Markov Chain Theory with Applications to Baseball

Cal D. Thomay

The College of Wooster, cal.thomay@gmail.com

Follow this and additional works at: https://openworks.wooster.edu/independentstudy

Part of the Statistics and Probability Commons

Recommended Citation
https://openworks.wooster.edu/independentstudy/5722

© Copyright 2014 Cal D. Thomay
Markov Chain Theory with Applications to Baseball

Independent Study Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Bachelor of Arts in the Department of Mathematics and Computer Science at The College of Wooster

by
Cal Thomay
The College of Wooster
2014

Advised by:
Dr. Robert Wooster
Abstract

The applications of Markov chains span a wide range of fields to which models have been designed and implemented to simulate random processes. Markov chains are stochastic processes that are characterized by their memoryless property, where the probability of the process being in the next state of the system depends only on the current state and not on any of the previous states. This property is known as the Markov property. This thesis paper will first introduce the theory of Markov chains, along with explaining two types of Markov chains that will be beneficial in creating a model for analyzing baseball as a Markov chain. The final chapter describes this Markov chain model for baseball, which we will use to calculate the expected number of runs scored for the 2013 College of Wooster baseball team. This paper finishes by displaying an analysis of sacrifice bunt and stolen base strategies through using the Markov chain model.
Acknowledgments

I would first like to thank God for all of the blessings He has given to me and for providing the opportunity for me to attend The College of Wooster and play baseball for such a prestigious program. I would like to thank my advisor Dr. Robert Wooster in assisting and guiding me through this thesis paper, as well as the mathematics department in preparing me for completion of this paper. I want to thank my grandparents and parents for the financial sacrifices that they have made to send me to The College of Wooster and receive a great education. Also, I am truly thankful for the love and encouragement that I have received from them and my two sisters over the years. I would also like to thank all of the great people I have met while in college, both students and professors, that have made my college experience unforgettable. I would especially like to thank all of the people involved in The College of Wooster baseball program over the last 4 years for the great friendships I have made and for everything that I have learned from them, both on and off the field.
Contents

Abstract iii

Acknowledgments v

1 Introduction 1

2 Markov Chains 5
   2.1 Transition Matrix . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
   2.2 Long-term Behavior of Markov Chains . . . . . . . . . . . . . . . . 9

3 Absorbing Markov Chains 17
   3.1 Drunkard’s Walk . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
   3.2 Canonical Form . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
   3.3 Probability of Absorption . . . . . . . . . . . . . . . . . . . . . . . . 20
   3.4 The Fundamental Matrix . . . . . . . . . . . . . . . . . . . . . . . . 21
   3.5 Expected Number of Steps to Absorption . . . . . . . . . . . . . . 25
   3.6 Absorption Probabilities . . . . . . . . . . . . . . . . . . . . . . . . 26
4 Ergodic Markov Chains

4.1 Fundamental Limit Theorem for Regular Markov Chains . . . . . 31
4.2 Examples of Ergodic Markov chains . . . . . . . . . . . . . . . 37

5 Baseball as a Markov chain

5.1 Setting up the Model . . . . . . . . . . . . . . . . . . . . . . . . . 46
5.2 Baseball Transition Probabilities . . . . . . . . . . . . . . . . . . . 48
5.3 Calculating Expected Number of Runs . . . . . . . . . . . . . . . 53
5.4 Player Value . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
5.5 Using Markov Chains for Baseball Strategy . . . . . . . . . . . . 57
  5.5.1 Sacrifice Bunt Strategy . . . . . . . . . . . . . . . . . . . . . 58
  5.5.2 Stolen Base Strategy . . . . . . . . . . . . . . . . . . . . . . 64

6 Conclusion

A Maple Worksheets

B Excel Worksheets
List of Figures

3.1 Drunkard’s walk at Ohio State ........................................ 19

4.1 The maze problem ........................................................ 42

A.1 Powers of the Cleveland weather transition matrix .......... 76
A.2 Powers of the Cleveland weather transition matrix .......... 77
A.3 Eigenvalues and eigenvectors of the Cleveland weather transi-
tion matrix ................................................................. 78
A.4 The fixed probability vector for the maze problem .......... 79

B.1 The first row of the fundamental matrix N in the baseball model . 82
B.2 The first row entries of the matrix B in the baseball model . . . . . . 83
B.3 Probability of scoring at least one run with and without a bunt
attempt starting with a runner on first base and no outs for The
College of Wooster ............................................................. 85
B.4 Probability of scoring at least one run with and without a bunt attempt starting with a runner on first base and no outs for Major League Baseball ................. 86

B.5 Probability of scoring at least one run with and without a stolen base attempt starting with a runner on first base and no outs, one out, and two outs for The College of Wooster ................. 87

B.6 Probability of scoring at least one run with and without a stolen base attempt starting with a runner on first base and no outs, one out, and two outs for Major League Baseball ................. 88
List of Tables

2.1 Powers of the Cleveland weather transition matrix ......... 13

5.1 The transient states of the baseball Markov chain ........ 48
5.2 Transition probabilities for The College of Wooster baseball team 52
5.3 Individual batter’s expected number of runs scored ........ 57
5.4 Run expectancies for each of the 24 states ............... 59
5.5 Transition probabilities for Major League Baseball in 2013 .... 61
5.6 Probability of scoring at least one run with and without a bunt attempt starting from the state with a runner on first base and no outs ......................................................... 61
5.7 Probability of scoring at least one run with a stolen base attempt . 65
5.8 Probability of scoring at least one run without a stolen base attempt 65
5.9 How successful a runner must be in stealing second base in terms of expected number of runs and probability of scoring at least one run ................................................................. 69
B.1 Transition probabilities from the 8th state to any of the 28 states in
the baseball transition matrix and powers of the transition matrix   84
Chapter 1

Introduction

Many aspects of life are characterized by randomly occurring events. It seems as though the world just doesn’t work as perfectly as we hope it would. In an effort to help quantify, model, and forecast the randomness of our world, the theory of probability and stochastic processes has been developed and may help answer some questions about how the world works [5]. The focus of this thesis paper is on one unique type of stochastic process known as Markov chains.

The theory of Markov chains developed during the early 20th century by a Russian mathematician named Andrei Andreyevich Markov. Learning mathematics under some famous Russian mathematicians such as Aleksandr Korkin and Pafnuty Chebyshev, Markov advanced his knowledge particularly in the fields of algebraic continued fractions and probability theory. His early work was dedicated mostly to number theory and analysis, continued fractions, limits of integrals, approximation theory and the convergence of series. Later in his life, however, he applied the method of continued fractions
CH**APTER 1. INTRODUCTION**

to probability theory guided by the influence of his teacher Pafnuty Chebyshev. Markov’s interest in the Law of Large Numbers and its extensions eventually led him to the development of what is now known as the theory of Markov chains, named after Andrei Markov himself [9].

A Markov chain is simply a sequence of random variables that evolves over time. It is a system that undergoes transitions between states in the system and is characterized by the property that the future is independent of the past given the present [5]. What this means is that the next state in the Markov chain depends only on the current state and not on the sequence of events that preceded it. This type of “memoryless” property of the past is known as the Markov property.

The changes between states of the system are known as transitions, and probabilities associated with various state changes are known as transition probabilities. A Markov chain is characterized by three pieces of information: a state space, a transition matrix with entries being transition probabilities between states, and an initial state or initial distribution across the state space. A state space is the set of all values which a random process can take. Furthermore, the elements in a state space are known as states and are a main component in constructing Markov chain models. With these three pieces, along with the Markov property, a Markov chain can be created and can model how a random process will evolve over time.

There are many interesting applications of Markov chains to other academic disciplines and industrial fields. For example, Markov chains have been used in Mendelian genetics to model and predict what future generations of a gene will look like. Another example of where Markov chains have been
applied to is in the popular children’s board game Chutes and Ladders. At each turn, a player is residing in a state in the state space (one square on the board), and from there the player has transition probabilities of moving to any other state in the state space. In fact, the transition probabilities are fixed since they are determined by the roll of a fair die. Nevertheless, the probability of moving to the next state is determined only by the current state and not how the player arrived there, and is therefore capable of being modeled as a Markov chain. In addition to both of these examples, Markov chains have been applied to areas as disparate as chemistry, statistics, operations research, economics, finance, and music. The application that we will focus on in this paper, however, is in baseball and the numerous aspects of the game that can be analyzed using Markov chains.

In Chapter 2 we provide an introduction to Markov chains and explain some properties about their long-term behavior. Following this chapter, we will discuss two important types of Markov chains that have been used in Markov chain models, namely absorbing Markov chains in Chapter 3 and ergodic Markov chains in Chapter 4. The theory that we present on absorbing Markov chains will be especially important when we discuss our Markov chain model for baseball in Chapter 5. This paper finishes with analysis of some baseball strategies using the Markov chain baseball model.
Chapter 2

Markov Chains

The following chapter will mainly focus on the basics of Markov chains, in which we will provide some useful definitions, properties, and theorems about Markov chains that will enable us to better understand them for analysis in the later chapters. Before we discuss Markov chains, however, we must first define the terms stochastic process and Markov process.

Definition 1. A stochastic process is a collection of random variables \( \{X_t, t \in T\} \) that are defined on the same probability space, where \( T \subseteq \mathbb{R} \). [2]

Stochastic processes are most commonly analyzed over discrete time \( \{X_t, t \in \mathbb{N}\} \), and continuous time \( \{X_t, 0 \leq t < \infty\} \). In this thesis paper, we will only consider stochastic processes in discrete time. There are many types of stochastic processes, but we will discuss only one area in particular: Markov processes. The following definition is found in [3].

Definition 2. Let \( \{X_1, X_2, \ldots\} \) be a sequence of random variables defined on a common probability space \( \Omega \) (i.e., \( X_j : \Omega \to S \subseteq \mathbb{R} \) for \( j \in \mathbb{N} = \{1, 2, 3, \ldots\} \)). We call S the
state space. Then \( \{X_1, X_2, \ldots \} \) is called a **Markov process** with state space \( S \) if

\[
P(X_{n+1} = s_{n+1} \mid X_n = s_n, \ldots, X_2 = s_2, X_1 = s_1) = P(X_{n+1} = s_{n+1} \mid X_n = s_n) \quad (2.1)
\]

holds for any \( n = 1, 2, \ldots \) and any \( s_1, s_2, \ldots, s_{n+1} \) with \( s_k \in S \) for \( 1 \leq k \leq n + 1 \).

A **Markov chain** is a Markov process with a finite number of states while still satisfying equation (2.1), which is known as the Markov property. These states, \( S = \{s_1, s_2, \ldots, s_r\} \), are the main components of stochastic processes. All of the possible values that each \( X_n \) can take on are called states. A stochastic process begins in a state and moves from one state to another by what is called a **step**. The Markov property indicates that future outcomes of a Markov process can be predicted by using only the outcome of the current state and neglecting any other information obtained about past states. The current state of the system is the only essential state in a Markov process, and the steps leading up to the current state are meaningless. It is somewhat of a “memoryless” property that characterizes a Markov process [6].

Before presenting Proposition 1 and Theorem 1 and the proofs of each, it is necessary to recall the definition of conditional probability. If \( A \) and \( B \) are two events in the state space \( S \) with \( P(B) \neq 0 \), then the **conditional probability** \( A \) given \( B \) is the probability that \( A \) will occur, given that \( B \) is known to occur or has already occurred. It is defined as

\[
P(A \mid B) = \frac{P(A \cap B)}{P(B)}.
\]

If \( A \) and \( B \) are independent events, then \( P(A \mid B) = P(A) \).
CHAPTER 2. MARKOV CHAINS

Let us now consider a discrete-time Markov chain that is currently in state \( s_i \) and moves to state \( s_j \) with a certain probability denoted by \( p_{ij} \). This probability, \( p_{ij} \), does not depend upon which states the chain was in before arriving at the current state. It only considers the probability of moving from state \( s_i \) to state \( s_j \). The probabilities \( p_{ij} \) are called transition probabilities. For a unique example, the transition probability \( p_{ii} \) represents the probability that the process remains in state \( s_i \) after one step.

2.1 Transition Matrix

Let \( \{X_k\} \) be a discrete-time Markov chain with a finite state space, \( S = \{1, 2, \ldots, r\} \). Consider the transition probabilities \( p_{ij} \), where \( i = 1, 2, \ldots, r; j = 1, 2, \ldots, r \). Thus, there are \( r^2 \) transition probabilities. To easily group the transition probabilities, a transition matrix \( P \) is formed with these values as entries. The matrix \( P \), where the \( i \)th row and \( j \)th column of \( P \) represent the probabilities of moving from the \( i \)th state to the \( j \)th state of \( S \), results in

\[
P = \begin{pmatrix}
    p_{1,1} & p_{1,2} & \cdots & p_{1,r} \\
p_{2,1} & p_{2,2} & \cdots & p_{2,r} \\
    \vdots & \vdots & \ddots & \vdots \\
p_{r,1} & p_{r,2} & \cdots & p_{r,r}
\end{pmatrix}.
\]

One comment about notation is that we will use the letter \( P \) for both a transition matrix and for taking the probability of some event. However, it should be clear from the context which reference we mean.


**Proposition 1.** Every transition matrix has the following properties:

1. \( p_{ij} \geq 0; \ i, j \in S \)

2. \( \sum_{j=1}^{r} p_{ij} = 1 \) for all \( i \in S \)

The proof we present follows what is shown in [7].

**Proof.** Property 1 is clearly true since the probability of moving from one state to another cannot be negative. To prove property 2, fix \( i \in S \) and \( k \in \mathbb{N} \). Now, the sum of the \( i \)th row of \( P \) can be written as \( p_{i1} + p_{i2} + \cdots + p_{in} \). So we have

\[
\sum_{j=1}^{r} p_{ij} = p_{i1} + p_{i2} + \cdots + p_{in}
\]

\[
= \frac{P(X_k = 1 \mid X_{k-1} = i) + P(X_k = 2 \mid X_{k-1} = i) + \cdots + P(X_k = n \mid X_{k-1} = i)}{P(X_{k-1} = i)}
\]

\[
= \frac{P(X_k = 1 \cap X_{k-1} = i) + P(X_k = 2 \cap X_{k-1} = i) + \cdots + P(X_k = n \cap X_{k-1} = i)}{P(X_{k-1} = i)}
\]

\[
= \frac{P[(X_k = 1 \cap X_{k-1} = i) \cup (X_k = 2 \cap X_{k-1} = i) \cup \cdots \cup (X_k = n \cap X_{k-1} = i)]}{P(X_{k-1} = i)}
\]

\[
= \frac{P[(X_k = 1) \cup (X_k = 2) \cup \cdots \cup (X_k = n)] \cap (X_{k-1} = i)}{P(X_{k-1} = i)}
\]

\[
= P[X_k \in S \mid X_{k-1} = i] = 1.
\]

\[\square\]
2.2 Long-term Behavior of Markov Chains

Example 1. Suppose the weather in Cleveland, Ohio is recorded according to whether it is a nice, rainy, or snowy day. If it is raining in Cleveland one day, then it will not snow the next day. Also, the probability of it raining the day after a rainy day is 2/3, with a 1/3 probability of it being nice the next day. If it is a nice day in Cleveland, then it will be nice the next day half of the time, with the other half split evenly between being a rainy or snowy day. If it is snowing in Cleveland, then it will be snowing the next day with a probability of 2/3, with the other 1/3 probability split evenly between being a rainy or nice day. With this information, we can form a Markov chain with the types of weather as the three states. We will assume that state 1 is rainy weather, state 2 is nice weather, and state 3 is snowy weather. From the above information, we can establish the transition probabilities and form the transition matrix \( P \) which is shown below.

\[
P = \begin{pmatrix}
1 & 2 & 3 \\
1 & \frac{2}{3} & \frac{1}{3} & 0 \\
2 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
3 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{pmatrix}
\]

Let \( A_{ij}^{(k)} \) denote the event that a Markov chain moves from state \( s_i \) to state \( s_j \) in exactly \( k \) steps. Using the information from Example 1, consider \( P(A_{ij}^{(2)}) \). To determine this probability, we can find the probability of the weather being nice two days from now given that it is snowy today, for example. This probability is the disjoint union of the following three events:

1. It is rainy tomorrow and nice two days from now.
2. It is nice tomorrow and nice two days from now.

3. It is snowy tomorrow and nice two days from now.

The probability of the first event is the product of the conditional probability that it is rainy tomorrow, given that it is snowy today, and the conditional probability that it is nice two days from now, given that it is rainy tomorrow. Based on the transition matrix $P$, this product can be written as $p_{31}p_{12}$. Similarly, the other two events can be written as a product of entries of $P$. Therefore, we have

$$P(A_{32}^{(2)}) = p_{31}p_{12} + p_{32}p_{22} + p_{33}p_{32}.$$  

This equation can also be derived by calculating the dot product of the third row of $P$ with the second column of $P$. In general, if a Markov chain has $r$ states, then

$$P(A_{ij}^{(2)}) = \sum_{k=1}^{r} p_{ik}p_{kj}.$$  

As we will show in Theorem 1, the powers of a transition matrix give us very useful information about a Markov chain. Each power of the matrix shows the individual probabilities of going from one state to another after a certain number of steps corresponding to the power of the matrix. For example, the third power of a transition matrix gives the probabilities of moving between states after three steps in a Markov chain.

In order to prove Theorem 1, we give the following lemma and proof found in [8].
Lemma 1. Let $B_1, B_2, \ldots, B_r$ be a set of events that partition the sample space $S$, and $P(B_l) > 0$ for $l = 1, 2, \ldots, r$. The law of total probability states that for any event $A$,

$$P(A) = \sum_{l=1}^{r} P(A \mid B_l)P(B_l).$$

Proof. We have

$$P(A) = P(A \cap S) = P \left[ A \cap \left( \bigcup_{l=1}^{r} B_l \right) \right] = P \left[ \bigcup_{l=1}^{r} (A \cap B_l) \right]$$

$$= \sum_{l=1}^{r} P(A \cap B_l) = \sum_{l=1}^{r} P(A \mid B_l)P(B_l)$$

since $B_1, B_2, \ldots, B_r$ form a partition of $S$, and $A \cap B_1, A \cap B_2, \ldots, A \cap B_r$ are disjoint events.

Using Lemma 1, the following theorem from [6] shows us that the properties we have stated about transition matrices hold for any natural number power of a transition matrix.

Theorem 1. Let $P$ be the transition matrix of a Markov chain and let $n \in \mathbb{N}$. The $ij$th entry $p_{ij}^{(n)}$ of the matrix $P^n$ gives the probability that the Markov chain, starting in state $s_i$, will be in state $s_j$ after $n$ steps.

Proof. We will prove this theorem by mathematical induction.

Initial Case ($n=1$): By definition of a transition probability, the $ij$th entry $p_{ij}^{(1)}$ of the matrix $P^1$ gives the probability that the Markov chain, starting in state $s_i$, will be in state $s_j$ after 1 step.

Induction Step: Assume that $p_{ij}^{(k)}$ gives the probability that the Markov chain moves from state $s_i$ to state $s_j$ in exactly $k$ steps, for some $k \in \mathbb{N}$. Also, assume
that the Markov chain has \( r \) states. We want to show that \( p_{ij}^{(k+1)} \) gives the probability that the Markov chain moves from state \( s_i \) to \( s_j \) in exactly \( k + 1 \) steps.

The probability \( p_{ij}^{(k+1)} \) represents the \( ij \)th entry of the matrix \( P^{k+1} \), where \( P^{k+1} = P^k P \). Knowing this, we can say

\[
p_{ij}^{(k+1)} = p_{i1}^{(k)} p_{1j} + p_{i2}^{(k)} p_{2j} + \cdots + p_{ir}^{(k)} p_{rj} = \sum_{l=1}^{r} p_{il}^{(k)} p_{lj},
\]

(2.2)

Now, one way to move from state \( s_i \) to state \( s_j \) is to stop in one of the \( r \) states after \( k \) steps, and then move from that state to state \( s_j \) in one step. Let \( A_{ij}^{(k)} \) denote the event that the Markov chain moves from state \( s_i \) to state \( s_j \) in exactly \( k \) steps. Now by Lemma 1,

\[
P \left( A_{ij}^{(k+1)} \right) = \sum_{l=1}^{r} P \left( A_{ij}^{(k+1)} \mid A_{il}^{(k)} \right) P \left( A_{il}^{(k)} \right).
\]

Furthermore, by the Inductive Hypothesis and the Markov Property, we have

\[
P \left( A_{ij}^{(k+1)} \right) = \sum_{l=1}^{r} p_{lj} p_{il}^{(k)}.
\]

(2.3)

Since equation (2.2) shows that \( p_{ij}^{(k+1)} = \sum_{l=1}^{r} p_{il}^{(k)} p_{lj} \) and equation (2.3) shows that \( P \left( A_{ij}^{(k+1)} \right) = \sum_{l=1}^{r} p_{lj} p_{il}^{(k)} \), we can say that \( p_{ij}^{(k+1)} = P \left( A_{ij}^{(k+1)} \right) \). In other words, \( p_{ij}^{(k+1)} \) gives the probability that the Markov chain will move from state \( s_i \) to state \( s_j \) in exactly \( k + 1 \) steps.

**Conclusion:** Therefore, the \( ij \)th entry \( p_{ij}^{(n)} \) of the matrix \( P^n \) gives the probability that the Markov chain, starting in state \( s_i \), will be in state \( s_j \) after \( n \) steps for all \( n \in \mathbb{N} \). \( \square \)
CHAPTER 2. MARKOV CHAINS

Recall Example 1 and the strange weather it describes for Cleveland. Theorem 1 gives us information about finding the probability of the Markov chain being in a certain state after a specified number of steps. Using Maple, the powers of $P$ can be obtained and analyzed to see how the Markov chain will evolve after many steps. The following matrices in Table 2.1 show a few of the powers of $P$, with entries rounded to three decimal places. Note that states 1, 2, and 3 are in the same positions as in Example 1 and other powers of $P$ can be seen in Figures A.1 and A.2 in Appendix A.

Table 2.1: Powers of the Cleveland weather transition matrix

\[
P = \begin{pmatrix} 0.667 & 0.333 & 0 \\ 0.250 & 0.500 & 0.250 \\ 0.167 & 0.167 & 0.667 \end{pmatrix} \quad P^3 = \begin{pmatrix} 0.463 & 0.384 & 0.153 \\ 0.319 & 0.294 & 0.387 \\ 0.393 & 0.349 & 0.257 \end{pmatrix}
\]

\[
P^6 = \begin{pmatrix} 0.387 & 0.345 & 0.267 \\ 0.379 & 0.338 & 0.283 \\ 0.392 & 0.348 & 0.260 \end{pmatrix} \quad P^9 = \begin{pmatrix} 0.391 & 0.347 & 0.262 \\ 0.389 & 0.346 & 0.265 \\ 0.391 & 0.348 & 0.261 \end{pmatrix}
\]

\[
P^{12} = \begin{pmatrix} 0.391 & 0.348 & 0.261 \\ 0.391 & 0.348 & 0.262 \end{pmatrix} \quad P^{14} = \begin{pmatrix} 0.391 & 0.348 & 0.261 \\ 0.391 & 0.348 & 0.261 \end{pmatrix}
\]

Based on this data, after only 14 steps the probabilities of each weather pattern occurring are approximately 0.391, 0.348, and 0.261, regardless of where the chain started. In other words, after 14 days, the predictions for rainy, snowy, or nice weather for the subsequent day are independent of the first day’s weather. So for example, the initial probability of it raining in Cleveland a day after a rainy day is 0.667. But the probability of it raining 15, 24, or 29 days after a rainy day is approximately 0.391 for each set of days.

While considering the long-term behavior of Markov processes, it is also necessary to define the term \textit{probability vector}. 

**Definition 3.** A *probability vector* is a vector whose entries are nonnegative and sum to 1.

In Markov chain theory, a probability vector serves as the probability distribution of each state after a certain step. Knowing the definition of a probability vector, the following theorem found in [6] relates to the long-term behavior of a Markov chain. For the remainder of this thesis paper, vectors will be denoted as boldface lower case letters.

**Theorem 2.** Let $P$ be the transition matrix of a Markov chain, and let $\mathbf{u} = \langle u_1, u_2, \ldots, u_r \rangle$ be the probability vector which represents the starting distribution of states. Then the probability that the chain is in state $s_i$ after $n$ steps is the $i$th entry in the row vector

$$\mathbf{u}^{(n)} = \mathbf{u}^P.$$ 

**Proof.** Let $B$ be the event that the Markov chain is in state $s_i$ after $n$ steps given that $\mathbf{u}$ is the probability vector which represents the starting distribution of states. Let $A_k$ be the event that the Markov chain starts in state $s_k$. By Lemma 1 and Theorem 1,

$$P(B) = \sum_{k=1}^r P(B \mid A_k)P(A_k) = \sum_{k=1}^r p_{ki}^{(n)}u_k. \quad (2.4)$$

Let $\mathbf{e}_k$ denote the standard basis vector $\mathbf{e}_k = \langle 0, \ldots, 0, 1, 0, \ldots, 0 \rangle$, where the $k$th entry is 1. Now, consider $\mathbf{u}^P$. We have

$$\mathbf{u}^P = \left( \sum_{k=1}^r u_k \mathbf{e}_k \right) \mathbf{P}^{(n)} = \sum_{k=1}^r u_k (\mathbf{e}_k \mathbf{P}^{(n)}) = \sum_{k=1}^r u_k \left( \sum_{j=1}^r p_{kj}^{(n)} \mathbf{e}_j \right).$$
To find the $i$th entry of $\mathbf{u}P^n$, we can dot it with $\mathbf{e}_i$. So we have

$$\mathbf{e}_i \cdot \mathbf{u}P^n = \mathbf{e}_i \cdot \left[ \sum_{k=1}^{r} u_k \left( \sum_{j=1}^{r} p_{kj}^{(n)} \mathbf{e}_j \right) \right] = \sum_{k=1}^{r} \sum_{j=1}^{r} u_k p_{kj}^{(n)} \mathbf{e}_i \cdot \mathbf{e}_j = \sum_{k=1}^{r} u_k p_{ki}^{(n)}. \quad (2.5)$$

Since equations (2.4) and (2.5) are equivalent, we have shown that the probability that the Markov chain is in state $s_i$ after $n$ steps is the $i$th entry of the row vector $\mathbf{u}P^n$. \hfill \Box

In the following two chapters, we will discuss in detail two different types of Markov chains, namely absorbing Markov chains and ergodic Markov chains.
Chapter 3

Absorbing Markov Chains

There are many different types of Markov chains, and the first ones that we will cover are absorbing Markov chains. Much of the following presented theory on absorbing Markov chains will be implemented in the applications to baseball in Chapter 5.

**Definition 4.** A state $s_i$ of a Markov chain is called **absorbing** if it is impossible to leave it (i.e., $p_{ii}=1$). A Markov chain is **absorbing** if it has at least one absorbing state, and if from every state it is possible to go to an absorbing state, but not necessarily in one step. [6]

**Definition 5.** In an absorbing Markov chain, a state which is not absorbing is called **transient**. [6]

Suppose an absorbing Markov chain has state space $S = \{s_1, s_2, \ldots, s_r\}$. Suppose further that there exist $m$ transient states. Labeling each transient state as $t_i \in S$, we have $\{t_1, t_2, \ldots, t_m\}$ as the set of transient states. Then, labeling each absorbing state as $a_i \in S$, we have $\{a_1, a_2, \ldots, a_{r-m}\}$ as the set of absorbing
states. Therefore, \( \{t_1, t_2, \ldots, t_m\} \cup \{a_1, a_2, \ldots, a_{r-m}\} = S \).

3.1 Drunkard’s Walk

**Example 2.** Consider a male student at The Ohio State University walking along a four-block stretch of High Street on a Friday night. He continues walking until he either reaches the bar at corner 4 or his apartment at corner 0. He notices many female students outside of the bar along corners 1, 2, and 3, which increases the probability of him staying at these corners, and decreases the probability of him walking either towards the bar or towards his apartment. Also, he is always in the process of moving from corner to corner. If he reaches the bar at corner 4 or his apartment at corner 0, then he stays there. We can form a Markov chain consisting of 5 states (labeled 0-4), with states 0 and 4 as absorbing states. The transition matrix for this Markov chain is given as

\[
P = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 1/4 & 0 & 3/4 & 0 \\
1 & 1/2 & 0 & 1/2 & 0 \\
3 & 0 & 0 & 3/4 & 0 & 1/4 \\
4 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\]

Notice from this transition matrix that the transient states are 1, 2, and 3, and from these states it is possible to reach the absorbing states 0 and 4, but not necessarily in one step. Therefore, this Markov chain is an absorbing chain. When a Markov chain reaches an absorbing state, it is referred to as being
absorbed. See Figure 3.1 below for the digraph of the Drunkard’s Walk. The arrows show which state he is going to and from with the appropriate probability labeled. [6]

![Figure 3.1: Drunkard’s walk at Ohio State](image)

3.2 Canonical Form

Suppose we have an arbitrary absorbing Markov chain. Renumber the states so that the transient states are listed first and the absorbing states listed last. Let there be $a$ absorbing states and $t$ transient states. The transition matrix’s canonical form is

$$P = \begin{pmatrix} TR & ABS \\ TR & Q & R \\ ABS & 0 & I_a \end{pmatrix}$$

Here, $Q$ is a $t \times t$ matrix, $R$ is a nonzero $t \times a$ matrix, $0$ is an $a \times t$ zero matrix and $I_a$ is the $a \times a$ identity matrix [6].

We have shown before that the entry $p_{ij}^{(n)}$ of the transition matrix $P^n$ gives the probability of being in state $s_j$ after $n$ steps, given that the chain starts in
state $s_i$. From matrix algebra, we find that $P^n$ is of the form

$$P^n = \begin{pmatrix} \text{TR.} & \text{ABS.} \\ \text{TR.} & Q^n \\ \text{ABS.} & 0 \end{pmatrix}$$

where the symbol * stands for a $t \times a$ matrix with entries consisting of components from both $Q$ and $R$ [6]. What we can gain from this matrix is that the sub-matrix $Q^n$ gives the probabilities of moving from a transient state to another transient state after $n$ steps. This matrix will be useful in the theory presented later in this chapter.

### 3.3 Probability of Absorption

Theorem 3 below on the probability that an absorbing Markov chain will eventually be absorbed follows that given in [6]. This theorem will be useful to know when we explain our Markov chain model for baseball in Chapter 5.

**Theorem 3.** In an absorbing Markov chain, the probability that the process will be absorbed is 1 (i.e., $\lim_{n \to \infty} Q^n = 0$).

**Proof.** Since we have an absorbing Markov chain, from each non-absorbing state $s_j \in S$ it is possible to reach an absorbing state. Let $m_j$ be the minimum number of steps required to reach an absorbing state, assuming we start from state $s_j$. Also, let $p_j$ be the probability that, starting from $s_j$, the process will not reach an absorbing state in $m_j$ steps. It follows that $p_j < 1$. Now, let $m$ be the largest of $m_j$ and let $p$ be the largest of $p_j$ as well. The probability of $s_j$ not
being absorbed in \( m \) steps is less than or equal to \( p \), in \( 2m \) steps less than or equal to \( p^2 \), etc. Since we know that \( p < 1 \), these probabilities approach 0. Therefore, since the probability of not being absorbed in \( n \) steps is monotone decreasing and is approaching 0, we have \( \lim_{n \to \infty} Q^n = 0 \). Thus, the probability that an absorbing Markov chain will eventually be absorbed is equal to 1. \( \square \)

### 3.4 The Fundamental Matrix

Recall that \( Q \) is a \( t \times t \) matrix consisting of transition probabilities between \( t \) transient states of a Markov chain. The theorem and proof below follow from [6] and uses this matrix.

**Theorem 4.** Consider an absorbing Markov chain. The matrix \( I - Q \) has an inverse \( N \), with \( N = I + Q + Q^2 + \ldots \). The \( ij \)th entry \( n_{ij} \) of the matrix \( N \) is the expected number of times the chain is in state \( s_j \), given that it started in state \( s_i \). If \( i = j \), then the initial state is counted.

**Proof.** Let \((I - Q)x = 0\). This can be equivalently written as \( x = Qx \). If we multiply by \( Q \) on the left, we are left with \( Qx = Q^2x \). Since \( x = Qx \), this equation now becomes \( x = Q^2x \). We can iterate this absorbing Markov chain and see that after \( n \) steps, we have \( x = Q^n x \). From Theorem 3, we know that \( \lim_{n \to \infty} Q^n = 0 \). Thus we have \( \lim_{n \to \infty} Q^n x = 0 \), and so \( x = 0 \). Now, if \( (I - Q)x = 0x \) and we were to have \( x \neq 0 \), then \( I - Q \) would have 0 as an eigenvalue. But, we know \( x \) must be 0, so 0 is not an eigenvalue for \( I - Q \). Thus, we know the matrix \( I - Q \) is invertible. So we have that \( N = (I - Q)^{-1} \) exists.
CHAPTER 3. ABSORBING MARKOV CHAINS

Next, note that

\[(I - Q)(I + Q + Q^2 + \cdots + Q^n) = I - Q^{n+1}.\]

Now, multiplying on the left by \(N\) gives

\[I + Q + Q^2 + \cdots + Q^n = N(I - Q^{n+1}).\]

Taking the limit as \(n\) approaches infinity, we have

\[N = I + Q + Q^2 + \cdots\]

since \(\lim_{n \to \infty} Q^n = 0.\)

Let \(s_i\) and \(s_j\) be transient states, where \(i\) and \(j\) are fixed. Let \(X^{(k)}\) be a 0-1 random variable which equals 1 if the chain is in state \(s_j\) after \(k\) steps starting from state \(s_i\), and equals 0 otherwise. For every \(k\), \(X^{(k)}\) depends upon both \(i\) and \(j\). Now, we have

\[P(X^{(k)} = 1) = q_{ij}^{(k)},\]

and

\[P(X^{(k)} = 0) = 1 - q_{ij}^{(k)},\]

where \(q_{ij}^{(k)}\) is the \(ij\)th entry of the matrix \(Q^k\). Also note that these equations hold for \(k = 0\) since \(Q^0 = I\). Now, the expected value of \(X^{(k)}\), written as \(E(X^{(k)})\), is

\[E(X^{(k)}) = 1 \left(q_{ij}^{(k)}\right) + 0 \left(1 - q_{ij}^{(k)}\right) = q_{ij}^{(k)}.\]
Furthermore, the expected number of times the chain is in state $s_j$ in the first $n$ steps, given that it starts in state $s_i$, is

$$E(X^{(0)} + X^{(1)} + \cdots + X^{(n)}) = E(X^{(0)}) + E(X^{(1)}) + \cdots + E(X^{(n)}) = q^{(0)}_{ij} + q^{(1)}_{ij} + \cdots + q^{(n)}_{ij}.$$

Letting $n$ tend to infinity, we have

$$E(X^{(0)} + X^{(1)} + \cdots) = q^{(0)}_{ij} + q^{(1)}_{ij} + \cdots = n_{ij}. \qedhere$$

**Definition 6.** Let $P$ be an absorbing Markov chain. Then the fundamental matrix for $P$ is the matrix

$$N = (I - Q)^{-1},$$

where $Q$ consists only of transition probabilities between transient states of $P$.

Knowing the fundamental matrix will be very useful when we describe our baseball Markov chain model in Chapter 5. The property that the entries of the fundamental matrix give the expected number of times that the Markov chain will be in a transient state until absorption will be beneficial in calculating the expected number of runs scored for The College of Wooster baseball team during the 2013 baseball season.

**Example 3.** Let us consider Example 2 again and find the fundamental matrix
of its transition matrix. The canonical form of its transition matrix is

\[
P = \begin{pmatrix}
0 & 3/4 & 0 & 1/4 & 0 \\
1/2 & 0 & 1/2 & 0 & 0 \\
0 & 3/4 & 0 & 0 & 1/4 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix},
\]

where the rows and columns are in the order of states 1, 2, 3, 0, and 4 from top to bottom and left to right, respectively.

From this, we can see that the matrix \( Q \) is

\[
Q = \begin{pmatrix}
0 & 3/4 & 0 \\
1/2 & 0 & 1/2 \\
0 & 3/4 & 0 \\
\end{pmatrix},
\]

and

\[
I - Q = \begin{pmatrix}
1 & -3/4 & 0 \\
-1/2 & 1 & -1/2 \\
0 & -3/4 & 1 \\
\end{pmatrix}.
\]

Computing \( N = (I - Q)^{-1} \) in Maple, we find that

\[
N = (I - Q)^{-1} = \begin{pmatrix}
1 & 2 & 3 \\
1 & 5/2 & 3 & 3/2 \\
2 & 4 & 2 \\
3 & 3/2 & 3 & 5/2 \\
\end{pmatrix}.
\]
We know that the entry $n_{ij}$ of matrix $N$ gives the expected number of times that the Markov chain is in the transient state $s_j$, given that it started in the transient state $s_i$. So for example, if we start in state 3, the expected number of times the chain will be in states 1, 2, and 3 before being absorbed are $3/2$, $3$, and $5/2$, respectively.

### 3.5 Expected Number of Steps to Absorption

Let us now consider finding the expected number of steps before an absorbing Markov chain is absorbed. The theorem and proof below follow that given from [6].

**Theorem 5.** Let $x_i$ be the expected number of steps before an absorbing Markov chain is absorbed, given that the chain starts in state $s_i$. Also, let $x$ be the column vector whose $i$th entry is $x_i$. Then we have

$$x = Nc,$$

where $c$ is a column vector all of whose entries are 1.

**Proof.** Suppose we add all of the entries in the $i$th row of $N$ together. This will give us the expected number of times the Markov chain will be in any of the transient states, starting in state $s_i$, before being absorbed. Therefore, $x_i$ is the sum of the entries in the $i$th row of $N$. Writing this statement in matrix form for all $i$, we have $x = Nc$, as desired. \qed
3.6 Absorption Probabilities

Another interesting characteristic that a transition matrix reveals is the probability that an absorbing chain will be absorbed in one of the absorbing states, given that it starts in a transient state. The theorem below gives an equation to find these probabilities. The theorem and proof follow that given in [6].

**Theorem 6.** Let $b_{ij}$ be the probability that an absorbing chain will be absorbed in the absorbing state $s_j$, given that it starts in the transient state $s_i$. Then $B$ is a $t \times a$ matrix with entries $b_{ij}$ and

$$B = NR,$$

where $N$ is the fundamental matrix and $R$ is as in canonical form.

**Proof.** Fix $1 \leq i \leq t$ and $1 \leq j \leq a$ where $t_i$ is a transient state and $a_j$ is an absorbing state, with $t_i, a_j \in S$. We have

$$b_{ij} = \sum_{n=0}^{\infty} \sum_{k=1}^{t} q_{ik}^{(n)} r_{kj} = \sum_{n=0}^{\infty} \sum_{k=1}^{t} q_{ik}^{(n)} r_{kj} = \sum_{k=1}^{t} n_{ik} r_{kj} = (NR)_{ij}.$$ 

Thus, $B = NR$ for all $i$ and $j$. \hfill \square

Similar to the fundamental matrix of an absorbing Markov chain, the matrix $B = NR$ will be very useful in our model in determining the expected number of runs scored for The College of Wooster baseball team.

**Example 4.** Continuing on again with the Drunkard’s Walk example in Example 2, we can find the expected number of steps to absorption and
absorption probabilities. From Example 3, we found that

\[
N = \begin{pmatrix}
1 & 2 & 3 \\
5/2 & 3 & 3/2 \\
2 & 4 & 2 \\
3/2 & 3 & 5/2
\end{pmatrix}.
\]

Now, to find the expected number of steps to absorption starting from each state, we compute

\[
x = Nc = \begin{pmatrix}
5/2 & 3 & 3/2 \\
2 & 4 & 2 \\
3/2 & 3 & 5/2
\end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 8 \\ 7 \end{pmatrix}.
\]

Thus, the expected times to absorption, starting in states 1, 2, and 3, are 7, 8, and 7, respectively. From the canonical form, we have

\[
R = \begin{pmatrix}
0 & 4 \\
1/4 & 0 \\
0 & 0 \\
0 & 1/4
\end{pmatrix}.
\]

Therefore,

\[
B = NR = \begin{pmatrix}
5/2 & 3 & 3/2 \\
2 & 4 & 2 \\
3/2 & 3 & 5/2
\end{pmatrix} \begin{pmatrix}
1/4 & 0 \\
0 & 0 \\
0 & 1/4
\end{pmatrix} = \begin{pmatrix} 5/8 & 3/8 \\ 1/2 & 1/2 \\ 3/8 & 5/8 \end{pmatrix}.
\]
Each row of this matrix tells us the absorption probabilities starting in the corresponding state. For example, the third row reveals that, starting in state 3, there is a probability of $3/8$ of absorption in state 0 and a probability of $5/8$ of absorption in state 4.
Chapter 4

Ergodic Markov Chains

Another important type of Markov chain that we will investigate further is an ergodic Markov chain. A Markov chain is considered *ergodic* if it is possible to reach every state from any one state, but not necessarily in one step [6]. In many sources, ergodic Markov chains are also referred to as being *irreducible*. A Markov chain is considered *regular* if some power of its transition matrix has all of its entries as being greater than 0. In other words, a Markov chain is regular if for some natural number $n$, it is possible to start at any state and reach any state in exactly $n$ steps. So clearly a regular Markov chain is ergodic. However, it is not necessarily true that an ergodic Markov chain is regular. The following example shown in [6] is known as the Ehrenfest Model, and is an example of an ergodic but non-regular Markov chain.

**Example 5.** Suppose we have two urns that contain four balls between them. At each step, one of the four balls is chosen at random and moved to the opposite urn than it was in previously. A Markov chain is formed with states chosen as being the number of balls present in the first urn. The transition
matrix for this chain is

\[
P = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 0 & 1 & 0 & 0 \\
1 & 1/4 & 0 & 3/4 & 0 \\
2 & 0 & 1/2 & 0 & 1/2 \\
3 & 0 & 0 & 3/4 & 0 & 1/4 \\
4 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}.
\]

In this example, if we were to start in state 0, then we will be in state 0, 2, or 4 only after an even number of steps, and in state 1 or 3 only after an odd number of steps. So we can reach every state from state 0, but not in the same number of steps. Similarly, we can reach every state if we were to start in state 1, 2, 3, or 4, but we cannot reach every state in the same number of steps from each of these starting states. Thus, we can reach every state from every state, but we can never reach every state in the same number of steps. Therefore, the Ehrenfest Model is a non-regular ergodic Markov chain.

If some power of the transition matrix of a Markov chain contains no zeros, then it is a regular Markov chain. However, it is possible for a regular Markov chain to have a transition matrix with some entries as zeros. The transition matrix from Example 1 has \( p_{13} = 0 \), but Table 2.1 shows that the third power of its transition matrix contains no zeros. Thus, this is a regular Markov chain.

**Definition 7.** Let \( P \) be the transition matrix for a Markov chain. Then a row vector \( \mathbf{w} \) satisfying the equation \( \mathbf{w}P = \mathbf{w} \) is called a **fixed row vector** for \( P \). Likewise, a column vector \( \mathbf{x} \) satisfying \( Px = x \) is called a **fixed column vector** for \( P \). [6]
One piece to note is that a fixed column vector has components that are all equal to each other. We can also say that a fixed row vector is a left eigenvector of the matrix $P$ corresponding to an eigenvalue of 1. Similarly, a fixed column vector is a right eigenvector of $P$ corresponding to an eigenvalue of 1.

### 4.1 Fundamental Limit Theorem for Regular Markov Chains

The Fundamental Limit Theorem for regular Markov chains is one of the main theorems regarding Markov chains. It shows that the long term behavior of a regular Markov chain has an equilibrium type of behavior. Before presenting the theorem, the following example and lemma from [6] will aid us in proving the Fundamental Limit Theorem for regular Markov chains.

**Example 6.** Consider the following transition matrix for a regular Markov chain:

$$P = \begin{pmatrix} 1/3 & 1/6 & 1/2 \\ 1/4 & 1/3 & 5/12 \\ 1/2 & 1/4 & 1/4 \end{pmatrix}.$$
Suppose we have a column vector \( y = \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} \), and consider the vector \( Py \). We have

\[
Py = \begin{pmatrix} 1/3 & 1/6 & 1/2 \\ 1/4 & 1/3 & 5/12 \\ 1/2 & 1/4 & 1/4 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} 35/6 \\ 23/4 \\ 19/4 \end{pmatrix}.
\]

Since each row of \( P \) is a probability vector, the vector \( Py \) replaces \( y \) with weighted averages of its components and tends \( y \) toward a constant column vector. The components of \( Py \), from top to bottom approximated to 3 decimal places, are 5.833, 5.750, and 4.750. These components are closer to each other than those of \( y \).

The following lemma says that if an \( r \times r \) transition matrix \( P \) has entries strictly greater than zero, and \( y \) is any column vector with \( r \) components, then the vector \( Py \) has components which are “closer together” than the components are in \( y \), similar to how we showed in Example 6. The components of \( y \) are replaced with a weighted average of its previous components.

**Lemma 2.** Let \( P \) be an \( r \times r \) transition matrix with entries strictly greater than zero. Let \( d \) be the smallest entry of \( P \) and let \( y \) be a column vector with \( r \) components, the largest of which we will denote \( M_0 \) and the smallest \( m_0 \). Also, let \( M_1 \) and \( m_1 \) be the largest and smallest component, respectively, of the vector \( Py \). We have \( r \geq 2 \), so
therefore \( d \leq 1/2 \) and then \( 1 - 2d \geq 0 \). Then

\[
M_1 - m_1 \leq (1 - 2d)(M_0 - m_0).
\]

**Proof.** Recall that each entry of the vector \( Py \) is a weighted average of the entries of the vector \( y \). The largest weighted average that could be reached would occur if all but one of the entries of \( y \) had a value of \( M_0 \) and one entry had a value \( m_0 \), with \( m_0 \) weighted by the smallest possible weight, namely \( d \).

For this case, the weighted average equals \( dm_0 + (1 - d)M_0 \). Likewise, the smallest possible weighted average equals \( dM_0 + (1 - d)m_0 \). Therefore,

\[
M_1 - m_1 \leq \left( dm_0 + (1 - d)M_0 \right) - \left( dM_0 + (1 - d)m_0 \right)
= dm_0 + M_0 - dM_0 - dM_0 - m_0 + dm_0
= M_0 - m_0 - 2dM_0 + 2dm_0
= (1 - 2d)(M_0 - m_0),
\]

as desired. \( \square \)

We can now present the Fundamental Limit Theorem for regular Markov chains and proof given in [6].

**Theorem 7.** Let \( P \) be the transition matrix for a regular Markov chain. The Fundamental Limit Theorem for regular Markov chains states that

\[
\lim_{n \to \infty} P^n = W,
\]

where \( W \) is a matrix with rows equal to the fixed row vector \( \mathbf{w} \) for \( P \). Additionally, all
entries of \( W \) are strictly positive.

**Proof.** We will prove this theorem using two cases, one where \( P \) has no 0 entries and the other where \( P \) could have some 0 entries.

Let us first consider the case where \( P \) has no 0 entries. Let \( y \) be any \( r \)-component column vector, with \( r \) corresponding to the number of states in the chain. First, assume that this regular chain has more than one state, \( r > 1 \), since otherwise this theorem is trivial. Now, let \( M_n \) and \( m_n \) be the maximum and minimum components of the vector \( P^n y \), respectively. Note that \( P^n y \) is obtained by multiplying the vector \( P^{n-1} y \) on the left by \( P \). It has been shown before that multiplying a column vector by \( P \) will result in a vector with averages of its previous components. Thus, each component of \( P^n y \) is an average of the components of \( P^{n-1} y \). Thus, we have \( M_0 \geq M_1 \geq M_2 \geq \cdots \) and \( m_0 \leq m_1 \leq m_2 \leq \cdots \). Each sequence \( \{M_0, M_1, M_2, \ldots\} \) and \( \{m_0, m_1, m_2, \ldots\} \) is therefore monotone. Both sequences are also bounded since \( M_n \) and \( m_n \) are the maximum and minimum of each sequence, respectively. Thus we have

\[
m_0 \leq m_n \leq M_n \leq M_0.
\]

Since both sequences are monotone and bounded, the Monotone Convergence Theorem states that each sequence converges to a limit as \( n \) tends to infinity.

Now, let \( M \) be the limit of \( M_n \) and \( m \) the limit of \( m_n \). We know that \( m \leq M \). We want to show that \( M - m = 0 \) to prove that \( P^n y \) converges to a constant column vector. We will have \( M - m = 0 \) if \( M_n - m_n \) tends to 0 as \( n \) approaches infinity. Now, let \( d \) be the smallest element of \( P \). Since all entries of \( P \) are strictly positive, we have \( d > 0 \). By Lemma 2, we know

\[
M_n \geq d^n M_0 \quad \text{and} \quad m_n \leq d^n m_0.
\]
$M_n - m_n \leq (1 - 2d)(M_{n-1} - m_{n-1})$. We shall use induction to show that $M_n - m_n \leq (1 - 2d)^n(M_0 - m_0)$.

**Initial Case** ($n=1$):
From Lemma 2, we have shown that $M_1 - m_1 \leq (1 - 2d)(M_0 - m_0)$.

**Induction Step:**
Assume $M_k - m_k \leq (1 - 2d)^k(M_0 - m_0)$ for some $k \in \mathbb{N}$. We want to show that $M_{k+1} - m_{k+1} \leq (1 - 2d)^{k+1}(M_0 - m_0)$. So we have

\[
M_{k+1} - m_{k+1} \leq (1 - 2d)(M_k - m_k)
\]

\[
\leq (1 - 2d)(1 - 2d)^k(M_0 - m_0)
\]

\[
= (1 - 2d)^{k+1}(M_0 - m_0),
\]

as desired.

**Conclusion:**
Therefore, we know that $M_n - m_n \leq (1 - 2d)^n(M_0 - m_0)$ for all $n \in \mathbb{N}$.

Now, since $r \geq 2$, we must have $d \leq 1/2$ and thus $0 \leq 1 - 2d < 1$. Therefore, since we have that $M_n - m_n \leq (1 - 2d)^n(M_0 - m_0)$ and $0 \leq 1 - 2d < 1$, we know that $M_n - m_n$ tends to 0 as $n$ approaches infinity. Since $M_n$ and $m_n$ are the maximum and minimum components of $P^n y$, each component must approach the same number $\bar{u} = M = m$. Therefore, we have

\[
\lim_{n \to \infty} P^n y = u,
\]

where $u$ is a column vector all of whose components equal $\bar{u}$.

Now, let $y$ be a column vector whose $j$th component is equal to 1 and
remaining components are equal to 0. Then \( P^n y \) is the \( j \)th column of \( P^n \). If this process is done for \( j = 1, \ldots, r \), then we see that the columns of \( P^n \) approach constant column vectors. In other words, the rows of \( P^n \) approach a common row vector \( w \), or

\[
\lim_{n \to \infty} P^n = W,
\]

where \( W \) is a matrix with the row vector \( w \) as its rows.

The final portion of the first case is to show that every entry of \( W \) is strictly positive. Here, let \( y \) be the vector with all components equal to 0 except with \( j \)th component equal to 1. Then \( Py \) yields the \( j \)th column of \( P \), whose entries are all strictly positive since all entries of \( P \) are strictly positive. The minimum component of the vector \( Py \), labeled \( m_1 \), is thus greater than 0. Since \( m_1 \leq m \), we therefore have \( m > 0 \). Furthermore, since \( m \) is the \( j \)th component of \( W \), we have that all components of \( W \) are strictly positive.

Now consider the case where \( P \) could have some entries as 0. We know that \( P \) is the transition matrix for a regular Markov chain. Therefore, for some \( N \in \mathbb{N} \), \( P^N \) has no zeros. In the first case, we showed that the difference \( M_n - m_n \) tends to 0 as \( n \) tends to infinity. Therefore, we can say that the difference \( M_{nN} - m_{nN} \) also approaches 0 as \( n \) tends to infinity. Furthermore, we know that the difference \( M_n - m_n \) can never increase. This is because earlier in the proof we showed that \( M_n - m_n \leq (1 - 2d)(M_{n-1} - m_{n-1}) \). In this case, we have \( d \geq 0 \). If \( d = 0 \) then we have \( M_n - m_n \leq M_{n-1} - m_{n-1} \). Thus, the difference \( M_n - m_n \) never increases if \( d = 0 \), and we showed in the first case that the difference does not increase if \( d > 0 \). Hence, if we know that the differences obtained at every \( N \)th step tend to 0, then the entire sequence \( M_n - m_n \) must
also tend to 0. Therefore, \( \lim_{n \to \infty} P^n = W \) for any regular transition matrix. □

## 4.2 Examples of Ergodic Markov chains

We will now revisit the idea of fixed row and column vectors, and use the probabilities from the Cleveland weather Markov chain to find these vectors.

**Example 7.** Recall the transition matrix for the Cleveland weather Markov chain:

\[
P = \begin{pmatrix}
\frac{2}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{pmatrix}.
\]

To find the fixed row vector \( w \) for Cleveland’s weather, start with the fact that we know \( w_1 + w_2 + w_3 = 1 \), where \( w_1, w_2, \) and \( w_3 \) are the entries of the vector \( w \).

Then for \( wP = w \) we have

\[
\begin{pmatrix}
w_1 & w_2 & w_3
\end{pmatrix}
\begin{pmatrix}
\frac{2}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{pmatrix}
= (w_1, w_2, w_3).
\]

From this information we have the following four equations with three unknowns:
\begin{align*}
w_1 + w_2 + w_3 &= 1, \\
(2/3)w_1 + (1/4)w_2 + (1/6)w_3 &= w_1, \\
(1/3)w_1 + (1/2)w_2 + (1/6)w_3 &= w_2, \\
(0)w_1 + (1/4)w_2 + (2/3)w_3 &= w_3.
\end{align*}

To solve this system of equations, set \( w_2 = 1 \) and solve for the third and fourth equations. We have

\begin{align*}
(1/3)w_1 + (1/2)(1) + (1/6)w_3 &= 1, \\
(1/4)(1) + (2/3)w_3 &= w_3.
\end{align*}

From this we first see that \( w_3 = 3/4 \), which leaves us with

\[(1/3)w_1 + (1/2) + (1/6)(3/4) = 1.\]

Therefore, \( w_1 = 9/8 \). So we have \( (w_1 \quad w_2 \quad w_3) = (9/8 \quad 1 \quad 3/4) \). This vector satisfies the last three equations, but not the first equation. In order to obtain a solution to this equation, simply divide the vector by the sum of its components, which in this case is 23/8. So we have

\[w = (9/23 \quad 8/23 \quad 6/23) \approx (0.391 \quad 0.348 \quad 0.261).\]

The Fundamental Limit Theorem for regular Markov chains says that this
fixed row vector \( w \) comprises every row of the limiting matrix \( W \). Table 2.1 confirms these results by showing that each row of the limiting matrix is the vector \((0.391 \quad 0.348 \quad 0.261)\).

The following theorem from [6] formally defines fixed row and column vectors and states a property about them that will be useful.

**Theorem 8.** Let \( P \) be a regular transition matrix. By the Fundamental Limit Theorem we have \( \lim_{n \to \infty} P^n = W \). Let \( w \) be a fixed row vector for \( P \) and let \( c \) be a column vector all of whose components are 1. Then

1. \( wP = w \), and any row vector \( v \) such that \( vP = v \) is a constant multiple of \( w \).

2. \( Pc = c \), and any column vector \( x \) such that \( Px = x \) is a multiple of \( c \).

The column vector \( x \) is a fixed column vector for \( P \).

**Proof.** For the proof of 1., first recall that the Fundamental Limit Theorem states \( \lim_{n \to \infty} P^n = W \). Consider \( P^{n+1} \). We have

\[ P^{n+1} = P^n P \to WP. \]

But \( P^{n+1} \to W \), so \( W = WP \). Thus \( w = wP \).

Let \( v \) be any vector with \( vP = v \). Then we have \( v = vP^n \). Taking the limit as \( n \) approaches infinity results in \( v = vW \). Let \( h \) be the sum of the components of \( v \). Then we see that \( vW = hw \). Thus, \( v = hw \) and is therefore a constant multiple of \( w \).

For the proof of part 2., assume that \( x = Px \). Then \( x = P^n x \). Taking the limit as \( n \) approaches infinity, we have \( x = Wx \). Since we know that all of the rows of
CHAPTER 4. ERGODIC MARKOV CHAINS

$W$ are the same, the components of $Wx$ are all equal. Thus, $x$ is a multiple of the column vector $c$. □

From this theorem, we also obtain the result that there exists only one probability vector $u$ such that $uP = u$.

Another method to find the fixed row vector of a transition matrix is to use eigenvalues and eigenvectors. It was mentioned before that a fixed row vector $w$ is a left eigenvector of the transition matrix $P$ corresponding to the eigenvalue 1. By the definition of a fixed row vector, we have $wP = w$. We can extend this equation to $wP = wI$, where $I$ is the identity matrix. From here, we can say $w(P - I) = 0$. Hence, $w$ is in the left null space of the matrix $P - I$. We can also use Maple to solve for eigenvalues and eigenvectors of a matrix. From the Cleveland weather matrix in Example 1, the eigenvector corresponding to the eigenvalue 1 is the vector

$$(3/2 \quad 4/3 \quad 1).$$

Dividing by the sum of this vector’s components so that its sum becomes 1 results in the vector

$$(9/23 \quad 8/23 \quad 6/23) \approx (0.391 \quad 0.348 \quad 0.261),$$

which is exactly the same fixed row vector that we found in Example 7. These computations can be found in Figure A.3 in Appendix A.

Up to this point we have assumed that a Markov chain starts in a specific state. The following theorem from [6] considers the case where the starting
state is determined by the probability vector itself.

**Theorem 9.** Let $P$ be the transition matrix for a regular Markov chain and let $v$ be an arbitrary probability vector. Then

$$\lim_{n \to \infty} v P^n = w,$$

where $w$ is the unique fixed probability vector for $P$.

**Proof.** By the Fundamental Limit Theorem for regular Markov chains, we know that $\lim_{n \to \infty} P^n = W$. Multiplying by $v$ on the left of both sides yields

$$v \left( \lim_{n \to \infty} P^n \right) = v W.$$

Since the entries of $v$ sum to 1 and each row of $W$ equals the vector $w$, we have $vW = w$. Thus, we can say that

$$\lim_{n \to \infty} v P^n = w.$$

What this theorem tells us is that given any starting probability vector $v$, the probability vector $v P^n$ gives the probabilities of being in any of the states after $n$ steps. This also tells us that in a more general case, the probability of being in state $s_i$ approaches the $i$th entry of the vector $w$ during the Markov chain process.

The final part of this chapter on ergodic Markov chains is on an example of an ergodic chain and its fixed row vector found in [6].
Example 8. Suppose a white rat is put into the maze shown in Figure 4.1. There exist nine compartments with connections between the compartments as seen in the figure. The rat moves through the maze at random. So for example, if the rat is in compartment 7, then he has a probability of $1/2$ of entering compartment 6 and a probability of $1/2$ of entering compartment 8.

![Figure 4.1: The maze problem](image)

We can represent the movements of the rat by a Markov chain process with a transition matrix given as

$$
P = \begin{pmatrix}
0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\
1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \\
0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\
1/3 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\
0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0
\end{pmatrix}.
$$
This chain is ergodic because the rat can reach every state from any state, but it is not regular. This is because from every odd-numbered state, the process can go only to an even-numbered state. Similarly, every even-numbered state can go only to an odd-numbered state.

To find the fixed probability vector for this matrix, a system of ten equations with nine unknowns would need to be solved. Rather than solve this system of equations, we can make a guess that the times spent in each compartment by the rat would be proportional to the number of entries to each compartment. So, we try the vector whose $i$th component is the number of entries to the $i$th compartment:

$$
\mathbf{w} = (2 \ 3 \ 2 \ 3 \ 4 \ 3 \ 2 \ 3 \ 2).
$$

Using Maple to solve $\mathbf{w} P$ shows that $\mathbf{w} P = \mathbf{w}$ and therefore $\mathbf{w}$, normalized to have a sum of 1, is

$$
\mathbf{w} = (1/12 \ 1/8 \ 1/12 \ 1/8 \ 1/6 \ 1/8 \ 1/12 \ 1/8 \ 1/12).
$$

This computation can be found in Figure A.4 in Appendix A. Also, $\mathbf{w}$ is a fixed row vector for $P$. So for example, regardless of where the rat starts in the maze, he has a probability of 1/6 of being in the fifth compartment after many steps in the Markov chain.
Chapter 5

Baseball as a Markov chain

The theory of Markov chains has applications spread out to many different academic disciplines, such as physics, biology, statistics, and the social sciences. Markov chains have also been applied to many popular board games, including Risk and Monopoly. The one area that we will be investigating further is the sport of baseball and using a Markov chain to model various aspects of the game. We will first outline how baseball can be modeled using a Markov chain, followed by analyzing some strategies of the game using the created model. The entries for the transition matrix that will be used in the model are probabilities taken from the statistics of The College of Wooster Fighting Scots 2013 baseball season. Our goal is to determine how many runs The College of Wooster baseball team was expected to score during the 2013 season using Markov chains. But before we can calculate this, the Markov chain model to analyze baseball must be presented.
5.1 Setting up the Model

The following Markov chain model presented is based off of the model used in [4]. At any point during an inning of a baseball game, a team currently on offense finds itself in one of 24 possible states. There are 8 possibilities for the distribution of runners on the bases: bases empty, one man on first, second, or third, men on first and second, men on first and third, men on second and third, and men on first, second, and third. Also, there can be zero, one, or two outs at any point during an inning. Thus, we have 24 states when considering the occupied bases and the number of outs. Furthermore, we need to account for when the half inning ends with three outs. For analysis that will be shown later, we will denote the $25^{th}$ state as being three outs and zero runners left on base when the inning ends. Likewise, the $26^{th}$, $27^{th}$, and $28^{th}$ states will denote the situations in which the inning ends with one, two, and three runners left on base, respectively.

To construct a Markov chain, we need the process to satisfy the Markov property (see equation (2.1)), the number of states needs to be finite, and we need a transition matrix with positive entries and rows sums equal to 1. We can certainly create a Markov chain for an inning of a baseball game satisfying these criteria. We have 28 states in the state space described before and can create a transition matrix determining how the process will move from state to state. We can also implement the Markov property, saying that the probability of moving to the next state in an inning of a baseball game is determined only by the current state the batter is in, and not how the team got to that state. For example, if a batter is facing a situation with runners on first and second base
with one out, the probability of the batter hitting a single is the same as if he
were batting with no one on base and zero outs. The interesting concept with
applying the Markov property is that it does not take into consideration
whether a team has momentum in an inning, or whether the pitcher is
becoming rattled by the number of runners reaching base safely, for example.
If there are runners on first and second base with one out, the Markov chain
model is not concerned with whether there were two singles followed by a fly
out, a single followed by a strikeout followed by a walk, etc. [11]. A Markov
chain simply uses probabilities of moving from one state to another and is
memoryless of how it reached the current state.

The Markov chain that we will be constructing is an absorbing chain, with
the states with zero, one, or two outs being the transient states, and the states
with three outs being the absorbing states. The listing of each transient state
with its corresponding number of bases occupied and number of outs in the
inning are shown in Table 5.1. As we have done before, we will put the
transition matrix into canonical form, with the transient states listed first
followed by the absorbing states. So we have the first 24 states as states with
less than three outs, and the last 4 states comprising the absorbing states each
with an associated number of runners left on base when the third out is
reached. In canonical form, $Q$ is a $24 \times 24$ matrix, $R$ is a $24 \times 4$ nonzero matrix,
$0$ is a $4 \times 24$ zero matrix, and $I_4$ is a $4 \times 4$ identity matrix.
CHAPTER 5. BASEBALL AS A MARKOV CHAIN

Table 5.1: The transient states of the baseball Markov chain

<table>
<thead>
<tr>
<th>State</th>
<th>Bases Occupied</th>
<th>Number of Outs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1st</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>2nd</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>3rd</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1st and 2nd</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1st and 3rd</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>2nd and 3rd</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1st, 2nd, and 3rd</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>None</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1st</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2nd</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>3rd</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1st and 2nd</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1st and 3rd</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>2nd and 3rd</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>1st, 2nd, and 3rd</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>None</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>1st</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>2nd</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>3rd</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>1st and 2nd</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>1st and 3rd</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>2nd and 3rd</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>1st, 2nd, and 3rd</td>
<td>2</td>
</tr>
</tbody>
</table>

5.2 Baseball Transition Probabilities

Once we have the transition matrix into canonical form, we can then begin to enter the transition probabilities for the baseball Markov chain into the matrix. Our first goal in the analysis of this Markov chain is to determine the expected number of runs that The 2013 College of Wooster baseball team would have scored during that year. We will do this using statistics from the 2013 season.
In order to accomplish this, we need to gather statistics on baseball events that occurred during the season and determine the probabilities of moving from one state to another by means of these events. The following method of determining which baseball events to use in the model follows that in [4].

The baseball events that we will use to establish transition probabilities are limited to singles, walks, hit batsmen, doubles, triples, home runs, plays in which only one out is made, and plays in which exactly two outs are made (also known as double plays). The first assumption we are making when creating this model is that after each transition, the current batter changes to the next batter in the lineup [11]. That is, the current batter either reaches base safely or makes an out. Therefore, we are not considering plays such as stealing a base or advancement on a passed ball. In addition, we are not considering triple plays, sacrifice flies, errors, plays where a runner advances from 1st base to 3rd base on a single, and other plays that are so rare and do not occur very often in a baseball game. We are making these assumptions in order to simplify the model, and, more importantly, we simply do not have access to most of the probabilities of these events occurring. Since the incidence of these events are fairly uncommon during the course of a baseball season, we can assume that they do not occur for the formulation of the model, and we will still have a fairly accurate model for run production. The following list describes other assumptions made for each baseball event that we are using in the model:

- **Single**: A base runner on first base moves to second base, other base runners score, and batter moves to first base.
• **Walk or Hit Batsman**: Runners advance one base only if forced (e.g. runner on first base will advance to second base, but runner only on second base will not advance), and batter moves to first base

• **Double**: A runner on first base advances to third base, other runners score, and batter moves to second base

• **Triple**: All runners score and batter moves to third base

• **Home Run**: All runners and batter score

• **Out**: Runners do not advance and the number of outs increases by one

• **Double Play**: Can only occur when there is a force out at second base. If there is more than one base runner and a force at second base exists, only the runner on first base and batter are out, and other runners advance one base or score from third if play does not result in three outs

The data needed to calculate the transition probabilities for this Markov chain are taken from The College of Wooster baseball’s home page under 2013 Statistics - Standard \[10\]. By using the entire team’s statistics, we are modeling and calculating the expected number of runs that the team would score for the season. This method of using the entire team’s statistics is useful and will yield fairly accurate results compared to the actual number of runs scored by the team. However, it does not exactly simulate how a baseball game actually works because every batter is identical in this model with exactly the same transition probabilities. For this analysis we will only consider 9 identical batters with transition probabilities based on the entire team’s performance during the season.
In a later section of this chapter we will use only one individual player’s statistics from the season and compute the transition probabilities for that player. With this information, we can run the model as if every batter were identical to him and calculate that player’s own expected number of runs scored for the season. These results will give us how many runs each individual player is worth to the team.

Before we can calculate expected number of runs, we first need to determine the transition probabilities for the entire team during the 2013 season. We must collect the number of times each of the following baseball events occurred: singles, walks, hit batsmen, doubles, triples, home runs, single outs, and double plays. In order to calculate the probability of each event occurring, we must collect the number of plate appearances that the team had and divide each event by this number to determine each event’s probability of occurring. The number of plate appearances was found to be 1,799 for the team throughout the season. Therefore, the number of occurrences for each event will be divided by 1,799 to find the each event’s probability of occurring. Table 5.2 displays the baseball events with the number of times each occurred and its probability of occurring.

One note about Table 5.2 is that walks and hit batsmen are considered under the same event (Walk or Hit Batsman) since both outcomes result in the same state after occurring. Also, the probability for the event Single or Walk or Hit Batsman is needed for some situations during a baseball game. For example, to transition from the state with no runners on base and zero outs to the state with a runner on first base with zero outs, we would need the probability of the batter getting a single, walk, or hit batsman. Furthermore,
when constructing the transition matrix, we need to keep in mind that each row of probabilities must sum to 1. So in situations where a double play can occur from a particular state, we assume that the probability of obtaining exactly one out from that state is the probability of a single out minus the probability of a double play. Here, the probability would be $0.590 - 0.014 = 0.576$. Now that we have the probability of each event occurring, we can fill in the transition matrix with the transition probabilities between each state following the list of states given in Table 5.1.

Additionally, Table B.1 in Appendix B shows the eighth row of each power of the transition matrix up to the eighth power. In other words, the probabilities shown in the table are transitions from the 8th state to any of the 28 states after one step and up to eight steps. When looking at this table, the 0.000 entry in the 10th state under the column $P$, for example, means that there is no probability of transitioning from state 8 to state 10 after one step. However, in this same row there is a 0.032 probability under the column $P^3$. This means that transitioning from state 8 to state 10 after three steps has a
probability of 0.032. Another piece of information to notice from this table is that the transition probabilities of the transient states are approaching 0 as the power of the transition matrix increases, just as Theorem 3 in Section 3.3 indicates. For example, notice that the probabilities in row 1 are decreasing after each increase in the power of the transition matrix. This means that the probability of transitioning from state 8 to state 1 after more and more steps is decreasing and the probability of the process being in one of the absorbing states is increasing.

5.3 Calculating Expected Number of Runs

The model that we present for calculating the expected number of runs for The 2013 College of Wooster baseball team for one inning in length follows that given in [12]. In any inning of a baseball game, it is true that every batter that comes to the plate either makes an out, scores a run, or is left on base at the end of the inning. Let $B$ be the number of batters who come to the plate in an inning, let $R$ be the number of runs scored, and let $L$ be the number of runners left on base at the end of the inning. Then, $B = 3 + R + L$. This can be equivalently written as $R = B - L - 3$. Taking the expected value of this equation, we are then left with $E(R) = E(B) - E(L) - 3$. Therefore, if we find the expected number of batters that appear in an inning and the expected number of runners left on base during that inning, we can find the expected number of runs for an inning.

Recall the fundamental matrix $N$ of an absorbing Markov chain discussed in section 3.4. The fundamental matrix $N$ for this Markov chain is a $24 \times 24$
matrix with entries $n_{ij}$ representing the number of times the chain is in state $s_j$, given that it started in state $s_i$. Also, we know that the expected number of steps from a transient state $s_i$ until absorption is the sum of the $i$th row’s entries of the matrix $N$. What this means in terms of baseball is that the sum of the $i$th row of $N$ is the expected number of batters that will come to bat during the remainder of the inning starting from state $s_i$. Since we are looking for the expected number of batters to come to the plate starting from the beginning of an inning, we only need to look at the sum of the first row of the fundamental matrix $N$. Therefore, $E(B)$ is the sum of the first row of $N$. In order to calculate and analyze $N$, we used the mathematical software MatLab in addition to Microsoft Excel. Once the data was compiled, we computed the matrix $N$ in Matlab and its row sums in Microsoft Excel which can be seen in Figure B.1 in Appendix B. It was found that $E(B) = 5.051$.

In order to calculate the expected number of runners left on base during an inning, we must use the matrix $B$ explained in Section 3.6. Recall that we have $B = NR$, where $N$ is the fundamental matrix and $R$ is an in canonical form of the transition matrix. Also, recall that the entries $b_{ij}$ give the probability that an absorbing Markov chain will be absorbed in an absorbing state $s_j$, given that the chain starts in the transient state $s_i$. For our baseball model, the $i$th row of the matrix $B$ contains the probabilities of the chain being absorbed with zero, one, two, and three runners left on base, starting from state $s_i$. Therefore, the first row of $B$ will be used to calculate the expected number of runners left on base starting from the beginning of the inning. The Excel worksheet showing the entries of the first row of $B$ can be found in Figure B.2 in Appendix B. Each part of the formula to calculate $E(L)$ will need to be appropriately weighted to
determine the expected number of runners left on base. The formula is

\[ E(L) = 0 \cdot P(0 \ left) + 1 \cdot P(1 \ left) + 2 \cdot P(2 \ left) + 3 \cdot P(3 \ left), \]

where \( P(0 \ left) \) represents the probability of ending the inning with no runners left on base, \( P(1 \ left) \) represents the probability of ending the inning with 1 runner left on base, and so on.

After calculating \( B \) in MatLab, the first row’s entries were found to be 0.240, 0.335, 0.345, and 0.081, rounded to the nearest thousandths place.

Applying the formula to find the expected number of runners left on base, we have that \( E(L) = 1.267 \). Thus, we have

\[ E(R) = E(B) - E(L) - 3 = 5.051 - 1.267 - 3 = 0.784. \]

According to this model, the team would be expected to score 0.784 runs every inning. During the 2013 season, The College of Wooster baseball team scored 364 runs in a total of 375 innings of play [10]. At an average of 0.784 runs per inning, the model estimates that the team would score a little more than 294 runs in 375 innings. This estimation is much lower than the actual number of runs scored likely due to the assumptions made in formulating the model. We did not consider stolen bases, sacrifice flies, and moving from first base to third base on a single, for example. Had these events been included, the results for expected number of runs would be closer to the actual number of runs scored. However, these assumptions were needed to be made due to the limitations in the statistics given for The College of Wooster baseball team.
5.4 Player Value

In the previous section we calculated the expected number of runs that would be scored for The College of Wooster baseball team during the 2013 season using the team’s statistics from that season. In this model we assumed that every batter was identical with the exact same transition probabilities. One interesting extension that we can make from this is that if we were to use one individual player’s statistics and formulate his transition probabilities, we can run the model and determine his expected number of runs scored if he batted all the time. The results will essentially tell us what each player’s run production value is to the team.

The batters that we will use in the model to determine player value are the batters that recorded the most at bats during the 2013 season. The method we will use to determine each individual player’s run expectancy is the same as the method used to calculate the team’s expected number of runs scored for the season. The transition probabilities for each player were computed using the 2013 season statistics found in [10]. As before, we will determine the expected number of runs both for one inning of play and for a full season of 375 innings. Recall that we have \( E(R) = E(B) - E(L) - 3 \). Therefore we must find \( E(B) \) and \( E(L) \) before we can determine \( E(R) \). Table 5.3 displays the 9 batters that were used in the model and the following data for each batter: the expected number of batters in one inning, the expected number of runners left on base after one inning, the expected number of runs during that inning, and the expected number of runs for a season of 375 innings.
Table 5.3: Individual batter’s expected number of runs scored

<table>
<thead>
<tr>
<th>Batter</th>
<th>E(B)</th>
<th>E(L)</th>
<th>E(R)</th>
<th>Expected Runs per Season</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Mancine</td>
<td>5.8257</td>
<td>1.5483</td>
<td>1.2774</td>
<td>479.025</td>
</tr>
<tr>
<td>J. McLain</td>
<td>5.1235</td>
<td>1.2846</td>
<td>0.8389</td>
<td>314.588</td>
</tr>
<tr>
<td>E. Reese</td>
<td>5.4759</td>
<td>1.2224</td>
<td>1.2535</td>
<td>470.063</td>
</tr>
<tr>
<td>Z. Mathie</td>
<td>5.1813</td>
<td>1.2467</td>
<td>0.9346</td>
<td>350.475</td>
</tr>
<tr>
<td>F. Vance</td>
<td>4.8994</td>
<td>1.2033</td>
<td>0.6961</td>
<td>261.038</td>
</tr>
<tr>
<td>C. Thomay</td>
<td>5.4081</td>
<td>1.4062</td>
<td>1.0019</td>
<td>375.713</td>
</tr>
<tr>
<td>C. Day</td>
<td>4.8082</td>
<td>1.1679</td>
<td>0.6403</td>
<td>240.113</td>
</tr>
<tr>
<td>R. Miner</td>
<td>4.6225</td>
<td>1.1508</td>
<td>0.4717</td>
<td>176.888</td>
</tr>
<tr>
<td>B. Miller</td>
<td>5.2032</td>
<td>1.3220</td>
<td>0.8812</td>
<td>330.450</td>
</tr>
</tbody>
</table>

Recall that the model estimated that the team would score 294 runs in 375 innings of play during the season. The interesting part of this analysis is that 6 of the 9 batters have a run expectancy higher than the team’s run expectancy, with two of them resulting in over 150 more expected runs than that modeled for the team during the season.

5.5 Using Markov Chains for Baseball Strategy

In recent years, there has been an increase in the amount of research done analyzing strategies within baseball games. Two such strategies that we will look into further are the sacrifice bunt and stolen base attempt. Baseball managers utilize both strategies in order to give their team the best chance to win the game. We will focus our analysis first on the sacrifice bunt.
5.5.1 Sacrifice Bunt Strategy

A sacrifice bunt is used most often when there is a runner on first base and no outs in order to move the runner to second base for the cost of one out. Managers make use of this strategy to move their runner into scoring position so that they have a chance to score from a hit by one of the following batters. However, there have been times in the past where managers have essentially “killed a rally” by calling for a sacrifice bunt and giving up an out for an extra base [13]. So, is it worth sacrificing an out for an advancement of one base for a runner? There is much disagreement on the answer to this question, but we will use mathematics to make a determination on whether it is a good decision to utilize the sacrifice bunt, and if so, when to employ it.

In order to win baseball games, teams need to score runs. Therefore, a manager’s objective is to give his team the best chance to score runs. To do this, managers use every possible strategic move to maximize their teams’ opportunities to score as many runs as they can. From the lineups that they create to the decisions to sacrifice bunt or attempt to steal a base, managers are trying to score runs to win baseball games. In section 5.3, we calculated the expected number of runs that The College of Wooster baseball team would score during the 2013 season. Using the same matrices $N$ and $B$, we can use Markov chains to determine the expected number of runs that the team would score from any of the 24 transient states of a baseball game. For example, to find the expected number of batters that come to bat starting from the second state, we would sum the second row of $N$ for the answer. Then to find the expected number of runners left on base from the second state we would use
the numbers found in the second row of the matrix $B$. Table 5.4 shows the run expectancies from each state until the end of the inning.

Table 5.4: Run expectancies for each of the 24 states

<table>
<thead>
<tr>
<th>Number of Outs</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bases Occupied</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xxx</td>
<td>0.784</td>
<td>0.413</td>
<td>0.145</td>
</tr>
<tr>
<td>1xx</td>
<td><strong>1.265</strong></td>
<td>0.736</td>
<td>0.297</td>
</tr>
<tr>
<td>x2x</td>
<td>1.480</td>
<td><strong>0.959</strong></td>
<td>0.471</td>
</tr>
<tr>
<td>xx3</td>
<td>1.481</td>
<td>0.959</td>
<td>0.471</td>
</tr>
<tr>
<td>12x</td>
<td>1.975</td>
<td>1.294</td>
<td>0.632</td>
</tr>
<tr>
<td>1x3</td>
<td>1.984</td>
<td>1.294</td>
<td>0.632</td>
</tr>
<tr>
<td>x23</td>
<td>2.196</td>
<td>1.521</td>
<td>0.806</td>
</tr>
<tr>
<td>123</td>
<td>2.771</td>
<td>1.940</td>
<td>1.042</td>
</tr>
</tbody>
</table>

For the Bases Occupied column, an “x” represents an empty base. For example, “12x” represents the states consisting of runners on first and second base with no runner on third base. Looking at this table we see that the team has a run expectancy of 1.265 runs from the state with a runner on first base with no outs. If the team were to use the sacrifice bunt from this state, they would find themselves in the state with a runner on second base with 1 out for a run expectancy of 0.959 runs from this state. From this perspective, managers lose 0.306 expected runs for the inning by calling for a sacrifice bunt. This loss in expected runs may not seem like a lot, but over the course of a season it could result in a considerable amount of runs left on the table.

One other noteworthy piece of information from Table 5.4 is that there is virtually no difference in the run expectancies between the states with a runner on second only and a runner on third only, as well as between the
states with runners on first and second base and the states with runners on first and third base. This is due to the fact that sacrifice flies are not included in the model and our assumption that a single will score both a runner from second base and from third base.

From analyzing Table 5.4, it is clear that managers should not employ the sacrifice bunt strategy if their goal is to maximize their team’s number of expected runs during an inning. Giving the defense a free out is too costly for a team trying to maximize the amount of runs scored during an inning.

Early on during a baseball game, a manager’s goal is to maximize his team’s expected number of runs, which will give his team the best chance to win the game at that point. We have determined that a manager should not call for a sacrifice bunt early on during a game. There may, however, be a time in which the sacrifice bunt would be a good strategy to use.

During a baseball game, there are times in which it would be necessary to maximize a team’s probability of scoring at least one run as opposed to maximizing their expected number of runs. For example, late in a game (7th inning or later) a team may be down one run and needs to score at least one more run to avoid losing the game. Instead of attempting to maximize their expected number of runs during that inning, it would be more beneficial to maximize the probability of getting at least one run across the board. As we showed before, the expected number of runs scored for a team decreases after a successful sacrifice bunt. However, the act of moving this runner into scoring position will likely increase the probability of the team scoring at least one run during that inning. Table 5.6 shows the probability of scoring at least 1 run with and without a bunt attempt from the state with a runner on first base and
no outs. We have also included the probabilities for Major League Baseball players using the statistics from the 2013 season found from the website FanGraphs.com [1]. The reason for including MLB statistics is so that we can make comparisons between Wooster’s team and Major League Baseball, and ultimately make better conclusions about sacrifice bunt strategies. Table 5.5 shows the number of occurrences of each baseball event during the 2013 Major League Baseball season found from [1], along with each event’s probability of occurring. These probabilities are needed to create a transition matrix with MLB data, which we will use in this sacrifice bunt attempt analysis.

Table 5.5: Transition probabilities for Major League Baseball in 2013

<table>
<thead>
<tr>
<th>Events</th>
<th>Number of Occurrences</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>28,438</td>
<td>0.154</td>
</tr>
<tr>
<td>Walk or Hit Batsman</td>
<td>16,176</td>
<td>0.087</td>
</tr>
<tr>
<td>Single or Walk or Hit Batsman</td>
<td>44,614</td>
<td>0.241</td>
</tr>
<tr>
<td>Double</td>
<td>8,222</td>
<td>0.044</td>
</tr>
<tr>
<td>Triple</td>
<td>772</td>
<td>0.004</td>
</tr>
<tr>
<td>Home Run</td>
<td>4,661</td>
<td>0.025</td>
</tr>
<tr>
<td>Single Out</td>
<td>126,603</td>
<td>0.685</td>
</tr>
<tr>
<td>Double Play</td>
<td>3,739</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Table 5.6: Probability of scoring at least one run with and without a bunt attempt starting from the state with a runner on first base and no outs

<table>
<thead>
<tr>
<th></th>
<th>Bunt Attempt</th>
<th>No Bunt Attempt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wooster</td>
<td>0.476</td>
<td>0.486</td>
</tr>
<tr>
<td>MLB</td>
<td>0.370</td>
<td>0.348</td>
</tr>
</tbody>
</table>

To compute these probabilities, we first need to find the probabilities of scoring at least one run given a successful bunt and an unsuccessful bunt. Let
\( R_n \) be the event that \( n \) or more runs scored, let \( A \) be the event of an attempt to sacrifice bunt, and let \( S \) be the event of a successful sacrifice bunt. The formula for finding the probability of scoring at least one run given a sacrifice bunt attempt is

\[
P(R_1|A) = P(R_1|S)(P(S)) + P(R_1|S^C)(P(S^C)),
\]

where \( S^C \) is the complement of the event \( S \).

We see here that both \( P(R_1|S) \) and \( P(R_1|S^C) \) must be weighted by the probability of a successful bunt and the probability of an unsuccessful bunt, respectively. To find \( P(R_1|S) \), we can find the probability of scoring no runs given a successful bunt attempt and then subtract it from 1. So, the probability of scoring no runs from a successful bunt attempt is \( p^{(2)}_{11,26} + p^{(3)}_{11,27} + p^{(4)}_{11,28} \). Recall that \( p^{(n)}_{ij} \) is the \( ij \)th entry of the baseball transition matrix \( P^n \). Thus, for example, \( p^{(2)}_{11,26} \) is the (11,26) entry of the transition matrix \( P^2 \) and is the probability of starting from the state with a runner on second base and one out and ending the inning with one runner left on base after exactly 2 batters. Therefore,

\[
P(R_1|S) = 1 - (p^{(2)}_{11,26} + p^{(3)}_{11,27} + p^{(4)}_{11,28}) = 0.546.
\]

From [13], a typical batter advances a runner on a sacrifice bunt 69\% of the time. Therefore, the probability of scoring at least one run given a successful bunt attempt and weighted by the average success rate of a sacrifice bunt is

\[
P(R_1|S) = 0.546 \times 0.69 = 0.377.
\]

Following the same method, we have that

\[
P(R_1|S^C) = 1 - (p^{(1)}_{10,25} + p^{(2)}_{10,25} + p^{(2)}_{10,26} + p^{(3)}_{10,26} + p^{(3)}_{10,27} + p^{(4)}_{10,28}) = 0.319.
\]
Knowing that $P(S^C) = 1 - P(S)$, we have

$$P(R_1|S^C)(P(S^C)) = 0.319(0.31) = 0.099.$$ 

Thus, $P(R_1|A) = 0.377 + 0.099 = 0.476$.

To find the probability of scoring at least one run without a sacrifice bunt attempt, we simply need to compute

$$P(R_1|A^C) = 1 - (p^{(2)}_{2,25} + p^{(3)}_{2,26} + p^{(4)}_{2,27} + p^{(5)}_{2,28}) = 0.486.$$ 

The other probabilities found in Table 5.6 are computed in a similar fashion and the Excel worksheets can be found in Figures B.3 and B.4 in Appendix B.

The interesting information given in Table 5.6 is that for Wooster’s team, the probability of scoring at least one run after a bunt attempt is lower than without a bunt attempt. This result is opposite of what we would expect to occur. One reason for this could be due to a low sample size of data for Wooster’s team during the 2013 season. This is another reason why we collected data from the 2013 Major League Baseball season. Given that there were 184,872 total plate appearances in the MLB compared to 1,799 plate appearance for The College of Wooster, using the MLB’s statistics should give us adequate results to analyze.

From Table 5.6 we see that with a great enough sample size, a sacrifice bunt does indeed give a team a greater probability of scoring at least one run during the inning. Therefore, if a team is down one run late in the game with a runner on first base and no outs, the Markov chain model says that a sacrifice bunt
will increase the team’s probability of scoring that one run to tie the game.

5.5.2 Stolen Base Strategy

The next piece of baseball strategy analysis that we will investigate is on the stolen base attempt. Utilizing the attempt to steal a base can be a useful strategy for managers to give their team the best chance to win a game. Similar to our sacrifice bunt analysis, we can explore the probabilities of scoring at least one run with and without at stolen base attempt from various states of a baseball game. Also, we can determine how successful a base stealer must be to make it worth attempting to steal a base. For this analysis, we will only consider attempting to steal second base.

To determine the probability of scoring at least one run with and without a stolen base attempt, we follow the same process as we did for sacrifice bunt attempts. We must first compute the probabilities of scoring at least one run given a successful steal and an unsuccessful steal and then weight each by the probability of a successful steal and an unsuccessful steal, respectively. We compute these probabilities using the statistics from [10] for The College of Wooster and from [1] for Major League Baseball. The probability of a successful stolen base for Wooster was 0.718 and the probability for the MLB was 0.728. Table 5.7 displays the computed probabilities given a stolen base attempt and Table 5.8 displays the probabilities given no stolen base attempt. The Excel worksheets used to find these numbers can be found in Figures B.5 and B.6 in Appendix B.

From this information, we can see that the probability of scoring at least
Table 5.7: Probability of scoring at least one run with a stolen base attempt

<table>
<thead>
<tr>
<th>Situation</th>
<th>Wooster</th>
<th>MLB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man on First, No Outs</td>
<td>0.558</td>
<td>0.460</td>
</tr>
<tr>
<td>Man on First, 1 Out</td>
<td>0.416</td>
<td>0.335</td>
</tr>
<tr>
<td>Man on First, 2 Outs</td>
<td>0.234</td>
<td>0.182</td>
</tr>
</tbody>
</table>

Table 5.8: Probability of scoring at least one run without a stolen base attempt

<table>
<thead>
<tr>
<th>Situation</th>
<th>Wooster</th>
<th>MLB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man on First, No Outs</td>
<td>0.486</td>
<td>0.348</td>
</tr>
<tr>
<td>Man on First, 1 Out</td>
<td>0.326</td>
<td>0.224</td>
</tr>
<tr>
<td>Man on First, 2 Outs</td>
<td>0.151</td>
<td>0.102</td>
</tr>
</tbody>
</table>

one run given an attempt to steal second base is greater than without an attempt to steal the base. So if it is late in the game, a team is down one run, and there is a runner on first base whose probability of a successful stolen base is at or higher than the team average, then the manager will increase his team’s probability of scoring one run by calling for the runner to attempt to steal second base.

There are times, however, when a manager should not give the steal sign to his runner on first base. For example, if the runner is not a very quick runner, then the probability of a successful steal will not be very high and the runner will likely be thrown out at second base, lowering the team’s chances of scoring a run in the inning. So, how successful must a runner be in stealing a base to make it worth attempting to steal? We can answer this question in two different ways. We can either seek an increase in the number of expected runs from a stolen base attempt, or an increase in the probability of scoring at
least one run from a stolen base attempt, depending on the situation of the
game. As we said before, a team’s goals early in the game are to maximize
their number of expected runs, and late in the game possibly to maximize their
probability of scoring at least one run.

Let $E_t$ be the expected number of runs scored from state $t$, let $R_n$ be the
event that $n$ or more runs scored, let $A$ be the event of an attempt to steal a
base, and let $S$ be the event of a successful stolen base. From an expected
number of runs standpoint, we would want the expected number of runs from
a stolen base attempt to be greater than the expected number of runs without a
stolen base attempt. For example, the formula for this when starting from the
state with a runner on first base and no outs is

$$E_3(P(S)) + E_9(1 - P(S)) > E_2.$$ 

The expected number of runs from state 3 and state 9 must be weighted by the
probability of a successful steal and an unsuccessful steal, respectively. From
Table 5.4 and the listing of the states in Table 5.1, we see that $E_2 = 1.265$,
$E_3 = 1.480$, and $E_9 = 0.413$. Solving for $P(S)$, we have that $P(S) > 0.7985$.
Therefore, a base stealer must be at least 79.85% successful to increase the
team’s expected number of runs scored given a stolen base attempt. The
results for when starting with one or two outs are also shown in Table 5.9.

We can also determine how successful a runner must be in stealing second
base from the standpoint of maximizing the probability of scoring at least one
run. Suppose a team is down one run in the 8th inning and they have a runner
on first base with no outs. We can determine how successful the runner needs
to be in stealing second base so that the probability of his team scoring at least one run from an attempt to steal is higher than the probability of his team scoring at least one run without an attempt to steal. The formula for this situation is

\[ P(R_1|S)(P(S)) + P(R_1|S^C)(P(S^C)) > P(R_1|A^C), \]

where \( A^C \) and \( S^C \) are the complements of the events \( A \) and \( S \). Let’s break this down into the three pieces of the inequality:

1. The probability of scoring at least 1 run given a successful steal weighted by the probability of a successful steal

   plus

2. The probability of scoring at least one run given an unsuccessful steal weighted by the probability of an unsuccessful steal

   must be greater than

3. The probability of scoring at least one run from no attempt to steal a base.

We can find these probabilities using the same technique as we did in analyzing sacrifice bunt strategies. First we have

\[ P(R_1|S) = 1 - (p_{3,26}^{(3)} + p_{3,27}^{(4)} + p_{3,28}^{(5)}) = 0.695. \]

During the 2013 season, Wooster had a stolen base success rate of 71.8\% [10].

Thus, the first piece of the inequality is \((0.695)(0.718) = 0.499\).

To compute \( P(R_1|S^C) \) we follow the same pattern as in computing \( P(R_1|S) \).
The probability of scoring at least one run from an unsuccessful steal is

\[ P(R_1|S^c) = 1 - (p_{9,26}^{(2)} + p_{9,26}^{(3)} + p_{9,27}^{(4)} + p_{9,28}^{(5)}) = 0.209. \]

So the second piece of the inequality is \((0.209)(0.282) = 0.059\). Lastly, the probability of scoring at least one run without an attempt to steal is

\[ P(R_1|A^c) = 1 - (p_{2,25}^{(2)} + p_{2,26}^{(3)} + p_{2,27}^{(4)} + p_{2,28}^{(5)}) = 0.486. \]

Now that we have all of the pieces we can solve for \(P(S)\). Plugging these numbers into the inequality results in \(P(S) > 0.5704\). Therefore, a base stealer must be at least 57.04% successful when stealing second base with no outs in order to increase the team’s probability of scoring at least one run given the attempt to steal second base. Table 5.9 shows how successful a runner must be from three situations of a baseball game, both in terms of looking at the expected number of runs scored and the probability of scoring at least one run. For this part of the analysis, we only used the data for The College of Wooster baseball team. To help read Table 5.9, consider the situation where the College of Wooster has a man on first base with 1 out. Runners that are greater than 72.65% successful at stealing second base will increase the team’s expected number of runs for the inning, runners less than 72.65% successful will decrease the team’s expected number of runs for the inning, and runners exactly 72.65% successful will break even on the team’s expected number of runs for the inning.
Table 5.9: How successful a runner must be in stealing second base in terms of expected number of runs and probability of scoring at least one run

<table>
<thead>
<tr>
<th>Situation</th>
<th>Expected Runs</th>
<th>Probability of One Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man on First, No Outs</td>
<td>79.85%</td>
<td>57.04%</td>
</tr>
<tr>
<td>Man on First, 1 Out</td>
<td>72.65%</td>
<td>52.11%</td>
</tr>
<tr>
<td>Man on First, 2 Outs</td>
<td>63.01%</td>
<td>46.49%</td>
</tr>
</tbody>
</table>

The interesting part of these results is that a base runner does not need to be nearly as successful in stealing second base if the team is trying to maximize their probability of scoring at least one run as opposed to maximizing their expected number of runs scored. This indicates that if a team is down one run late in the game with less than two outs, the base runner only needs to be at least 50% successful in stealing second base to maintain a higher probability of scoring at least one run than without attempting to steal at all. On the other hand, a base runner must be more than 70% successful in stealing second base if the team is attempting to maintain a higher number of expected runs from a stolen base attempt than without a stolen base attempt.
Chapter 6

Conclusion

Andrei Andreyevich Markov had a profound impact on the field of mathematics, ranging from probability theory to his interest in the Law of Large Numbers. Markov’s biggest contribution to mathematics was in what is now known as the theory of Markov chains. His work in stochastic processes that have the special “memoryless” property of the past has had a tremendous effect on the way we can model the randomness of this world.

Throughout this thesis paper, we developed the important pieces of the theory of Markov chains, including showing the long-term behavior of a Markov chain and the interesting properties that come along with it. Following the introductory chapter on the theory of Markov chains, we discussed both absorbing and ergodic Markov chains - two very important types of Markov chains that have many applications to the real world. Absorbing Markov chains were seen to have a great impact in the development of the baseball Markov chain model. For example, the fundamental matrix was used to calculate the expected number of batters to come to bat in one half
inning of play. The definitions and theorems shown in Chapter 3 were found to be beneficial in calculating the expected number of runs that would score for The College of Wooster baseball team during the 2013 season.

Once we calculated the expected number of runs scored for the entire College of Wooster baseball team, we determined the player value of each of the top 9 batters for Wooster’s team by calculating each player’s expected number of runs scored had each player been the batter every time up at the plate. We found that 6 of the 9 batters had a higher run expectancy than the team’s run expectancy, with the team’s top two hitters having more than 150 more expected runs than modeled for the team.

The final piece of our application of Markov chains to baseball dealt with analyzing strategies within a baseball game, namely sacrifice bunt attempts and stolen base attempts. We concluded that early in a baseball game, a team should not sacrifice bunt with a runner on first base and no outs, as it will decrease the team’s number of expected runs. However, late in a baseball game when a team may want to maximize their probability of scoring at least one run to possibly tie the game, we found that this probability will increase slightly given a sacrifice bunt attempt. In our analysis on stolen base attempts, we found that an attempt to steal second base will increase a team’s probability of scoring at least one run during that inning.

In stolen base attempts, however, there is a limit to how successful a runner must be to make it worth attempting. Our results showed that a base stealer must be around 20% more successful at stealing second base if the team is trying to maximize their expected number of runs scored as opposed to trying to maximize their probability of scoring at least one run late in a game.
The application of Markov chains to baseball has allowed us to perform these types of analyses and it has given us an opportunity to further investigate baseball strategies and when and whether managers should employ them.
Appendix A

Maple Worksheets
# This is the Cleveland weather transition matrix

\[
\begin{bmatrix}
\frac{2}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{bmatrix}
\]  

# This is the transition matrix raised to the second power and approximated to 3 decimal places

\[
\begin{bmatrix}
\frac{2}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{bmatrix}
\]  

\[ev alf(\mathbf{P}^2, 3)\]

# This is the transition matrix raised to the second power and approximated to 3 decimal places

\[
\begin{bmatrix}
0.528 & 0.389 & 0.0833 \\
0.333 & 0.375 & 0.292 \\
0.264 & 0.250 & 0.486
\end{bmatrix}
\]  

\[ev alf(\mathbf{P}^3, 3)\]

\[
\begin{bmatrix}
0.463 & 0.384 & 0.153 \\
0.365 & 0.347 & 0.288 \\
0.319 & 0.294 & 0.387
\end{bmatrix}
\]  

\[ev alf(\mathbf{P}^4, 3)\]

\[
\begin{bmatrix}
0.430 & 0.372 & 0.198 \\
0.378 & 0.343 & 0.279 \\
0.351 & 0.318 & 0.331
\end{bmatrix}
\]  

\[ev alf(\mathbf{P}^5, 3)\]

\[
\begin{bmatrix}
0.413 & 0.362 & 0.225 \\
0.384 & 0.344 & 0.272 \\
0.369 & 0.331 & 0.300
\end{bmatrix}
\]  

\[ev alf(\mathbf{P}^6, 3)\]

\[
\begin{bmatrix}
0.403 & 0.356 & 0.241 \\
0.387 & 0.345 & 0.267 \\
0.379 & 0.338 & 0.283
\end{bmatrix}
\]  

Figure A.1: Powers of the Cleveland weather transition matrix
The matrix reaches an equilibrium after being raised to the 14 power.

![Figure A.2: Powers of the Cleveland weather transition matrix](image-url)
with(LinearAlgebra):

\[
P := \begin{bmatrix}
\frac{2}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{bmatrix}
\]

# This is the Cleveland weather transition matrix

\[
\begin{bmatrix}
\frac{2}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{bmatrix}
\]

Eigenvalues\( (P^{547}) \) # This command finds the eigenvalues for this matrix

\[
\begin{bmatrix}
1 \\
\frac{5}{12} - \frac{1}{12}\sqrt{3} \\
\frac{5}{12} + \frac{1}{12}\sqrt{3}
\end{bmatrix}
\]

EigenVectors\( (P^{547}) \) # This command finds the eigenvectors corresponding to the eigenvalues listed above and puts them into a matrix. Each column is a different eigenvector

\[
\begin{bmatrix}
1 \\
\frac{5}{12} + \frac{1}{12}\sqrt{3} \\
\frac{5}{12} - \frac{1}{12}\sqrt{3}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\frac{3}{2} & \frac{3}{4} & \frac{2}{3} + \frac{1}{3}\sqrt{3} \\
\frac{5}{2} & \frac{3}{4} & \frac{2}{3} - \frac{1}{3}\sqrt{3} \\
\frac{4}{3} & \frac{3 + 2\sqrt{3}}{3} & \frac{3}{2} + \frac{3}{2}\sqrt{3} \\
1 & \frac{1}{3} & \frac{5}{2} + \frac{3}{2}\sqrt{3} \\
1 & \frac{1}{3} & \frac{5}{2} - \frac{3}{2}\sqrt{3}
\end{bmatrix}
\]

Figure A.3: Eigenvalues and eigenvectors of the Cleveland weather transition matrix
Figure A.4: The fixed probability vector for the maze problem

\[
\text{with(LinearAlgebra):}
\begin{align*}
P &:= \\
\begin{bmatrix}
0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\
\frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\
0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\
\frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\
\end{bmatrix}
\end{align*}
\]

\#This is the transition matrix for the maze problem in Example 8

\[
x := \begin{bmatrix} 2 & 3 & 2 & 3 & 4 & 3 & 2 & 3 & 2 \end{bmatrix}
\]

\[
xP \quad \text{#This command shows that } xP = x
\begin{bmatrix} 2 & 3 & 3 & 2 & 3 & 2 & 3 & 2 \end{bmatrix}
\]

\[
xP \quad \text{#This command shows that } xP = x
\begin{bmatrix} 2 & 3 & 2 & 3 & 4 & 3 & 2 & 3 & 2 \end{bmatrix}
\]
Appendix B

Excel Worksheets
Figure B.1: The first row of the fundamental matrix $N$ in the baseball model

<table>
<thead>
<tr>
<th>Column</th>
<th>First Row Entries of $N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0220339</td>
</tr>
<tr>
<td>2</td>
<td>0.35551356</td>
</tr>
<tr>
<td>3</td>
<td>0.07447458</td>
</tr>
<tr>
<td>4</td>
<td>0.01355932</td>
</tr>
<tr>
<td>5</td>
<td>0.16279154</td>
</tr>
<tr>
<td>6</td>
<td>0.00172203</td>
</tr>
<tr>
<td>7</td>
<td>0.03569492</td>
</tr>
<tr>
<td>8</td>
<td>0.02912541</td>
</tr>
<tr>
<td>9</td>
<td>0.6248645</td>
</tr>
<tr>
<td>10</td>
<td>0.44580009</td>
</tr>
<tr>
<td>11</td>
<td>0.09728173</td>
</tr>
<tr>
<td>12</td>
<td>0.02145508</td>
</tr>
<tr>
<td>13</td>
<td>0.33240408</td>
</tr>
<tr>
<td>14</td>
<td>0.00371669</td>
</tr>
<tr>
<td>15</td>
<td>0.07704076</td>
</tr>
<tr>
<td>16</td>
<td>0.07932159</td>
</tr>
<tr>
<td>17</td>
<td>0.39543958</td>
</tr>
<tr>
<td>18</td>
<td>0.43226849</td>
</tr>
<tr>
<td>19</td>
<td>0.09872791</td>
</tr>
<tr>
<td>20</td>
<td>0.02874116</td>
</tr>
<tr>
<td>21</td>
<td>0.46350684</td>
</tr>
<tr>
<td>22</td>
<td>0.00579094</td>
</tr>
<tr>
<td>23</td>
<td>0.11296349</td>
</tr>
<tr>
<td>24</td>
<td>0.13704057</td>
</tr>
<tr>
<td>Column</td>
<td>First Row Entries of B</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------</td>
</tr>
<tr>
<td>1</td>
<td>0.23955055</td>
</tr>
<tr>
<td>2</td>
<td>0.33495086</td>
</tr>
<tr>
<td>3</td>
<td>0.34464465</td>
</tr>
<tr>
<td>4</td>
<td>0.08085394</td>
</tr>
</tbody>
</table>

Figure B.2: The first row entries of the matrix B in the baseball model
Table B.1: Transition probabilities from the 8th state to any of the 28 states in the baseball transition matrix and powers of the transition matrix

<table>
<thead>
<tr>
<th>State</th>
<th>P</th>
<th>P^2</th>
<th>P^3</th>
<th>P^4</th>
<th>P^5</th>
<th>P^6</th>
<th>P^7</th>
<th>P^8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.013</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.000</td>
<td>0.019</td>
<td>0.008</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>0.000</td>
<td>0.006</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>0.008</td>
<td>0.003</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>0.197</td>
<td>0.064</td>
<td>0.029</td>
<td>0.012</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>6</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>7</td>
<td>0.065</td>
<td>0.021</td>
<td>0.009</td>
<td>0.004</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>8</td>
<td>0.127</td>
<td>0.049</td>
<td>0.017</td>
<td>0.007</td>
<td>0.003</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>9</td>
<td>0.000</td>
<td>0.015</td>
<td>0.009</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>10</td>
<td>0.000</td>
<td>0.000</td>
<td>0.032</td>
<td>0.018</td>
<td>0.009</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>11</td>
<td>0.000</td>
<td>0.000</td>
<td>0.010</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>12</td>
<td>0.000</td>
<td>0.009</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>13</td>
<td>0.000</td>
<td>0.227</td>
<td>0.110</td>
<td>0.068</td>
<td>0.035</td>
<td>0.017</td>
<td>0.008</td>
<td>0.004</td>
</tr>
<tr>
<td>14</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>15</td>
<td>0.000</td>
<td>0.076</td>
<td>0.037</td>
<td>0.020</td>
<td>0.010</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>16</td>
<td>0.576</td>
<td>0.146</td>
<td>0.085</td>
<td>0.040</td>
<td>0.020</td>
<td>0.010</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>17</td>
<td>0.000</td>
<td>0.00</td>
<td>0.014</td>
<td>0.011</td>
<td>0.007</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>18</td>
<td>0.000</td>
<td>0.003</td>
<td>0.001</td>
<td>0.038</td>
<td>0.026</td>
<td>0.016</td>
<td>0.009</td>
<td>0.005</td>
</tr>
<tr>
<td>19</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.012</td>
<td>0.008</td>
<td>0.005</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>20</td>
<td>0.014</td>
<td>0.005</td>
<td>0.010</td>
<td>0.007</td>
<td>0.005</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>21</td>
<td>0.000</td>
<td>0.000</td>
<td>0.197</td>
<td>0.128</td>
<td>0.098</td>
<td>0.060</td>
<td>0.035</td>
<td>0.019</td>
</tr>
<tr>
<td>22</td>
<td>0.000</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>23</td>
<td>0.000</td>
<td>0.000</td>
<td>0.067</td>
<td>0.043</td>
<td>0.029</td>
<td>0.018</td>
<td>0.010</td>
<td>0.006</td>
</tr>
<tr>
<td>24</td>
<td>0.000</td>
<td>0.332</td>
<td>0.127</td>
<td>0.099</td>
<td>0.057</td>
<td>0.035</td>
<td>0.020</td>
<td>0.011</td>
</tr>
<tr>
<td>25</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.009</td>
<td>0.015</td>
<td>0.020</td>
<td>0.023</td>
<td>0.024</td>
</tr>
<tr>
<td>26</td>
<td>0.000</td>
<td>0.008</td>
<td>0.016</td>
<td>0.025</td>
<td>0.059</td>
<td>0.083</td>
<td>0.097</td>
<td>0.105</td>
</tr>
<tr>
<td>27</td>
<td>0.000</td>
<td>0.008</td>
<td>0.011</td>
<td>0.168</td>
<td>0.271</td>
<td>0.347</td>
<td>0.394</td>
<td>0.421</td>
</tr>
<tr>
<td>28</td>
<td>0.000</td>
<td>0.000</td>
<td>0.196</td>
<td>0.270</td>
<td>0.329</td>
<td>0.363</td>
<td>0.384</td>
<td>0.395</td>
</tr>
</tbody>
</table>
Figure B.3: Probability of scoring at least one run with and without a bunt attempt starting with a runner on first base and no outs for The College of Wooster
Man on First, No Outs

<table>
<thead>
<tr>
<th>Probability of Scoring At Least 1 Run Given A Bunt Attempt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.37011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probability of Scoring At Least 1 Run Without A Bunt Attempt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.347861</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prob. Of At Least 1 Run Given Successful Bunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.437689</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prob. Of At Least 1 Run Given Unsuccessful Bunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.219693</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prob. Of a Successful Bunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prob. Of Unsuccessful Bunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.31</td>
</tr>
</tbody>
</table>

Figure B.4: Probability of scoring at least one run with and without a bunt attempt starting with a runner on first base and no outs for Major League Baseball
Man on First, No Outs

**Probability of Scoring At Least 1 Run Given A Stolen Base Attempt**
0.557977

**Probability of Scoring At Least 1 Run Without A Stolen Base Attempt**
0.486354

Prob. Of At Least 1 Run Given Successful Steal
0.695296

Prob. Of At Least 1 Run Given Unsuccessful Steal
0.208959

Man on First, One Out

**Probability of Scoring At Least 1 Run Given A Stolen Base Attempt**
0.416026

**Probability of Scoring At Least 1 Run Without A Stolen Base Attempt**
0.325564

Prob. Of At Least 1 Run Given Successful Steal
0.545951

Prob. Of At Least 1 Run Given Unsuccessful Steal
0.0858

Man on First, Two Outs

**Probability of Scoring At Least 1 Run Given A Stolen Base Attempt**
0.233633

**Probability of Scoring At Least 1 Run Without A Stolen Base Attempt**
0.151342

Prob. Of At Least 1 Run Given Successful Steal
0.325554

Prob. Of At Least 1 Run Given Unsuccessful Steal
0

Figure B.5: Probability of scoring at least one run with and without a stolen base attempt starting with a runner on first base and no outs, one out, and two outs for The College of Wooster
**Man on First, No Outs**

| Probability of Scoring At Least 1 Run Given A Stolen Base Attempt | 0.460231 |
| Probability of Scoring At Least 1 Run Without A Stolen Base Attempt | 0.347861 |
| Prob. Of At Least 1 Run Given Successful Steal | 0.5788 |
| Prob. Of At Least 1 Run Given Unsuccessful Steal | 0.143146 |

| Prob. Of a Stolen Base | 0.727838 |
| Prob. Of Unsuccessful Steal | 0.272162 |

**Man on First, One Out**

| Probability of Scoring At Least 1 Run Given A Stolen Base Attempt | 0.335461 |
| Probability of Scoring At Least 1 Run Without A Stolen Base | 0.22431 |
| Prob. Of At Least 1 Run Given Successful Steal | 0.437689 |
| Prob. Of At Least 1 Run Given Unsuccessful Steal | 0.062073 |

**Man on First, Two Outs**

| Probability of Scoring At Least 1 Run Given A Stolen Base Attempt | 0.181976 |
| Probability of Scoring At Least 1 Run Without A Stolen Base | 0.102342 |
| Prob. Of At Least 1 Run Given Successful Steal | 0.250023 |
| Prob. Of At Least 1 Run Given Unsuccessful Steal | 0 |

Figure B.6: Probability of scoring at least one run with and without a stolen base attempt starting with a runner on first base and no outs, one out, and two outs for Major League Baseball
Bibliography


