#linear_algebra

The Markov chains described in this section are used as mathematical models of a wide variety of situations in biology, business, chemistry, engineering, physics, and elsewhere. In each case, the model is used to describe an experiment or measurement that is performed many times in the same way, where the outcome of each trial of the experiment will be one of several specified possible outcomes, and where the outcome of one trial depends only on the immediately preceding trial.

For example, if the population of a city and its suburbs were measured each year, then a vector such as

$$\mathbf{x}_0 = \begin{bmatrix} .60 \\ .40 \end{bmatrix} \tag{1}$$

could indicate that 60% of the population lives in the city and 40% in the suburbs. The decimals in $\mathbf{x}_0$ add up to 1 because they account for the entire population of the region. Percentages are more convenient for our purposes here than population totals.

A vector with nonnegative entries that add up to 1 is called a **probability vector**. A **stochastic matrix** is a square matrix whose columns are probability vectors. A **Markov chain** is a sequence of probability vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$, together with a stochastic matrix P, such that

$$\mathbf{x}_1 = P\mathbf{x}_0, \quad \mathbf{x}_2 = P\mathbf{x}_1, \quad \mathbf{x}_3 = P\mathbf{x}_2, \quad \ldots$$

Thus the Markov chain is described by the first-order difference equation

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \ldots$$

When a Markov chain of vectors in $\mathbb{R}^n$ describes a system or a sequence of experiments, the entries in $\mathbf{x}_k$ list, respectively, the probabilities that the system is in each of n possible states, or the probabilities that the outcome of the experiment is one of n possible outcomes. For this reason, $\mathbf{x}_k$ is often called a **state vector**.

EXAMPLE 1 Section 1.10 examined a model for population movement between a city and its suburbs. See Fig. 1. The annual migration between these two parts of the metropolitan region was governed by the *migration matrix* M:

$$\begin{array}{ccc} & \text{From:} & \\ & \begin{array}{cc}\text{City} & \text{Suburbs}\end{array} & \text{To:} \\ M = & \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} & \begin{array}{l}\text{City}\\ \text{Suburbs}\end{array} \end{array}$$

That is, each year 5% of the city population moves to the suburbs, and 3% of the suburban population moves to the city. The columns of M are probability vectors, so M is a stochastic matrix. Suppose the 2000 population of the region is 600,000 in the city and 400,000 in the suburbs. Then the initial distribution of the population in the region is given by $\mathbf{x}_0$ in (1) above. What is the distribution of the population in 2001? In 2002?

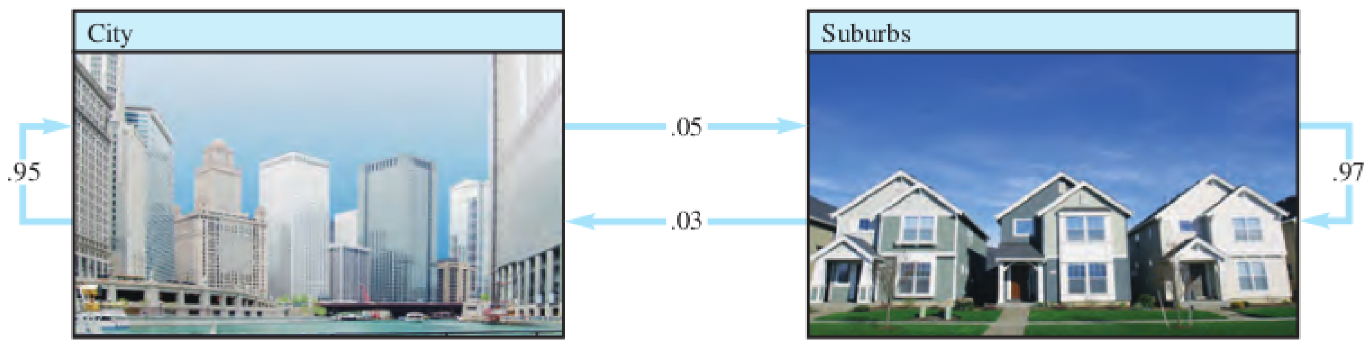

FIGURE 1 Annual percentage migration between city and suburbs.

SOLUTION In Example 3 of Section 1.10, we saw that after one year, the population vector $\begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$ changed to

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

If we divide both sides of this equation by the total population of 1 million, and use the fact that $kM\mathbf{x} = M(k\mathbf{x})$, we find that

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .600 \\ .400 \end{bmatrix} = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$$

The vector $\mathbf{x}_1 = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$ gives the population distribution in 2001. That is, 58.2% of the region lived in the city and 41.8% lived in the suburbs. Similarly, the population distribution in 2002 is described by a vector $\mathbf{x}_2$, where

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .582 \\ .418 \end{bmatrix} = \begin{bmatrix} .565 \\ .435 \end{bmatrix} \qquad \blacksquare$$

EXAMPLE 2 Suppose the voting results of a congressional election at a certain voting precinct are represented by a vector $\mathbf{x}$ in $\mathbb{R}^3$:

$$\mathbf{x} = \begin{bmatrix} \text{\% voting Democratic (D)} \\ \text{\% voting Republican (R)} \\ \text{\% voting Libertarian (L)} \end{bmatrix}$$

Suppose we record the outcome of the congressional election every two years by a vector of this type and the outcome of one election depends only on the results of the preceding election. Then the sequence of vectors that describe the votes every two years may be a Markov chain. As an example of a stochastic matrix P for this chain, we take

$$\begin{array}{cc} & \text{From:} \\ & \begin{array}{ccc} \text{D} & \text{R} & \text{L} \end{array} \\ P = & \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \end{array} \begin{array}{c} \text{To:} \\ \text{D} \\ \text{R} \\ \text{L} \end{array}$$

The entries in the first column, labeled D, describe what the persons voting Democratic in one election will do in the next election. Here we have supposed that 70% will vote D again in the next election, 20% will vote R, and 10% will vote L. Similar interpretations hold for the other columns of P. A diagram for this matrix is shown in Fig. 2.

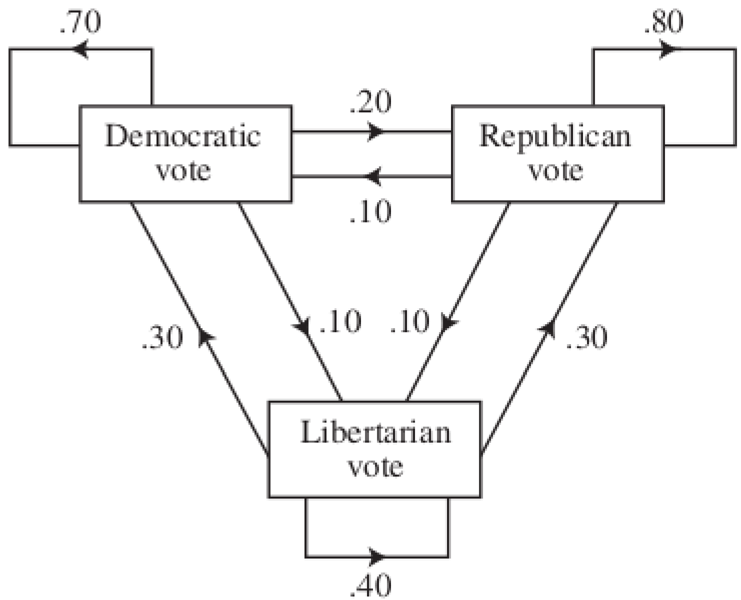

FIGURE 2 Voting changes from one election to the next.

If the "transition" percentages remain constant over many years from one election to the next, then the sequence of vectors that give the voting outcomes forms a Markov chain. Suppose the outcome of one election is given by

$$\mathbf{x}_0 = \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix}$$

Determine the likely outcome of the next election and the likely outcome of the election after that.

SOLUTION The outcome of the next election is described by the state vector $\mathbf{x}_1$ and that of the election after that by $\mathbf{x}_2$, where

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix} = \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} \quad \begin{array}{l} \text{44\% will vote D.} \\ \text{44.5\% will vote R.} \\ \text{11.5\% will vote L.} \end{array}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} = \begin{bmatrix} .3870 \\ .4785 \\ .1345 \end{bmatrix} \quad \begin{array}{l} \text{38.7\% will vote D.} \\ \text{47.8\% will vote R.} \\ \text{13.5\% will vote L.} \end{array}$$

To understand why $\mathbf{x}_1$ does indeed give the outcome of the next election, suppose 1000 persons voted in the "first" election, with 550 voting D, 400 voting R, and 50 voting L. (See the percentages in $\mathbf{x}_0$.) In the next election, 70% of the 550 will vote D again, 10% of the 400 will switch from R to D, and 30% of the 50 will switch from L to D. Thus the total D vote will be

$$.70(550) + .10(400) + .30(50) = 385 + 40 + 15 = 440 \tag{2}$$

Thus 44% of the vote next time will be for the D candidate. The calculation in (2) is essentially the same as that used to compute the first entry in $\mathbf{x}_1$. Analogous calculations could be made for the other entries in $\mathbf{x}_1$, for the entries in $\mathbf{x}_2$, and so on. ■

Predicting the Distant Future.

The most interesting aspect of markov chian is the study of a chain's long-term behavior.

$$\mathbf{x}_8 = \begin{bmatrix} .3008 \\ .5977 \\ .1016 \end{bmatrix}, \quad \mathbf{x}_9 = \begin{bmatrix} .3004 \\ .5988 \\ .1008 \end{bmatrix}, \quad \mathbf{x}_{10} = \begin{bmatrix} .3002 \\ .5994 \\ .1004 \end{bmatrix}, \quad \mathbf{x}_{11} = \begin{bmatrix} .3001 \\ .5997 \\ .1002 \end{bmatrix}$$

$$\mathbf{x}_{12} = \begin{bmatrix} .30005 \\ .59985 \\ .10010 \end{bmatrix}, \quad \mathbf{x}_{13} = \begin{bmatrix} .30002 \\ .59993 \\ .10005 \end{bmatrix}, \quad \mathbf{x}_{14} = \begin{bmatrix} .30001 \\ .59996 \\ .10002 \end{bmatrix}, \quad \mathbf{x}_{15} = \begin{bmatrix} .30001 \\ .59998 \\ .10001 \end{bmatrix}$$

These vectors seem to be approaching $\mathbf{q} = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$. The probabilities are hardly changing from one value of k to the next. Observe that the following calculation is exact (with no rounding error):

$$P\mathbf{q} = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix} = \begin{bmatrix} .15 + .12 + .03 \\ .09 + .48 + .03 \\ .06 + 0 + .04 \end{bmatrix} = \begin{bmatrix} .30 \\ .60 \\ .10 \end{bmatrix} = \mathbf{q}$$

When the system is in state $\mathbf{q}$, there is no change in the system from one measurement to the next. ■

Steady-state vectors

If P is a stochastic matrix, then a **steady-state vector** (or **equilibrium vector**) for P is a probability vector $\mathbf{q}$ such that

$$P\mathbf{q} = \mathbf{q}$$

It can be shown that every stochastic matrix has a steady-state vector. In Example 3, $\mathbf{q}$ is a steady-state vector for P.

EXAMPLE 4 The probability vector $\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$ is a steady-state vector for the population migration matrix M in Example 1, because

$$M\mathbf{q} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .375 \\ .625 \end{bmatrix} = \begin{bmatrix} .35625 + .01875 \\ .01875 + .60625 \end{bmatrix} = \begin{bmatrix} .375 \\ .625 \end{bmatrix} = \mathbf{q}$$

If the total population of the metropolitan region in Example 1 is 1 million, then $\mathbf{q}$ from Example 4 would correspond to having 375,000 persons in the city and 625,000 in the suburbs. At the end of one year, the migration *out of* the city would be $(.05)(375{,}000) = 18{,}750$ persons, and the migration *into* the city from the suburbs would be $(.03)(625{,}000) = 18{,}750$ persons. As a result, the population in the city would remain the same. Similarly, the suburban population would be stable.

How to find a steady-state vector.

EXAMPLE 5 Let $P = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$. Find a steady-state vector for P.

SOLUTION First, solve the equation $P\mathbf{x} = \mathbf{x}$.

$$\begin{aligned} P\mathbf{x} - \mathbf{x} &= \mathbf{0} \\ P\mathbf{x} - I\mathbf{x} &= \mathbf{0} \qquad \text{Recall from Section 1.4 that } I\mathbf{x} = \mathbf{x}. \\ (P - I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

For P as above,

$$P - I = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -.4 & .3 \\ .4 & -.3 \end{bmatrix}$$

To find all solutions of $(P - I)\mathbf{x} = \mathbf{0}$, row reduce the augmented matrix:

$$\begin{bmatrix} -.4 & .3 & 0 \\ .4 & -.3 & 0 \end{bmatrix} \sim \begin{bmatrix} -.4 & .3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then $x_1 = \frac{3}{4}x_2$ and x_2 is free. The general solution is $x_2 \begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$.

Next, choose a simple basis for the solution space. One obvious choice is $\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$ but a better choice with no fractions is $\mathbf{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ (corresponding to $x_2 = 4$).

Finally, find a probability vector in the set of all solutions of $P\mathbf{x} = \mathbf{x}$. This process is easy, since every solution is a multiple of the solution $\mathbf{w}$ above. Divide $\mathbf{w}$ by the sum of its entries and obtain

$$\mathbf{q} = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$$

As a check, compute

$$P\mathbf{q} = \begin{bmatrix} 6/10 & 3/10 \\ 4/10 & 7/10 \end{bmatrix} \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix} = \begin{bmatrix} 18/70 + 12/70 \\ 12/70 + 28/70 \end{bmatrix} = \begin{bmatrix} 30/70 \\ 40/70 \end{bmatrix} = \mathbf{q} \qquad \blacksquare$$

THEOREM 18

If P is an $n \times n$ regular stochastic matrix, then P has a unique steady-state vector $\mathbf{q}$. Further, if $\mathbf{x}_0$ is any initial state and $\mathbf{x}_{k+1} = P\mathbf{x}_k$ for $k = 0, 1, 2, \ldots$, then the Markov chain $\{\mathbf{x}_k\}$ converges to $\mathbf{q}$ as $k \to \infty$.