Markov Chains

Finding Probability with Matrix Algebra

Kailey Cozart, Sukhraj Kaur

June 7, 2019

1 Abstract

Markov chains are sequences and probabilities of events occurring, where each probability is based off of the state obtained from the previous event. Matrix algebra can be helpful in solving problems involving Markov chains because equations comprised of states and probabilities can be solved simultaneously in matrices. Because Markov chains can be used to predict various probabilities, they are applicable in countless scenarios. They can be used to solve problems such as page ranking, population processes, and algorithmic music composition to varying degrees of success. In this paper, we will discuss basic computations and applications of Markov chains.

2 Introduction

Markov chains, as defined in statistics, are "a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event" (Oxford 1). In other words, they are a system that is used to represent changes between states based on certain rules

(Maltby 1). They were named after Andrey Markov, a Russian mathematician who lived from the mid-1800's to the early 1900's (Gagniuc 1).

2.1 Applications

Markov chains can be used in anything from economics to genetics. In our paper, we will use Markov chains to solve market share and business applications. Some other ways that Markov chains can be used are discussed below.

Google's PageRank algorithm uses Markov chains and can be seen as a model of user behavior where an individual is given websites and randomly clicks on links while never clicking the "back" button (Jauregui 79). In other words, it is the probability of opening a certain page after the previous events have brought the user to a certain state. Using this process, PageRank is able to predict what websites users might be likely to click on.

Algorithmic music composition can also include Markov chains. Notes become states in the system, and every note corresponds to a value. A transition matrix is then created with values derived from a certain musical style. Both higher-order chains and first-order chains can be used to create music, but typically only higher-order chains produce musical phrases that sound pleasing (Mentes 1).

The applications of Markov chains are almost endless, as they can be used to solve various problems involving probability, as long as certain requirement are fulfilled.

2.2 Limitations

A few conditions must be satisfied in order for Markov chains to be used. First of all, the total population that the Markov process describes must be constant. Secondly, the population of a state must never be negative (Fraleigh

105). Markov chains must have a finite number of states, and the state space is the set of possible states (Maltby 1). They must also fulfill the "Markov property," which says that the probability of future events is not related to the events that have occurred prior to the current state (Gagniuc 1).

3 The Basics of Markov Chains

3.1 Transition Probability Matrix

In Markov processes, two types of matrices are used. The first is a transition probability matrix and the second is the initial state distribution matrix. The transition probability matrix describes the probability of a state staying the same or changing. The rows of a transition probability matrix must add up to 1 because the total probability must be 1. The columns of a transition matrix do not need to add up to any particular number.

The 3x3 matrix P is an initial state distribution matrix and shows the probabilities between 3 different states. P_{ij} is the probability of moving from state j to state i.

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \end{array}$$

In some cases, P_{ij} can be defined differently. We will discuss those cases in section 3.3.

3.2 Initial State Distribution Matrix

The initial state distribution matrix is used to calculate what the probabilities will look like over a particular time period. The vector $\mathbf{x_0}$ is an initial state distribution matrix. The sum of all the components of P and $\mathbf{x_0}$ must add up to 1:

$$\mathbf{x_0} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

3.3 Computational Methods

There are two way to compute Markov chains. The first method is to multiply the initial state matrix on the right of the transition matrix as a column vector. In this case, the transition matrix P_{ij} is the probability of moving from state j to state i. For example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The second method is to multiply the initial state matrix on the left of the transition matrix as a row. In this case, the transition matrix P_{ij} is the probability of moving from state i to state j. For example:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

In our paper, we will be using the first method to compute Markov chains, but both methods are widely used (Biezen 1).

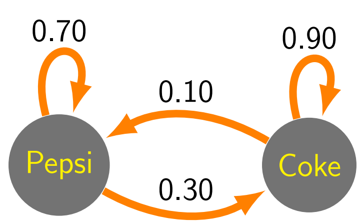
4 Examples

4.1 Sources

Examples 5.1 and 5.2 are based off of an example by Henry Maltby. See works cited for more information. The main application, Example 5.3, was created by us.

4.2 Transition Diagram Example

Suppose that we have two companies, Pepsi and Coke. If, at the initial state, you are at Pepsi, then there is a 70% chance that it will be Pepsi at the final state and a 30% chance that it will be Coke at the final state. Similarly, if you are at Coke at the initial state, there is a 90% chance that it will be Coke at the final state and a 10% chance that it will be Pepsi at the final state. With this information, we can create a transition diagram to model the information. Transition diagrams visually illustrate Markov processes:



As we can see above, the probabilities for possible events for Coke add up to 1, and the probabilities for possible events for Pepsi also add up to 1.

Problem: Given this information, create a transition diagram.

Solution: If we are using a column vector for our initial state matrix and will be multiplying the initial state matrix on the right of the transition matrix, we will have the initial state on the top of the matrix and the final state on the left of the matrix. In other words, M_{ij} is the probability of moving from state j to state i. Using this information and the chart above, we can create the transition probability matrix M below:

$$M = \begin{array}{c} \\ Pepsi \\ Coke \end{array} \begin{array}{c} \begin{array}{cc} Pepsi & Coke \end{array} \\ \begin{pmatrix} 0.70 & 0.10 \\ 0.30 & 0.90 \end{pmatrix} \end{array}$$

The matrix M shows the probability of moving from a state on the top to a state on the left. For example, if the initial state is Coke, then there is a 10% chance that the final state will be Pepsi (Maltby 1).

4.3 Market Share Prediction Example

Let's suppose that a market survey found that 55% of the market share belongs to Pepsi and 45% of the market share belongs to Coke. We can display this information in an initial state distribution matrix:

$$\mathbf{x_0} = \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix}$$

Problem: Given this information, as well as the information from Example 5.1, we want to find the percentage of the market share that Pepsi will have in 2 months, assuming that each time period is 1 month.

Solution: In order to find the market share that Pepsi will have in 2 months, we will first find the percentage of the market share that Pepsi will have in 1 month. To do this, we multiply the transition matrix and the initial state (current state) matrix:

$$\mathbf{x_1} = M\mathbf{x_0} \qquad \mathbf{x_1} = \begin{bmatrix} 0.70 & 0.10 \\ 0.30 & 0.90 \end{bmatrix} \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix} = \begin{bmatrix} 0.43 \\ 0.57 \end{bmatrix}$$

As we can see above, after 1 month, Pepsi has 43% of the market share. Now, if we want to know what Pepsi's market share will be in 2 months, we multiply the same transition matrix with the result for the first month, like this:

$$\mathbf{x_2} = M\mathbf{x_1} \qquad \mathbf{x_2} = \begin{bmatrix} 0.70 & 0.10 \\ 0.30 & 0.90 \end{bmatrix} \begin{bmatrix} 0.43 \\ 0.57 \end{bmatrix} = \begin{bmatrix} 0.36 \\ 0.64 \end{bmatrix}$$

Above, we see that, after 2 months, Pepsi has only 36% of the market share as opposed to the 55% of the market share that they had at the beginning of the 2 months (Maltby 1).

4.4 Steady State Market Share Example

Matrix S represents the matrix with target market share values between Microsoft, Amazon, and Google. S_{11} shows us that 50% of Microsoft's products are bought by customers brand-loyal to Microsoft.

$$S = \begin{array}{c} \\ M \\ A \\ G \end{array} \begin{array}{c} \begin{array}{ccc} M & A & G \end{array} \\ \begin{pmatrix} 0.50 & 0.30 & 0.20 \\ 0.20 & 0.50 & 0.30 \\ 0.30 & 0.20 & 0.50 \end{pmatrix} \end{array}$$

If Microsoft were to launch a new phone, the initial state of it's profitability can be represented by

$$\mathbf{x_0} = \begin{bmatrix} 0.60 \\ 0.20 \\ 0.20 \end{bmatrix}$$

. Assume this matrix is a prediction that is meant to hold for one week, and that the percentages were taken from a market survey.

If we want to know what the prediction matrix will be for who will buy the new phone in week 2, we must multiply S by $\mathbf{x_0}$ to get our prediction:

$$\mathbf{x_1} = S\mathbf{x_0} \qquad \mathbf{x_1} = \begin{bmatrix} 0.40 \\ 0.28 \\ 0.32 \end{bmatrix}$$

We can continue to repeat this process to get new predictions for each week till we reach a *steady state* i.e. when the probabilities reach a state that is constant.

$$\mathbf{x_2} = S\mathbf{x_1} \qquad \mathbf{x_2} = \begin{bmatrix} 0.35 \\ 0.32 \\ 0.33 \end{bmatrix}$$

At $\mathbf{x_2}$ the probabilities are still changing, so we find $\mathbf{x_3}$.

$$\mathbf{x_3} = S\mathbf{x_2} \qquad \mathbf{x_3} = \begin{bmatrix} 0.34 \\ 0.33 \\ 0.34 \end{bmatrix}$$

When we see that $\mathbf{x_3}$ and $\mathbf{x_4}$ are the same, we begin to suspect that we have reached the steady state.

$$\mathbf{x_4} = S\mathbf{x_3} \qquad \mathbf{x_4} = \begin{bmatrix} 0.34 \\ 0.33 \\ 0.34 \end{bmatrix}$$

We reach a steady state at Week 3. The probabilities will not change further. It will remain constant for week 4, week 5, and so on. (The steady state vector that we are left with is an Eigenvector.) This idea of a constant state is important in Markov chains because, at some point, many Markov chains no longer change. It is important to note that Markov processes do not always end in a steady state.

5 Limitations Revisited and Conclusion

Markov matrices can help us predict the probability of a future event. But there are limitations to this model. First, future probability is based off the previous events and the model assumes that probabilities remain constant throughout the computation. Second, with this model, it's impossible to compute every single factor that can effect an event. For example, in our market share problem, we are only looking at three companies, but, in real world, there are many other competitors and factors that can effect the outcomes of a future event. For instance, a sudden downfall in our economy can change the results. A natural disaster, launch of another product by a competitor, etc. can affect our numbers significantly, but, as discussed, our model doesn't account for these things. Hence, Markov matrices are tools to study the likelihood of an event, and they are not always completely accurate. Despite these limitations, Markov chains are an interesting concept in Linear Algebra, and there are still countless uses for Markov chains, many of which we have discussed in our paper.

6 Works Cited

Biezen, Michel. "Prob and Stats - Markov Chains: Method 2 (30 of 38) Basics." (Jul 2015). Web. 7 Jun 2019.

Fraleigh, John B. and Raymond A. Beauregrad. *Linear Algebra, Third Edition.* Addison-Wesley Publishing Company, Inc. (1995).

Gagniuc, Paul A. *Markov Chains: From Theory to Implementation and Experimentation.* USA, NJ: John Wiley & Sons. (2017): 1–235.

Jauregui, Jeff. "Math 312: Markov Chains, Google's PageRank Algorithm." (Oct 2012). Web. 21 May 2019.

Maltby, Henry. "Markov Chains." *Brilliant.* Web. 4 June 2019.

"Markov chain." Oxford Living Dictionary - English. Web. 4 June 2019.

Mentes, Selen. "Algorithmic Music Composition with Markov Chains." (May 2015) Web. 21 May 2019.

Srivastava, Tavish. "Introduction to Markov Chains: Simplified!" *Analytics Vidhya.* (July 2014). Web. 4 June 2019.

Kiser, Terry L., Thomas A. McCready, and Neil C. Schwertman. "Can the Committee Meet? A Markov Chain Analysis." *Mathematics Magazine* 78. (Feb 2005): 57-63.

Knill, Oliver. "Lecture 33: Markov Matrices." Web. 2 June 2019.