MATH 2320: Linear Algebra
Prof. Leanne Robertson
May 20, 2024

Lab 3: Markov Chains

Due date: Lab 3 is due Monday, June 3, 2024. The lab is worth 20 points.

Overview: Markov chains are used model systems that change from state to state when the current state depends only on the previous state. Here, you will consider the movement of a population and the change in weather, and use linear algebra is used to predict future conditions. Along the way, you will be introduced to the concept of a *steady-state vector*, which is an example an eigenvector (with eigenvalue 1), which is an important idea that we will study in Chapter 5.

Collaborative report: You will work in groups of 3 students on this lab. When working in a group, it is critical that you DO NOT simply divide up the work. You must discuss the problems, compare answers, and make sure that everyone shares equitably in solving the problems and writing and proofreading the report. Each group will hand in a single typed lab report.

What to hand in: A completed lab report consists of typed answers presented in *complete sentences*. Keep the numbering of the questions in your report and answer each in a paragraph, using complete sentences to clearly explain your reasoning. Be sure to state your assumptions and define all your variables. Do not include MATLAB input. Instead, you should edit your your MATLAB file, removing all of the MATLAB commands and typing errors, and present your answers and explanations in complete sentences. MATLAB is a tool for computation, so only the results of your MATLAB computations should be included in your report and they should be incorporated into the text using complete sentences.

Your lab report will be graded for presentation and mathematical correctness. You and your lab partners should all check the computations so that you will not lose points. Also make sure that your answers are clear and your work is explained using complete sentences. Please ask if you have any questions about these expectations.

Getting started: A with your previous labs, you will need to record the results of your MATLAB session in a document, which you can edit using your favorite text editor to produce your lab report. I think the easiest way to record your results is to use the **diary** command in MATLAB. To do this, begin your MATLAB session by typing the command **diary lab3.txt** followed by the **Enter** key. Then each subsequent computation you make in MATLAB will be saved in a text file named **lab2.txt**. I suggest that the first command you make is to type **format compact** so you will not have unnecessary spaces your diary file. When you finish your MATLAB session, please don't forget to turn off the recording by typing **diary off** at the MATLAB prompt, otherwise you may lose all your work. If you stop your MATLAB session before competing all the computations needed for your lab assignment, then you can reopen the diary file the next time you start MATLAB. If you use the same filename, the results of your new MATLAB session will be written at the end of the old diary file.

Lab 3: Markov Chains

(Problems 1 and 2 are based on a lab by Dr. MacLean)

Problem 1. There are no rural areas in the Land of Oz, only suburbs and cities. Each year 5% of the city population moves to the suburbs, and 95% stay in the city. Also, each year 3% of the suburban population moves to the city, while the other 97% stay in the suburbs.

Suppose in the year 2020, there were 6000 people living in the city and 4000 people living in the suburbs. For each nonnegative integer k, let the vector x_k be the vector in $\mathbb{R}^2$ where the first coordinate is the number of people living in the city and the second coordinate is the number of people living in the suburbs (where k is the number of years after the year 2020).

So $x_0 = (6000, 4000)$. In the year 2021, based on the above information, there were

$$(.95)(6000) + (.03)(4000) = 5820$$

people living in the city, and

$$(.05)(6000) + (.97)(4000) = 4180$$

people living in the suburbs. Thus, $x_1 = (5820, 4180)$. We can use linear algebra to predict the size of the future populations in the suburbs and cities in the Land of Oz.

Let A denote the following *transition matrix* (a square matrix with nonnegative entries such that the sum of the entries of each column is equal to 1):

$$A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}.$$

Then,

$$Ax_0 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 6000 \\ 4000 \end{bmatrix} = \begin{bmatrix} (.95)(6000) + (.03)(4000) \\ (.05)(6000) + (.97)(4000) \end{bmatrix} = \begin{bmatrix} 5820 \\ 4180 \end{bmatrix} = x_1.$$

(a) What is x_2? (Remember, each year 5% of the city population moves to the suburbs, and 95% stay in the city. And each year 3% of the suburban population moves to the city, while the other 97% stay in the suburbs.)

(b) Can you give a formula for x_2 in terms of the matrix A and x_0? How about for x_3? Explain your reasoning.

(c) Approximately how many people will be living in the city in the year 2025? How many people will be living in the suburbs in the year 2025? Show work and explain your reasoning.

(d) Suppose in a given year there are 3750 people living in the city and 6250 people living in the suburbs. How many people will be living in the city and how many in the suburbs the following year? Two years later?

(e) Lets be the vector $s = \begin{bmatrix} 3750 \\ 6250 \end{bmatrix}$. Why do you think s is called a *steady-state vector*? (If you have no idea, you may have done something wrong in your calculations.)

Problem 2. Here is a method for finding a steady-state vector s as in Problem 1(e). If A is a square matrix and v is a nonzero vector such that $Av = v$, then v is a steady-state vector for A. It is also called an *eigenvector* of A with *eigenvalue* 1 (more on this important topic in Sections 5.1 and 5.2 of your textbook, which we will cover at the end of the course). If we wish to find v, we can solve the equation $Av = v$ by rewriting it as $Av - v = 0$, and then as $(A - I)v = 0$. Now we simply need to solve the homogeneous system $(A - I)v = 0$. (You know how to do this! Don't be thrown off by the fact that our matrix is now $A - I$ instead of just A.) Use this idea to find a vector v such that $Av = v$, where A is the matrix

$$A = \begin{bmatrix} 0 & 1/3 & 1/3 \\ 0 & 2/3 & 1/3 \\ 1 & 0 & 1/3 \end{bmatrix}.$$

Note that there will be an infinite number of solutions. Pick one and call it s. Verify that $As = s$. Be sure to show this in your lab report. Can you find a different steady-state vector for A that is also a *probability vector*, that is a vector whose entries are nonnegative and add up to 1? Either find such a stead-state vector or explain why none exists. In applications you may need to pick a state-state vector that makes the most sense within the context of the problem, and often probability vectors are useful (as in the next problem).

Problem 3. Next we will investigate the weather in the Land of Oz. Here there are only three kids of weather – sunny, cloudy, and rainy – and it is never sunny two days in a row. We cannot determine exactly what the weather will be on a given day, but we can give the probability of sun, clouds, or rain on a given day because the weather follows the following behavior:

- After a sunny day, the next day is equally likely to be cloudy or rainy.
- After a cloudy day, there is a 1/4 probability that the next day will be sunny, 1/4 probability that it will be cloudy, and 1/2 probability that it will be rainy.
- After a rainy day, there is a 1/4 probability that the next day will be sunny, 1/2 probability that it will be cloudy, and 1/4 probability that it will be rainy.

Based on this information, the weather in the Land of Oz can be modeled by a Markov chain with transition matrix

$$P = \begin{bmatrix} 0 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/2 \\ 1/2 & 1/2 & 1/4 \end{bmatrix},$$

where the columns and rows give the probabilities from sunny, cloudy, and rainy as stipulated above. Suppose that on day 0 the weather is rainy. That is,

$$x_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then the probabilities for the weather the next day are given by

$$x_1 = Px_0 = \begin{bmatrix} 0 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/2 \\ 1/2 & 1/2 & 1/4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}.$$

That is, there is a 1/4 probability that the next day will be sunny, 1/2 probability that it will be cloudy, and 1/4 probability that it will be rainy, exactly as stipulated above. Note that the vectors x_0 and x_1 are examples of probability vectors, as defined in Problem 2. In this problem, all the vectors that you deal with should be probability vectors.

(a) Find the probabilities for sun, clouds, and rain for the next day. Find the probabilities for the weather a week after the initial rainy day. Find the probabilities for the weather two weeks after the initial rainy day. Be sure to show your work and explain how you got your answers.

(b) Use the method in Problem 2 to compute a steady-state vector for P. As in Problem 2, there will be infinitely many choices for your vector. Choose one that is a probability vector, *i.e*, so that all the entries add up to 1. How does this steady-state vector compare with your answers in part (a) above? What does the steady-state vector say about the weather in the Land of Oz in the long run?