UMUC Math 340

Markov Chains and Weather Prediction

An Application of Linear Algebra

Arun Kulkarni
5-10-2019

I. Introduction

Markov chains are named after Russian mathematician Andrey Andreyevich Markov (June 14, 1856 – July 20, 1922). The sum of his research areas have become known as Markov processes and Markov chains, which are widely used today for several applications, a lot of which are various forms of statistical modeling. Markov first introduced the Markov chains in 1906 when produced the first theoretical results for stochastic processes (randomly determined processes that can be modeled but not predicted with 100% precision) and used the term "chain" for the first time, and in 1913 he calculated letter sequences of the Russian language. One defining characteristic of Markov processes and chains that sets them apart from other stochastic models is their memorylessness, meaning that the condition of the future state $n+1$ is only dependent upon the condition of the current state n, and not any of the past states. Effectively, this means that the model's future and past are independent of one another.

This provides a way of dealing with a sequence of events based on the probabilities dictating the motion of a set of data among various states, even is little is known about the conditions of past events. Applications of this type of modeling range from minimizing energy consumption in a mobile wireless service network to predicting election results.

II. Background

A Markov chain can be described as follows: we have a set of states, $S = \{s_1, s_2, \ldots , s_r\}$. The Markov process starts in one of these states and moves successively from one state to another, each move here being called a step. If the chain is presently in state s_i, then it moves to state s_j at the next step with a probability denoted by p_{ij}, the transition probability. As previously stated due to the memorylessness property of Markov chains, this probability does not depend upon the states the chain was in before the current state. An initial probability distribution defined on the set *S* specifies the starting state, which is usually a particular state in *S*.

To demonstrate the use of Markov chains, let us consider the following example, taken from Jim Hefferon's text "Linear Algebra." A simple game exists in which one player bets money on the results of a coin toss. The player starts with one dollar, either loses or gains a dollar based on whether his or her prediction is correct, and the game ends when the player either reaches five dollars or is left with zero dollars. At any point, the player has $0, $1, … , $5, which correspond to the state that the player is in of s_0, s_1, … , s_5. The transition probability from any state s_i to the next state s_j is 0.5, as a coin toss has a 0.5 probability of being heads or tails. The starting state is s_1 or $1, and the boundary states are defined as s_0 and s_5, as these states are when the game ends. Let the transition probability of i $p_i(n)$ represent the probability that the player is in state s_i after n flips; then for example the probability of being in state s_2 after flip n+1 is $p_2(n+1) = p_2(n) + 0.5 * p_3(n)$. Using these rules, we can come up with the following linear system of probabilities:

$$\begin{pmatrix} 1.0 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.5 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.0 & 0.5 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 1.0 \end{pmatrix} \begin{pmatrix} p_0(n) \\ p_1(n) \\ p_2(n) \\ p_3(n) \\ p_4(n) \\ p_5(n) \end{pmatrix} = \begin{pmatrix} p_0(n+1) \\ p_1(n+1) \\ p_2(n+1) \\ p_3(n+1) \\ p_4(n+1) \\ p_5(n+1) \end{pmatrix}$$

(Hefferon p.305)

Solving this system yields the Markov chain:

$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$\cdots$	$n = 24$
0	0	0	0.125	0.125		0.39600
0	0	0.25	0	0.1875		0.00276
0	0.5	0	0.375	0		0
1	0	0.5	0	0.3125		0.00447
0	0.5	0	0.25	0		0
0	0	0.25	0.25	0.375		0.59676

(Hefferon p.306)

In this Markov chain, each vector is a probability vector whose components are all nonnegative real values that sum to 1. The matrix is the transition matrix or stochastic matrix, whose entries are nonnegative real values and whose columns sum to 1. The computation of the Markov chain is as follows: the vector for n=0 is computed, and then multiplied by the transition matrix to yield the vector for n=1, which is then multiplied by the transition matrix to yield the vector for n=2, and so on. This process can be repeated infinitely.

III. Solution of Basic Problem

Mankind has been predicting the weather for centuries. The earliest recorded weather forecasting occurred with early civilizations using reoccurring astronomical and meteorological events to help them monitor seasonal changes in the weather. Starting with the Babylonians and the Chinese, the art of weather prediction evolved over time from an art to a science, with numerical prediction and satellite-assisted observation-based forecasting being the main methods used today.

Markov chains can be used to model weather predictions of any variety, from chances of precipitation to tornado likelihood in a given area. One such example comes from J.T. Schoof and S. C. Pryor in their paper “On the Proper Order of Markov Chain Model for Daily Precipitation Occurrence in the Contiguous United States.”

These precipitation models are described using the aforementioned states, as well as the number of previous values used to determine the state-to-state transition probabilities or the order. In the case of these precipitation models, the variable X_t representing the current state is binary meaning that it only has two states; occurrence of precipitation denoted by $X_t = 1$, or nonoccurrence of precipitation denoted by $X_t = 0$. From this information, the simplest Markov model for precipitation occurrence is a two-state, first-order model, which can be described using the following four transition probabilities:

$$\begin{aligned}
p_{00} &= \Pr\{X_t = 0 | X_{t-1} = 0\}, \\
p_{01} &= \Pr\{X_t = 1 | X_{t-1} = 0\}, \\
p_{10} &= \Pr\{X_t = 0 | X_{t-1} = 1\}, \quad \text{and} \\
p_{11} &= \Pr\{X_t = 1 | X_{t-1} = 1\}.
\end{aligned}$$

(Schoof and Pryor, p. 2478)

Since the sum of P_{00} and P_{01} equals 1, as well as the sum of P_{10} and P_{11}, the model can be simplified and defined fully by two transition probabilities: P_{01}, the probability that it will rain tomorrow if it did not rain today, and P_{11}, the probability that that it will rain tomorrow if it rained today. To gather and compute the data for the Markov chain model, 831 different weather stations were monitored throughout the contiguous United States and used to determine whether or not a Markov chain would be appropriate to model the probability of rain occurrence, and which model-order would fit best. The resulting data is as follows:

Month	Zeroth order		First order		Second order		Third order	
	BIC	K-S	BIC	K-S	BIC	K-S	BIC	K-S
Jan	0	52 (73.7)	645 (91.4)	487 (84.8)	158 (74.4)	209 (86.6)	28 (82.1)	4 (61.5)
Feb	0	47 (72.3)	703 (93.0)	531 (87.3)	108 (75.7)	192 (87.2)	20 (76.5)	12 (67.2)
Mar	0	25 (67.5)	680 (90.1)	561 (86.5)	132 (74.1)	195 (88.3)	19 (69.8)	7 (63.6)
Apr	0	15 (76.4)	725 (89.0)	610 (85.7)	104 (74.7)	144 (90.6)	2 (65.0)	18 (59.1)
May	0	26 (83.5)	748 (93.1)	601 (85.1)	81 (77.0)	179 (88.0)	2 (56.5)	7 (61.9)
Jun	0	41 (80.8)	654 (87.0)	525 (82.4)	171 (75.2)	215 (86.1)	6 (71.5)	10 (59.9)
Jul	2 (75.5)	105 (79.0)	659 (89.7)	468 (85.0)	128 (75.9)	172 (85.5)	42 (81.4)	28 (63.0)
Aug	0	104 (77.8)	689 (91.5)	519 (83.6)	132 (75.0)	142 (84.9)	10 (78.8)	12 (60.9)
Sep	0	12 (71.6)	711 (88.6)	633 (86.7)	118 (71.6)	156 (84.0)	2 (63.0)	16 (61.7)
Oct	0	4 (70.5)	738 (91.6)	697 (89.6)	92 (73.5)	91 (87.7)	1 (38.0)	2 (52.0)
Nov	0	12 (70.5)	698 (90.4)	586 (87.6)	96 (74.7)	136 (86.5)	37 (84.8)	15 (62.5)
Dec	0	26 (72.4)	643 (90.8)	507 (85.2)	150 (78.9)	188 (87.2)	38 (86.3)	19 (62.7)

(Schoof and Pryor, p. 2478)

The numbers in the table represent how many weather stations out of 831 for each month of the year and order fit a Markov chain model, and whether the fit was determined using the Bayesian information criteria (BIC) or Kolmogorov-Smirnov (K-S) test. From the data, we can see that a significant percentage of the observed data from the weather stations fit a first-order Markov chain model. Using these findings, a 50-year projected Markov model of precipitation occurrence was created using 0th, 1st, 2nd, and 3rd order chains. The following plots show these models as well as cumulative-mass functions for precipitation during the month of July at a specific weather station in Sacaton.

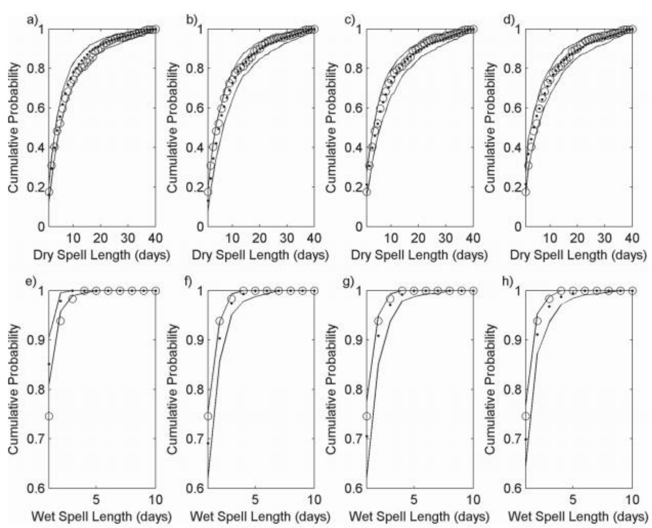


FIG. 4. Cumulative mass functions for (a)–(d) dry and (e)–(h) wet spells for July at Sacaton for observed data and Markov chains of order 0–3 (from left to right). The circles show the observed cumulative probabilities. The dots indicate the median of the modeled cumulative probabilities, and the solid lines indicate the 5th and 95th ordered cumulative probabilities based on bootstrap resampling.

(Schoof and Pryor, p. 2484).

The actual transition matrix for this study was not given, but the results of the Markov chain models were. We shall assume that the magnitude of the transition matrix would cause it to be too large to include in its entirety in a paper such as this, since it would include hundreds and hundreds of different probabilities. In this next section we will use MATLAB to create a Markov chain model for a much more simplified weather system.

IV. MATLAB Solution

Here we will consider a particular, simplified Markov chain model of the probability of precipitation over a discrete time interval. We will use MATLAB to create these models with two different methods. The first method will involve a matrix of transition probabilities for various states and using said Matrix to create a Markov chain, and the second will use a matrix of observed state changes. The latter method would be more applicable to the modeling of weather forecasts, since observed data is the starting point for these models in the real world.

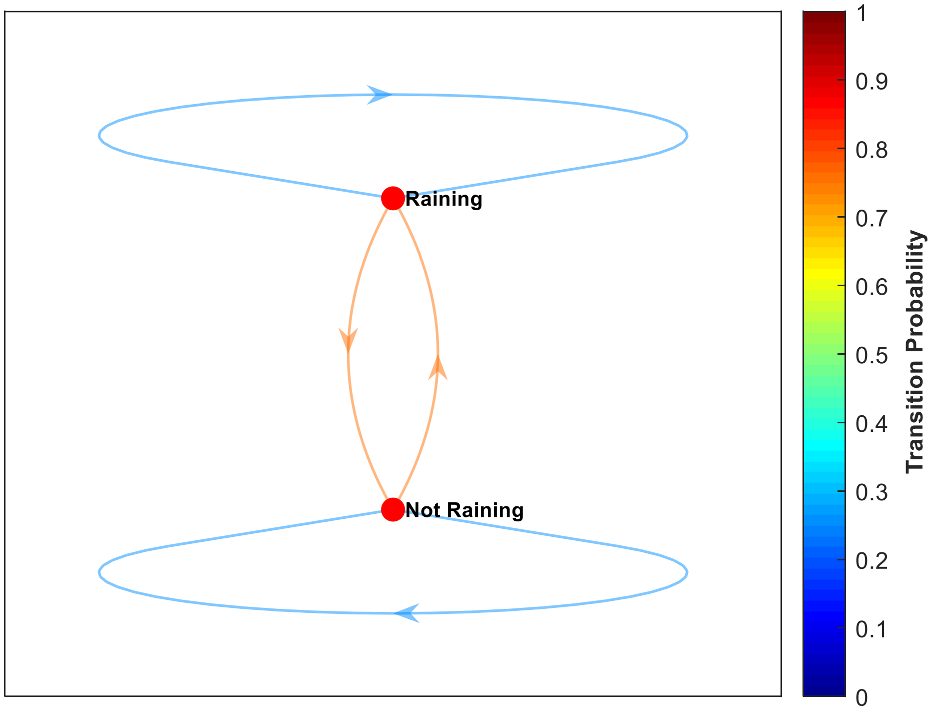
For the first example, let us consider the following 2x2 transition matrix for a Markov chain containing two states:

$$P = \begin{pmatrix} .25 & .75 \\ .75 & .25 \end{pmatrix}$$

For any entry P_{ij}, the value represents the probability that state i will transition to state j at time t+1 if the model is in state i at time t. If we say that state 1 is the absence of rain and state 2 is the presence of rain, from the transition matrix we see that at time t+1, there is a probability of .75 that it will rain if it is not raining at time t. Using the following MATLAB code

```
P = [.25 .75;.75 .25];
mc = dtmc(P,'StateNames',["Not Raining" "Raining"]);
figure(1);
graphplot(mc,'ColorEdges',true);
```

we can produce the following figure, which plots the Markov chain as a directed graph using colors to denote the transition probabilities.



(Figure 1)

We can observe from the plot that there is a .75 probability of state change from raining to not Raining at time *t*+1, and only a .25 probability of state steadiness. If this modeled a real weather system, it would mean that the weather would frequently switch from raining to not raining day after day, and only have the same weather two days in a row 25% of the time.

For the second example, we will produce a similar output using a matrix of observed transition counts. This is something that can be done more practically using observed data to predict future weather patterns. We will start with the following 2x2 matrix:

$$P = \begin{pmatrix} 33 & 27 \\ 48 & 12 \end{pmatrix}$$

Since this is a transition count matrix, the entries represent the number of state transitions from state *i* to state *j* over a given period of time. In this case, the sum of the entries in the matrix is

120, so it could represent a 120-day or roughly four month period. This would be interpreted as a transition from no rain to no rain for 33 days, no rain to rain for 27 days, rain to rain for 48 days, and rain to no rain for 12 days. Using the following MATLAB code,

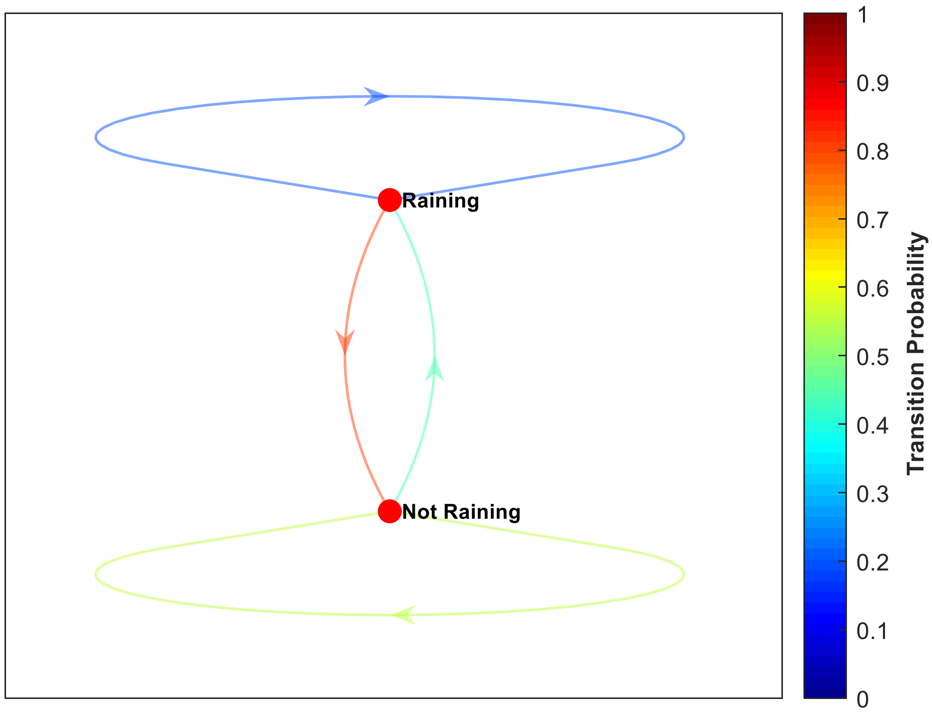
```
P = [33 27;
48 12];
mc = dtmc(P);
mc.P
```

we can use this data to produce a normalized transition matrix,

```
ans =

 0.5500    0.4500
 0.8000    0.2000
```

whose rows each sum to 1 (note that MATLAB is using row vectors as transition vectors as opposed to the column vectors that we used in previous examples). We can now plot the directed graph of the Markov chain using color to represent the transition probabilities along each directed edge between the nodes of the graph.



(Figure 2)

We can see from the colors the different probabilities of the state transitions, as determined by the normalized transition matrix created from the observed data. We can also observe the memorylessness demonstrated by this graph, as each state transition probability depends only on the previous state, not any of the others.

V. Conclusion and Summary

Markov chains are an effective tool for the modeling of stochastic systems. They can be used to model the behavior of any system that can be defined by a set of states and probabilities. Weather forecasting depends heavily on the use of Markov chains to define weather patterns and predict future weather states. Given any set of probabilities or observed numbers of state changes, a transition matrix can be formed to model future state changes of the system. This paper has merely scratched the surface of this application, as realistically there are hundreds of conditions, states, and probabilities involved in predicting the weather. In addition to merely predicting precipitation patterns, Markov chains can be used to protect communities of people by predicting the occurrences of weather disasters such as tornadoes and earthquakes, making their study a very productive use of time.

VI. References

Hefferon, J. (2009) *Linear Algebra: Third Edition.* Retrieved from http://joshua.smcvt.edu/linearalgebra/book.pdf

Leonard, G. (2011). History and Theoretical Basics of Hidden Markov Models. Hidden Markov Models, Theory and Applications. doi:10.5772/15205 http://cdn.intechweb.org/pdfs/15369.pdf (n.d.). Retrieved from https://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf

P. (n.d.). Retrieved from https://www.mathworks.com/help/econ/dtmc.html

Schoof, J. T., & Pryor, S. C. (2008). On the Proper Order of Markov Chain Model for Daily Precipitation Occurrence in the Contiguous United States. *Journal of Applied Meteorology & Climatology*, *47*(9), 2477–2486. https://doi-org.ezproxy.umuc.edu/10.1175/2008JAMC1840.1

Weather Forecasting Through the Ages. (n.d.). Retrieved from https://earthobservatory.nasa.gov/features/WxForecasting/wx2.php