a Markov chain be defined by

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 1 & 2 \end{matrix} \\ \begin{matrix} 1 \\ 2 \end{matrix} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{matrix} \quad (11.3.1)$$

Then it is clear that it is possible to move from any state to any state, so the chain is ergodic. However, if  $n$  is odd, then it is not possible to move from state 0 to state 0 in  $n$  steps, and if  $n$  is even, then it is not possible to move from state 0 to state 1 in  $n$  steps, so the chain is not regular.

A more interesting example of an ergodic, non-regular Markov chain is provided by the Ehrenfest urn model.

### ✓ Example 11.3.2

Recall the Ehrenfest urn model (Example 11.1.8). The transition matrix for this example is

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad (11.3.2)$$

In this example, if we start in state 0 we will, after any even number of steps, be in either state 0, 2 or 4, and after any odd number of steps, be in states 1 or 3. Thus this chain is ergodic but not regular.

## Regular Markov Chains

Any transition matrix that has no zeros determines a regular Markov chain. However, it is possible for a regular Markov chain to have a transition matrix that has zeros. The transition matrix of the Land of Oz example of Section 11.1 has  $p_{NN} = 0$  but the second power  $\mathbf{P}^2$  has no zeros, so this is a regular Markov chain.

An example of a nonregular Markov chain is an absorbing chain. For example, let

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix} \quad (11.3.3)$$

be the transition matrix of a Markov chain. Then all powers of  $\mathbf{P}$  will have a 0 in the upper right-hand corner.

We shall now discuss two important theorems relating to regular chains.

### Theorem 11.3.1

Let  $\mathbf{P}$  be the transition matrix for a regular chain. Then, as  $n \rightarrow \infty$ , the powers  $\mathbf{P}^n$  approach a limiting matrix  $\mathbf{W}$  with all rows the same vector  $\mathbf{w}$ . The vector  $\mathbf{w}$  is a strictly positive probability vector (i.e., the components are all positive and they sum to one).

In the next section we give two proofs of this fundamental theorem. We give here the basic idea of the first proof.

We want to show that the powers  $\mathbf{P}^n$  of a regular transition matrix tend to a matrix with all rows the same. This is the same as showing that  $\mathbf{P}^n$  converges to a matrix with constant columns. Now the  $j$ th column of  $\mathbf{P}^n$  is  $\mathbf{P}^n \mathbf{y}$  where  $\mathbf{y}$  is a column vector with 1 in the  $j$ th entry and 0 in the other entries. Thus we need only prove that for any column vector  $\mathbf{y}$ ,  $\mathbf{P}^n \mathbf{y}$  approaches a constant vector as  $n$  tend to infinity.

Since each row of  $\mathbf{P}$  is a probability vector,  $\mathbf{P}\mathbf{y}$  replaces  $\mathbf{y}$  by averages of its components. Here is an example:

$$\begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/2 \cdot 1 + 1/4 \cdot 2 + 1/4 \cdot 3 \\ 1/3 \cdot 1 + 1/3 \cdot 2 + 1/3 \cdot 3 \\ 1/2 \cdot 1 + 1/2 \cdot 2 + 0 \cdot 3 \end{pmatrix} = \begin{pmatrix} 7/4 \\ 2 \\ 3/2 \end{pmatrix}. \quad (11.3.4)$$

The result of the averaging process is to make the components of  $\mathbf{P}\mathbf{y}$  more similar than those of  $\mathbf{y}$ . In particular, the maximum component decreases (from 3 to 2) and the minimum component increases (from 1 to 3/2). Our proof will show that as we do more and more of this averaging to get  $\mathbf{P}^n \mathbf{y}$ , the difference between the maximum and minimum component will tend to 0 as  $n \rightarrow \infty$ . This means  $\mathbf{P}^n \mathbf{y}$  tends to a constant vector. The  $ij$ th entry of  $\mathbf{P}^n$ ,  $p_{ij}^{(n)}$ , is the probability that the process will be in state  $s_j$  after  $n$  steps if it starts in state  $s_i$ . If we denote the common row of  $\mathbf{W}$  by  $\mathbf{w}$ , then Theorem 11.3.1 states that the probability of being in  $s_j$  in the long run is approximately  $w_j$ , the  $j$ th entry of  $\mathbf{w}$ , and is independent of the starting state.

### Example 11.3.3

Recall that for the Land of Oz example of Section 1.1, the sixth power of the transition matrix  $\mathbf{P}$  is, to three decimal places,

$$\mathbf{P}^6 = \begin{matrix} & \begin{matrix} R & N & S \end{matrix} \\ \begin{matrix} R \\ N \\ S \end{matrix} & \begin{pmatrix} .4 & .2 & .4 \\ .4 & .2 & .4 \\ .4 & .2 & .4 \end{pmatrix} \end{matrix} \quad (11.3.5)$$

Thus, to this degree of accuracy, the probability of rain six days after a rainy day is the same as the probability of rain six days after a nice day, or six days after a snowy day. Theorem 11.3.1 predicts that, for large  $n$ , the rows of  $\mathbf{P}$  approach a common vector. It is interesting that this occurs so soon in our example.

### Theorem 11.3.2

Let  $\mathbf{P}$  be a regular transition matrix, let

$$\mathbf{W} = \lim_{n \rightarrow \infty} \mathbf{P}^n, \quad (11.3.6)$$

let  $\mathbf{w}$  be the common row of  $\mathbf{W}$ , and let  $\mathbf{c}$  be the column vector all of whose components are 1. Then

- $\mathbf{w}\mathbf{P} = \mathbf{w}$ , and any row vector  $\mathbf{v}$  such that  $\mathbf{v}\mathbf{P} = \mathbf{v}$  is a constant multiple of  $\mathbf{w}$ .

b.  $\mathbf{Pc} = \mathbf{c}$ , and any column vector such that  $\mathbf{Px} = \mathbf{x}$  is a multiple of  $\mathbf{c}$ .

**Proof.** To prove part (a), we note that from Theorem 11.3.1,

$$\mathbf{P}^n \rightarrow \mathbf{W}. \quad (11.3.7)$$

Thus,

$$\mathbf{P}^{n+1} = \mathbf{P}^n \cdot \mathbf{P} \rightarrow \mathbf{WP}. \quad (11.3.8)$$

But  $\mathbf{P}^{n+1} \rightarrow \mathbf{W}$ , and so  $\mathbf{W} = \mathbf{WP}$ , and  $\mathbf{w} = \mathbf{wP}$ .

Let  $\mathbf{v}$  be any vector with  $\mathbf{vP} = \mathbf{v}$ . Then  $\mathbf{v} = \mathbf{vP}^n$ , and passing to the limit,  $\mathbf{v} = \mathbf{vW}$ . Let  $r$  be the sum of the components of  $\mathbf{v}$ . Then it is easily checked that  $\mathbf{vW} = r\mathbf{w}$ . So,  $\mathbf{v} = r\mathbf{w}$ .

To prove part (b), assume that  $\mathbf{x} = \mathbf{Px}$ . Then  $\mathbf{x} = \mathbf{P}^n\mathbf{x}$ , and again passing to the limit,  $\mathbf{x} = \mathbf{Wx}$ . Since all rows of  $\mathbf{W}$  are the same, the components of  $\mathbf{Wx}$  are all equal, so  $\mathbf{x}$  is a multiple of  $\mathbf{c}$ .

Note that an immediate consequence of Theorem 11.3.2 is the fact that there is only one probability vector  $\mathbf{v}$  such that  $\mathbf{vP} = \mathbf{v}$ .

## Fixed Vectors

### Definition: 11.3.1: Row and Column Vectors

A row vector  $\mathbf{w}$  with the property  $\mathbf{wP} = \mathbf{w}$  is called a *fixed row vector* for  $\mathbf{P}$ . Similarly, a column vector  $\mathbf{x}$  such that  $\mathbf{Px} = \mathbf{x}$  is called a *fixed column vector* for  $\mathbf{P}$ .

Thus, the common row of  $\mathbf{W}$  is the unique vector  $\mathbf{w}$  which is both a fixed row vector for  $\mathbf{P}$  and a probability vector. Theorem 11.3.2 shows that any fixed row vector for  $\mathbf{P}$  is a multiple of  $\mathbf{w}$  and any fixed column vector for  $\mathbf{P}$  is a constant vector.

One can also state Definition 11.3.1 in terms of eigenvalues and eigenvectors. A fixed row vector is a left eigenvector of the matrix  $\mathbf{P}$  corresponding to the eigenvalue 1. A similar statement can be made about fixed column vectors.

We will now give several different methods for calculating the fixed row vector for a regular Markov chain.

### ✓ Example 11.3.4

By Theorem 11.3.1 we can find the limiting vector  $\mathbf{w}$  for the Land of Oz from the fact that

$$w_1 + w_2 + w_3 = 1 \quad (11.3.9)$$

and

$$\begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix} \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} = \begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix}. \quad (11.3.10)$$

These relations lead to the following four equations in three unknowns:

$$\begin{aligned} w_1 + w_2 + w_3 &= 1, \\ (1/2)w_1 + (1/2)w_2 + (1/4)w_3 &= w_1, \\ (1/4)w_1 + (1/4)w_3 &= w_2, \\ (1/4)w_1 + (1/2)w_2 + (1/2)w_3 &= w_3. \end{aligned}$$

Our theorem guarantees that these equations have a unique solution. If the equations are solved, we obtain the solution

$$\mathbf{w} = (.4 \quad .2 \quad .4), \quad (11.3.11)$$

in agreement with that predicted from  $\mathbf{P}^6$ , given in Example 11.1.2.

To calculate the fixed vector, we can assume that the value at a particular state, say state one, is 1, and then use all but one of the linear equations from  $\mathbf{wP} = \mathbf{w}$ . This set of equations will have a unique solution and we can obtain  $\mathbf{w}$  from this solution

by dividing each of its entries by their sum to give the probability vector  $\mathbf{w}$ . We will now illustrate this idea for the above example.

### ✓ Example 11.3.5

(Example 11.3.1 continued) We set  $w_1 = 1$ , and then solve the first and second linear equations from  $\mathbf{wP} = \mathbf{w}$ . We have

$$\begin{aligned} (1/2) + (1/2)w_2 + (1/4)w_3 &= 1, \\ (1/4) + (1/4)w_3 &= w_2. \end{aligned}$$

If we solve these, we obtain

$$(w_1 \ w_2 \ w_3) = (1 \ 1/2 \ 1). \quad (11.3.12)$$

Now we divide this vector by the sum of the components, to obtain the final answer:

$$\mathbf{w} = (.4 \ .2 \ .4). \quad (11.3.13)$$

This method can be easily programmed to run on a computer.

As mentioned above, we can also think of the fixed row vector  $\mathbf{w}$  as a left eigenvector of the transition matrix  $\mathbf{P}$ . Thus, if we write  $\mathbf{I}$  to denote the identity matrix, then  $\mathbf{w}$  satisfies the matrix equation

$$\mathbf{wP} = \mathbf{wI}, \quad (11.3.14)$$

or equivalently,

$$\mathbf{w}(\mathbf{P} - \mathbf{I}) = \mathbf{0}. \quad (11.3.15)$$

Thus,  $\mathbf{w}$  is in the left nullspace of the matrix  $\mathbf{P} - \mathbf{I}$ . Furthermore, Theorem 11.3.2 states that this left nullspace has dimension 1. Certain computer programming languages can find nullspaces of matrices. In such languages, one can find the fixed row probability vector for a matrix  $\mathbf{P}$  by computing the left nullspace and then normalizing a vector in the nullspace so the sum of its components is 1.

The program **FixedVector** uses one of the above methods (depending upon the language in which it is written) to calculate the fixed row probability vector for regular Markov chains.

So far we have always assumed that we started in a specific state. The following theorem generalizes Theorem [thm 11.3.6] to the case where the starting state is itself determined by a probability vector.

### 🔗 Theorem 11.3.3

Let  $\mathbf{P}$  be the transition matrix for a regular chain and  $\mathbf{v}$  an arbitrary probability vector. Then

$$\lim_{n \rightarrow \infty} \mathbf{vP}^n = \mathbf{w}, \quad (11.3.16)$$

where  $\mathbf{w}$  is the unique fixed probability vector for  $\mathbf{P}$ .

**Proof.** By Theorem 11.3.1,

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \mathbf{W}. \quad (11.3.17)$$

Hence,

$$\lim_{n \rightarrow \infty} \mathbf{vP}^n = \mathbf{vW}. \quad (11.3.18)$$

But the entries in  $\mathbf{v}$  sum to 1, and each row of  $\mathbf{W}$  equals  $\mathbf{w}$ . From these statements, it is easy to check that

$$\mathbf{vW} = \mathbf{w}. \quad (11.3.19)$$

If we start a Markov chain with initial probabilities given by  $\mathbf{v}$ , then the probability vector  $\mathbf{vP}^n$  gives the probabilities of being in the various states after  $n$  steps. Theorem [thm 11.3.9] then establishes the fact that, even in this more general class of processes, the probability of being in  $s_j$  approaches  $w_j$ .

## Equilibrium

We also obtain a new interpretation for  $\mathbf{w}$ . Suppose that our starting vector picks state  $s_i$  as a starting state with probability  $w_i$ , for all  $i$ . Then the probability of being in the various states after  $n$  steps is given by  $\mathbf{wP}^n = \mathbf{w}$ , and is the same on all steps. This method of starting provides us with a process that is called "stationary." The fact that  $\mathbf{w}$  is the only probability vector for which  $\mathbf{wP} = \mathbf{w}$  shows that we must have a starting probability vector of exactly the kind described to obtain a stationary process.

Many interesting results concerning regular Markov chains depend only on the fact that the chain has a unique fixed probability vector which is positive. This property holds for all ergodic Markov chains.

### Theorem 11.3.4

*For an ergodic Markov chain, there is a unique probability vector  $\mathbf{w}$  such that  $\mathbf{wP} = \mathbf{w}$  and  $\mathbf{w}$  is strictly positive. Any row vector such that  $\mathbf{vP} = \mathbf{v}$  is a multiple of  $\mathbf{w}$ . Any column vector  $\mathbf{x}$  such that  $\mathbf{Px} = \mathbf{x}$  is a constant vector.*

**Proof.** This theorem states that Theorem 11.3.2 is true for ergodic chains. The result follows easily from the fact that, if  $\mathbf{P}$  is an ergodic transition matrix, then  $\bar{\mathbf{P}} = (1/2)\mathbf{I} + (1/2)\mathbf{P}$  is a regular transition matrix with the same fixed vectors (see Exercises 11.3.25–[exer 11.3.28]).

For ergodic chains, the fixed probability vector has a slightly different interpretation. The following two theorems, which we will not prove here, furnish an interpretation for this fixed vector.

### Theorem 11.3.5

*Let  $\mathbf{P}$  be the transition matrix for an ergodic chain. Let  $\mathbf{A}_n$  be the matrix defined by*

$$\mathbf{A}_n = \frac{\mathbf{I} + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^n}{n+1}. \quad (11.3.20)$$

*Then  $\mathbf{A}_n \rightarrow \mathbf{W}$ , where  $\mathbf{W}$  is a matrix all of whose rows are equal to the unique fixed probability vector  $\mathbf{w}$  for  $\mathbf{P}$ .*

If  $\mathbf{P}$  is the transition matrix of an ergodic chain, then Theorem 11.3.2 states that there is only one fixed row probability vector for  $\mathbf{P}$ . Thus, we can use the same techniques that were used for regular chains to solve for this fixed vector. In particular, the program **FixedVector** works for ergodic chains.

To interpret Theorem 11.3.5 let us assume that we have an ergodic chain that starts in state  $s_i$ . Let  $X^{(m)} = 1$  if the  $m$ th step is to state  $s_j$  and 0 otherwise. Then the average number of times in state  $s_j$  in the first  $n$  steps is given by

$$H^{(n)} = \frac{X^{(0)} + X^{(1)} + X^{(2)} + \cdots + X^{(n)}}{n+1}. \quad (11.3.21)$$

But  $X^{(m)}$  takes on the value 1 with probability  $p_{ij}^{(m)}$  and 0 otherwise. Thus  $E(X^{(m)}) = p_{ij}^{(m)}$ , and the  $ij$ th entry of  $\mathbf{A}_n$  gives the expected value of  $H^{(n)}$ , that is, the expected proportion of times in state  $s_j$  in the first  $n$  steps if the chain starts in state  $s_i$ .

If we call being in state  $s_j$  and any other state we could ask if a theorem analogous to the law of large numbers for independent trials holds. The answer is yes and is given by the following theorem.

### Theorem (Law of Large Numbers for Ergodic Markov Chains) 11.3.6

*Let  $H_j^{(n)}$  be the proportion of times in  $n$  steps that an ergodic chain is in state  $s_j$ . Then for any  $\epsilon > 0$ ,*

$$P\left(|H_j^{(n)} - w_j| > \epsilon\right) \rightarrow 0, \quad (11.3.22)$$

*independent of the starting state  $s_i$ .*

We have observed that every regular Markov chain is also an ergodic chain. Hence, Theorems 11.3.5 and 11.3.6 apply also for regular chains. For example, this gives us a new interpretation for the fixed vector  $\mathbf{w} = (.4, .2, .4)$  in the Land of Oz example. Theorem [\PageIndex{5}](#) predicts that, in the long run, it will rain 40 percent of the time in the Land of Oz, be nice 20 percent of the time, and snow 40 percent of the time.

## Simulation

We illustrate Theorem 11.3.6 by writing a program to simulate the behavior of a Markov chain. **SimulateChain** is such a program.

### ✓ Example 11.3.6

In the Land of Oz, there are 525 days in a year. We have simulated the weather for one year in the Land of Oz, using the program **SimulateChain**. The results are shown in Table 11.3.2.

```
SSRNRNSSSSSSNRSNSSRNSRNSSSNSRRRNSSSNRRSSSSNRSSNSRRRRRRRNSSS
SSRRRSNSNRNRNRSSRNSNSRRNRNRNRSSNSRNRNSRRRSRNSSSNSRNRSSNRNR
RNSSSNSNSSNRSSRNSNSSRNSRRNRNRNRSSNSRNRNSRNSNRNRSSSRSS
NRSSNSSSSSSSNSNSRRNRNRNRNRSSSNRRSSSSNRSSSSRNRNRNRSSSR
RNRNRSSRRRRSSNRNRNRNRNRSSNRNSNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR
RRRSSNRNRNRNSNSSSSRRRRSRNRSSRRRRSSRRNRNRNRNRNRNRNRNRNRNRNR
RNSNRNSNRNRNRNRNRNRSSNRSSRNSRNSNSNRNSNSSSNRNRNRNRNRNRNRNR
SSNSRNSNRNRNRNRNRNRSSSRNSRRSSNRNRNRNRNRNRNRNRNRNRNRNRNRNR
NSRRRNSSRRRNSSSNRRSRNSRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR
SNS
```

Table 11.3.2: Weather in the Land of Oz.

State	Times	Fraction
R	217	.413
N	109	.208
S	199	.379

We note that the simulation gives a proportion of times in each of the states not too different from the long run predictions of .4, .2, and .4 assured by Theorem 11.3.1. To get better results we have to simulate our chain for a longer time. We do this for 10,000 days without printing out each day's weather. The results are shown in Table 11.3.3. We see that the results are now quite close to the theoretical values of .4, .2, and .4.

Table 11.3.3: Comparison of observed and predicted frequencies for the Land of Oz.

State	Times	Fraction
R	4010	.401
N	1902	.19
S	4088	.409

## Examples of Ergodic Chains

The computation of the fixed vector  $\mathbf{w}$  may be difficult if the transition matrix is very large. It is sometimes useful to guess the fixed vector on purely intuitive grounds. Here is a simple example to illustrate this kind of situation.

### ✓ Example 11.3.7

A white rat is put into the maze of Figure 11.3.1

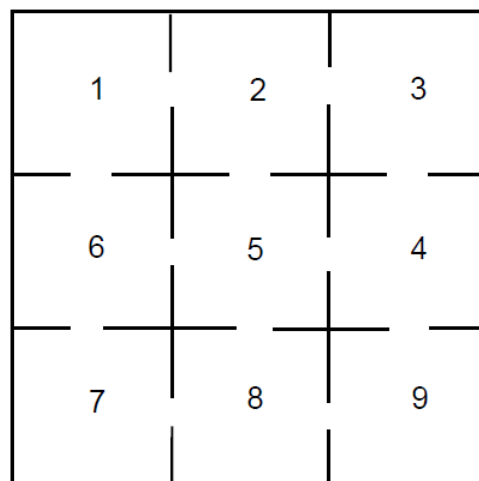


Figure 11.3.1: The maze problem.

There are nine compartments with connections between the compartments as indicated. The rat moves through the compartments at random. That is, if there are  $k$  ways to leave a compartment, it chooses each of these with equal probability. We can represent the travels of the rat by a Markov chain process with transition matrix given by

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{matrix} & \begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 & 1/3 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \end{pmatrix} \end{matrix} \quad (11.3.23)$$

That this chain is not regular can be seen as follows: From an odd-numbered state the process can go only to an even-numbered state, and from an even-numbered state it can go only to an odd number. Hence, starting in state  $i$  the process will be alternately in even-numbered and odd-numbered states. Therefore, odd powers of  $\mathbf{P}$  will have 0's for the odd-numbered entries in row 1. On the other hand, a glance at the maze shows that it is possible to go from every state to every other state, so that the chain is ergodic.

To find the fixed probability vector for this matrix, we would have to solve ten equations in nine unknowns. However, it would seem reasonable that the times spent in each compartment should, in the long run, be proportional to the number of entries to each compartment. Thus, we try the vector whose  $j$ th component is the number of entries to the  $j$ th compartment:

$$\mathbf{x} = (2 \ 3 \ 2 \ 3 \ 4 \ 3 \ 2 \ 3 \ 2) . \quad (11.3.24)$$

It is easy to check that this vector is indeed a fixed vector so that the unique probability vector is this vector normalized to have sum 1:

$$\mathbf{w} = \left( \frac{1}{12} \quad \frac{1}{8} \quad \frac{1}{12} \quad \frac{1}{8} \quad \frac{1}{6} \quad \frac{1}{8} \quad \frac{1}{12} \quad \frac{1}{8} \quad \frac{1}{12} \right) . \quad (11.3.25)$$

### ✓ Example 11.3.8

Add example text here.

#### Solution

Add example text here.

(Example 11.3.7 continued) We recall the Ehrenfest urn model of Example 11.1.6. The transition matrix for this chain is as follows:

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} .000 & 1.000 & .000 & .000 & .000 \\ .250 & .000 & .750 & .000 & .000 \\ .000 & .500 & .000 & .500 & .000 \\ .000 & .000 & .750 & .000 & .250 \\ .000 & .000 & .000 & 1.000 & .000 \end{pmatrix} \end{matrix} \quad (11.3.26)$$

If we run the program **FixedVector** for this chain, we obtain the vector

$$\mathbf{P}^6 = \begin{matrix} & 0 & 1 & 2 & 3 & 4 \end{matrix} \quad (11.3.27)$$

$$\begin{pmatrix} .0625 & .2500 & .3750 & .2500 & .0625 \end{pmatrix}$$

By Theorem 11.3.7, we can interpret these values for  $w_i$  as the proportion of times the process is in each of the states in the long run. For example, the proportion of times in state 0 is .0625 and the proportion of times in state 1 is .375. The astute reader will note that these numbers are the binomial distribution  $1/16, 4/16, 6/16, 4/16, 1/16$ . We could have guessed this answer as follows: If we consider a particular ball, it simply moves randomly back and forth between the two urns. This suggests that the equilibrium state should be just as if we randomly distributed the four balls in the two urns. If we did this, the probability that there would be exactly  $j$  balls in one urn would be given by the binomial distribution  $b(n, p, j)$  with  $n = 4$  and  $p = 1/2$ .

## Exercises

### Exercise 11.3.1

Which of the following matrices are transition matrices for regular Markov chains?

- $\mathbf{P} = \begin{pmatrix} .5 & .5 \\ .5 & .5 \end{pmatrix}.$
- $\mathbf{P} = \begin{pmatrix} .5 & .5 \\ 1 & 0 \end{pmatrix}.$
- $\mathbf{P} = \begin{pmatrix} 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \\ 0 & 1/5 & 4/5 \end{pmatrix}.$
- $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$
- $\mathbf{P} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$

### Exercise 11.3.2

Consider the Markov chain with transition matrix



$$\mathbf{P} = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 3/4 & 0 & 1/4 \\ 0 & 1 & 0 \end{pmatrix}. \quad (11.3.28)$$

- Show that this is a regular Markov chain.
- The process is started in state 1; find the probability that it is in state 3 after two steps.
- Find the limiting probability vector  $\mathbf{w}$ .

### Exercise 11.3.3

Consider the Markov chain with general  $2 \times 2$  transition matrix

$$\mathbf{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}. \quad (11.3.29)$$

- Under what conditions is  $\mathbf{P}$  absorbing?
- Under what conditions is  $\mathbf{P}$  ergodic but not regular?
- Under what conditions is  $\mathbf{P}$  regular?

### Exercise 11.3.4

Find the fixed probability vector  $\mathbf{w}$  for the matrices in Exercise [exer 11.3.3] that are ergodic.

### Exercise 11.3.5

Find the fixed probability vector  $\mathbf{w}$  for each of the following regular matrices.

- $\mathbf{P} = \begin{pmatrix} .75 & .25 \\ .5 & .5 \end{pmatrix}.$
- $\mathbf{P} = \begin{pmatrix} .9 & .1 \\ .1 & .9 \end{pmatrix}.$
- $\mathbf{P} = \begin{pmatrix} 3/4 & 1/4 & 0 \\ 0 & 2/3 & 1/3 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$

### Exercise 11.3.6

Consider the Markov chain with transition matrix in Exercise 11.3.3 with  $a = b = 1$ . Show that this chain is ergodic but not regular. Find the fixed probability vector and interpret it. Show that  $\mathbf{P}^n$  does not tend to a limit, but that

$$\mathbf{A}_n = \frac{\mathbf{I} + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^n}{n+1} \quad (11.3.30)$$

does.

### Exercise 11.3.7

Consider the Markov chain with transition matrix of Exercise 11.3.3 with  $a = 0$  and  $b = 1/2$ . Compute directly the unique fixed probability vector, and use your result to prove that the chain is not ergodic.

### Exercise 11.3.8

Show that the matrix

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 0 & 1 \end{pmatrix} \quad (11.3.31)$$

has more than one fixed probability vector. Find the matrix that  $\mathbf{P}^n$  approaches as  $n \rightarrow \infty$ , and verify that it is not a matrix all of whose rows are the same.

### Exercise 11.3.9

Prove that, if a 3-by-3 transition matrix has the property that its sums are 1, then  $(1/3, 1/3, 1/3)$  is a fixed probability vector. State a similar result for  $n$ -by- $n$  transition matrices. Interpret these results for ergodic chains.

### Exercise 11.3.10

Is the Markov chain in Example 11.1.8 ergodic?

### Exercise 11.3.11

Is the Markov chain in Example 11.1.9 ergodic?

### Exercise 11.3.12

Consider Example 11.2.1 (Drunkard's Walk). Assume that if the walker reaches state 0, he turns around and returns to state 1 on the next step and, similarly, if he reaches 4 he returns on the next step to state 3. Is this new chain ergodic? Is it regular?

### Exercise 11.3.13

For Example 11.1.2 when  $\mathbf{P}$  is ergodic, what is the proportion of people who are told that the President will run? Interpret the fact that this proportion is independent of the starting state.

### Exercise 11.3.14

Consider an independent trials process to be a Markov chain whose states are the possible outcomes of the individual trials. What is its fixed probability vector? Is the chain always regular? Illustrate this for Example 11.1.3.

### Exercise 11.3.15

Show that Example 11.1.6 is an ergodic chain, but not a regular chain. Show that its fixed probability vector  $\mathbf{w}$  is a binomial distribution.

### Exercise 11.3.16

Show that Example 11.1.7 is regular and find the limiting vector.

### Exercise 11.3.17

Toss a fair die repeatedly. Let  $S_n$  denote the total of the outcomes through the  $n$ th toss. Show that there is a limiting value for the proportion of the first  $n$  values of  $S_n$  that are divisible by 7, and compute the value for this limit. : The desired limit is an equilibrium probability vector for an appropriate seven state Markov chain.

### Exercise 11.3.18

Let  $\mathbf{P}$  be the transition matrix of a regular Markov chain. Assume that there are  $r$  states and let  $N(r)$  be the smallest integer  $n$  such that  $\mathbf{P}^n$  is regular if and only if  $\mathbf{P}^{N(r)}$  has no zero entries. Find a finite upper bound for  $N(r)$ . See if you can determine  $N(3)$  exactly.

### Exercise 11.3.19

Define  $f(r)$  to be the smallest integer  $n$  such that for all regular Markov chains with  $r$  states, the  $n$ th power of the transition matrix has all entries positive. It has been shown,<sup>14</sup> that  $f(r) = r^2 - 2r + 2$ .

- Define the transition matrix of an  $r$ -state Markov chain as follows: For states  $s_i$ , with  $i = 1, 2, \dots, r-2$ ,  $\mathbf{P}(i, i+1) = 1$ ,  $\mathbf{P}(r-1, r) = \mathbf{P}(r-1, 1) = 1/2$ , and  $\mathbf{P}(r, 1) = 1$ . Show that this is a regular Markov chain.
- For  $r = 3$ , verify that the fifth power is the first power that has no zeros.
- Show that, for general  $r$ , the smallest  $n$  such that  $\mathbf{P}^n$  has all entries positive is  $n = f(r)$ .

### Exercise 11.3.20

A discrete time queueing system of capacity  $n$  consists of the person being served and those waiting to be served. The queue length  $x$  is observed each second. If  $0 < x < n$ , then with probability  $p$ , the queue size is increased by one by an arrival and, independently, with probability  $r$ , it is decreased by one because the person being served finishes service. If  $x = 0$ , only an arrival (with probability  $p$ ) is possible. If  $x = n$ , an arrival will depart without waiting for service, and so only the departure (with probability  $r$ ) of the person being served is possible. Form a Markov chain with states given by the number of customers in the

queue. Modify the program **FixedVector** so that you can input  $n$ ,  $p$ , and  $r$ , and the program will construct the transition matrix and compute the fixed vector. The quantity  $s = p/r$  is called the Describe the differences in the fixed vectors according as  $s < 1$ ,  $s = 1$ , or  $s > 1$ .

### Exercise 11.3.21

Write a computer program to simulate the queue in Exercise 11.3.20 Have your program keep track of the proportion of the time that the queue length is  $j$  for  $j = 0, 1, \dots, n$  and the average queue length. Show that the behavior of the queue length is very different depending upon whether the traffic intensity  $s$  has the property  $s < 1$ ,  $s = 1$ , or  $s > 1$ .

### Exercise 11.3.22

In the queueing problem of Exercise 11.3.20 let  $S$  be the total service time required by a customer and  $T$  the time between arrivals of the customers.

- Show that  $P(S = j) = (1 - r)^{j-1}r$  and  $P(T = j) = (1 - p)^{j-1}p$ , for  $j > 0$ .
- Show that  $E(S) = 1/r$  and  $E(T) = 1/p$ .

Interpret the conditions  $s < 1$ ,  $s = 1$  and  $s > 1$  in terms of these expected values.

### Exercise 11.3.23

In Exercise 11.3.20 the service time  $S$  has a geometric distribution with  $E(S) = 1/r$ . Assume that the service time is, instead, a constant time of  $t$  seconds. Modify your computer program of Exercise [exer 11.3.20] so that it simulates a constant time service distribution. Compare the average queue length for the two types of distributions when they have the same expected service time (i.e., take  $t = 1/r$ ). Which distribution leads to the longer queues on the average?

### Exercise 11.3.24

A certain experiment is believed to be described by a two-state Markov chain with the transition matrix  $\mathbf{P}$ , where

$$\mathbf{P} = \begin{pmatrix} .5 & .5 \\ p & 1 - p \end{pmatrix} \quad (11.3.32)$$

and the parameter  $p$  is not known. When the experiment is performed many times, the chain ends in state one approximately 20 percent of the time and in state two approximately 80 percent of the time. Compute a sensible estimate for the unknown parameter  $p$  and explain how you found it.

### Exercise 11.3.25

Prove that, in an  $r$ -state ergodic chain, it is possible to go from any state to any other state in at most  $r - 1$  steps.

### Exercise 11.3.26

Let  $\mathbf{P}$  be the transition matrix of an  $r$ -state ergodic chain. Prove that, if the diagonal entries  $p_{ii}$  are positive, then the chain is regular.

### Exercise 11.3.27

Prove that if  $\mathbf{P}$  is the transition matrix of an ergodic chain, then  $(1/2)(\mathbf{I} + \mathbf{P})$  is the transition matrix of a regular chain. : Use Exercise 11.3.26

### Exercise 11.3.28

Prove that  $\mathbf{P}$  and  $(1/2)(\mathbf{I} + \mathbf{P})$  have the same fixed vectors.

### Exercise 11.3.29

In his book,<sup>15</sup> A. Engle proposes an algorithm for finding the fixed vector for an ergodic Markov chain when the transition probabilities are rational numbers. Here is his algorithm: For each state  $i$ , let  $a_i$  be the least common multiple of the denominators of the non-zero entries in the  $i$ th row. Engle describes his algorithm in terms of moving chips around on the states—indeed, for small examples, he recommends implementing the algorithm this way. Start by putting  $a_i$  chips on state  $i$  for all  $i$ . Then, at each state, redistribute the  $a_i$  chips, sending  $a_i p_{ij}$  to state  $j$ . The number of chips at state  $i$  after this redistribution need not be a multiple of  $a_i$ . For each state  $i$ , add just enough chips to bring the number of chips at state  $i$  up to a multiple of  $a_i$ . Then redistribute the

chips in the same manner. This process will eventually reach a point where the number of chips at each state, after the redistribution, is the same as before redistribution. At this point, we have found a fixed vector. Here is an example:

$$\mathbf{P}^6 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/4 & 1/4 \end{pmatrix} \end{matrix} \quad (11.3.33)$$

We start with  $\mathbf{a} = (4, 2, 4)$ . The chips after successive redistributions are shown in Table 11.3.1.

$$\begin{matrix} (4 & 2 & 4) \\ (5 & 2 & 3) \\ (8 & 2 & 4) \\ (7 & 3 & 4) \\ (8 & 4 & 4) \\ (8 & 3 & 5) \\ (8 & 4 & 8) \\ (10 & 4 & 6) \\ (12 & 4 & 8) \\ (12 & 5 & 7) \\ (12 & 6 & 8) \\ (13 & 5 & 8) \\ (16 & 6 & 8) \\ (15 & 6 & 9) \\ (16 & 6 & 12) \\ (17 & 7 & 10) \\ (20 & 8 & 12) \\ (20 & 8 & 12) . \end{matrix} \quad (11.3.34)$$

We find that  $\mathbf{a} = (20, 8, 12)$  is a fixed vector.

- Write a computer program to implement this algorithm.
- Prove that the algorithm will stop. : Let  $\mathbf{b}$  be a vector with integer components that is a fixed vector for  $\mathbf{P}$  and such that each coordinate of the starting vector  $\mathbf{a}$  is less than or equal to the corresponding component of  $\mathbf{b}$ . Show that, in the iteration, the components of the vectors are always increasing, and always less than or equal to the corresponding component of  $\mathbf{b}$ .

### Exercise 11.3.30

(Coffman, Kaduta, and Shepp<sup>16</sup>) A computing center keeps information on a tape in positions of unit length. During each time unit there is one request to occupy a unit of tape. When this arrives the first free unit is used. Also, during each second, each of the units that are occupied is vacated with probability  $p$ . Simulate this process, starting with an empty tape. Estimate the expected number of sites occupied for a given value of  $p$ . If  $p$  is small, can you choose the tape long enough so that there is a small probability that a new job will have to be turned away (i.e., that all the sites are occupied)? Form a Markov chain with states the number of sites occupied. Modify the program **FixedVector** to compute the fixed vector. Use this to check your conjecture by simulation.

### Exercise 11.3.31

(Alternate proof of Theorem [thm 11.3.8]) Let  $\mathbf{P}$  be the transition matrix of an ergodic Markov chain. Let  $\mathbf{x}$  be any column vector such that  $\mathbf{P}\mathbf{x} = \mathbf{x}$ . Let  $M$  be the maximum value of the components of  $\mathbf{x}$ . Assume that  $x_i = M$ . Show that if  $p_{ij} > 0$  then  $x_j = M$ . Use this to prove that  $\mathbf{x}$  must be a constant vector.

### Exercise 11.3.32

Let  $\mathbf{P}$  be the transition matrix of an ergodic Markov chain. Let  $\mathbf{w}$  be a fixed probability vector (i.e.,  $\mathbf{w}$  is a row vector with  $\mathbf{w}\mathbf{P} = \mathbf{w}$ ). Show that if  $w_i = 0$  and  $p_{ji} > 0$  then  $w_j = 0$ . Use this to show that the fixed probability vector for an ergodic chain cannot have any 0 entries.

### Exercise 11.3.33

Find a Markov chain that is neither absorbing or ergodic.

---

This page titled [11.3: Ergodic Markov Chains\\*\\*](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.