Spring 2018

Analyzing the Probabilistic Spread of a Virus on Various Networks

Teagan DeCusatis
Bard College

Analyzing the Probabilistic Spread of a Virus on Various Networks

A Senior Project submitted to
The Division of Science, Mathematics, and Computing
of
Bard College

by
Teagan DeCusatis

Annandale-on-Hudson, New York
May, 2018

Abstract

In this project we model the spread of a virus on networks as a probabilistic process. We assume the virus breaks out at one vertex on a network and then spreads to neighboring vertices in each time step with a certain probability. Our objective is to find probability distributions that describe the uncertain number of infected vertices at a given time step. The networks we consider are paths, cycles, star graphs, complete graphs, and broom graphs. Through the use of Markov chains and Jordan Normal Form we analyze the probability distribution of these graphs, characterizing the transition matrix for each graph as well as the Jordan Form matrices.

Contents

List of Figures

List of Tables

Dedication

To my Dad, for sparking my interest in math and helping me every step of the way.

Acknowledgments

I would like to thank my advisor, Lauren Rose, for all of her help and guidance during this project. Her comments and encouragement motivated me to work harder. Thank you to the members of my board, Maria Belk and Stefan Méndez-Diez, for their help and advice on my work. I would aslo like to thank the professors of the Bard Mathematics Department for inspiring me with their passion and knowledge these past four years.

Finally, I would like to thank my family. To my parents for always being there for me and supporting me, to my brothers for being my role models and always making me laugh, and to my grandfather whose generosity helped make this possible.

Thank you.

1

Introduction

In this project we model the spread of a virus on graphs as a probabilistic process. The analysis of this process is useful in understanding how a virus or information is transmitted through different types of networks. Our model assumes that at the initial time step a virus breaks out on a graph at one vertex. In the following time steps the virus spreads to neighboring vertices with probability p and doesn't spread with probability $(1 - p)$. Using this model our goal is to determine the probability distribution that describes the number of infected vertices at a given time step on different types of networks.

This project was inspired by Anam Nasim's senior project *Spread of a Virus on Networks: A Probabilistic Approach* [5] and Yushan Jiang's senior project *The Analysis of Probabilistic Spread on Complete Graphs* [4]. They introduced a probability model and used it to study the spread of a virus on different graphs. Anam Nasim studied the spread of a virus on paths, cycles, star graphs, and complete graphs while Yushan Jiang focused her study on complete graphs. In Anam Nasim's project she introduced a method of analyzing this model using Markov chains. Both Anam and Yushan used this method to study virus spread on complete graphs. In this project we use this same Markov chain method to study virus spread on paths, star graphs, cycles, complete graphs, and broom graphs which we define in Chapter 6. In our analysis of these graphs we also use the Jordan Normal Form to simplifiy our Markov chain computations.

Let's look at an example of a virus spreading on the star graph with 3 vertices denoted S_3 (see Figure 1.0.1).

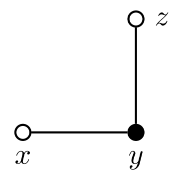
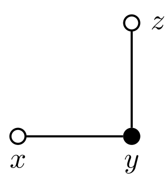
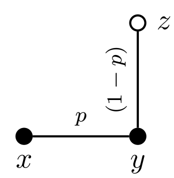


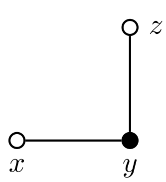
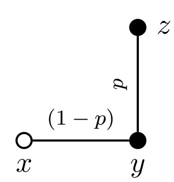
Figure 1.0.1. The star graph S_3

On this graph we assume the virus breakes out at the vertex y. Suppose we want to find the probability that the virus spreads to one more vertex after one time step. This can happen two ways.

<u>Case 1</u>: The virus spreads to x and not to z. The following table illustrates this probability.

$t = 0$	$t = 1$	Probability
 A virus breaks out at y	 The virus spreads to x and doesn't spread to z	$p(1-p)$

<u>Case 2</u>: The virus spreads to z and not to x. The following table illustrates this probability.

$t = 0$	$t = 1$	Probability
 A virus breaks out at y	 The virus spreads to z and doesn't spread to x	$p(1-p)$

Adding these two cases together we get that the probability of exactly one more vertex being infected after one time step is $p(1-p) + p(1-p) = 2p(1-p)$.

In this project we analyze one time-step probabilities like this one in our Markov chain process. If we arrange all the possible one time step probabilities into a matrix we get

$$T = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

where each element p_{ij}, is the probability of moving from i infected vertices to j infected vertices in one time step.

We find for paths with n vertices that this matrix is an $n \times n$ matrix of the form

$$T = \begin{pmatrix} (1-p) & p & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & (1-p) & p \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix},$$

for cycles with n vertices this matrix is an $n \times n$ matrix of the form

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & (1-p)^2 & 2p(1-p) & p^2 \\ \vdots & & & & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix},$$

and for star graphs with n vertices this matrix has elements such that

$$p_{ij} = \binom{n-i}{j-i} p^{j-i}(1-p)^{n-j}.$$

In Chapter 2 we introduce basic terminology for graph theory, linear algebra, probability, and Markov chains. In Chapter 3 we introduce our probability model and how to analyze it using Markov chains. In Chapter 4, we present theorems specific to our probability model that will be useful throughout this project. In Chapter 5 we analyze virus spread on paths, star graphs, cycles, and complete graphs. In Chapter 6 we define broom graphs and analyze their transition matricies. Finally, in Chapter 7 we see how our Markov chain method doesn't work with all graphs and how to modify it so that it does.

2
Preliminaries

In this chapter we review some basic concepts of graph theory, linear algebra, and probability that will be needed to understand this project.

2.1 Graph Theory

In this section we will review some basic elements of graph theory. Our main reference for this section is Wilson [7].

A **simple graph** G consists of a non-empty finite set $V(G)$ of elements called **vertices** (or **nodes**) and a finite set $E(G)$ of distinct unordered pairs of the elements in $V(G)$ called **edges**. An edge $\{v, w\}$ is said to join the vertices v and w and is often abbreviated to vw. For example, Figure 2.1.1 shows a simple graph G with vertex-set $V(G) = \{u, v, w, z\}$ and edge-set $E(G) = \{uv, uw, uz, wz, vw\}$. Note that in this project we will only be considering simple graphs but will refer to them as graphs.

Two vertices v and w are **adjacent** if there is an edge vw joining them, and the vertices v and w are then **incident** with that edge. The **degree** of a vertex v is the number of edges incident with v. For example, in Figure 2.1.1 the vertex u is adjacent to the vertex z since there is an edge uz joining them. Therefore, u and z are incident with the edge uz. Also, the vertex u has degree 3 since it is incident with the edges uv, uw and uz.

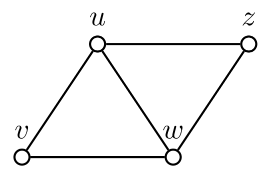


Figure 2.1.1. A simple graph G

Now we will discuss the types of graphs that will be used in this project. Definitions 2.1.1 and 2.1.2 were taken directly from Wilson [7, p. 9-10].

Definition 2.1.1. If two graphs G_1 and G_2 and their vertex sets $V(G_1)$ and $V(G_2)$ are disjoint, then their **union** $G_1 \cup G_2$ is the graph with vertex set $V(G_1) \cup V(G_2)$ and edge-family $E(G_1) \cup E(G_2)$.

The following figure illustrates the union of two graphs G_1 and G_2.

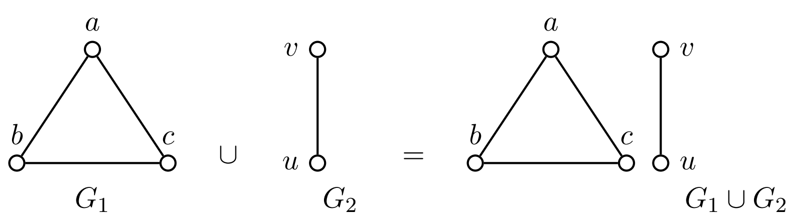


Figure 2.1.2. The union of G_1 and G_2

Definition 2.1.2. A graph is **connected** if it cannot be expressed as a union of graphs, and **disconnected** otherwise.

Definition 2.1.3. A **path** is a connected graph with two vertices of degree 1 and all other vertices of degree 2. A path can be drawn so that all of its vertices and edges lie on a straight line. We denote a path with n vertices by P_n.

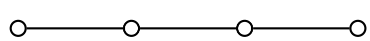


Figure 2.1.3. The path P_4

Definition 2.1.4. A **star graph** with n vertices is a connected graph that has one vertex of degree $n-1$ and $n-1$ vertices of degree 1. There is one vertex at the center of the graph that is adjacent to all of the other nodes. A star with n vertices is denoted S_n.

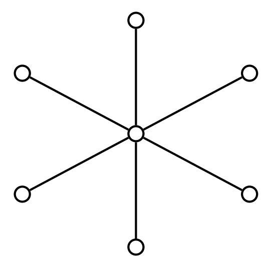


Figure 2.1.4. The star graph S_7

Definition 2.1.5. A **cycle** is a connected graph in which each vertex has degree 2. A cycle can also be thought of as a path with an edge connecting the endpoints. We denote a cycle with n vertices by C_n.

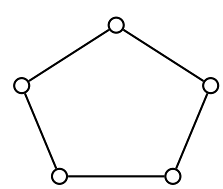


Figure 2.1.5. The cycle graph C_5

Definition 2.1.6. A **complete graph** is a connected graph in which all vertices are adjacent to each other. We denote a complete graph with n vertices as K_n.

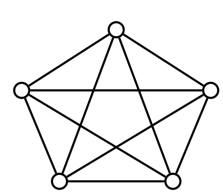


Figure 2.1.6. The complete graph K_5

2.2 Linear Algebra

In this section we review some basic concepts of linear algebra. Our main reference for this section is Edwards and Penney [1].

A **matrix** is an $n \times m$ array of numbers with n rows and m columns. The **identity matrix** is a square matrix I that has 1's on the diagonal and 0's everywhere else. The $n \times n$ identity matrix has the property that $AI = IA = A$ for any $n \times m$ matrix A. Note that in this project we only consider square $n \times n$ matrices.

The following definition is taken from Edwards and Penney [1, p. 190].

Definition 2.2.1. A matrix B such that $AB = BA = I$ is called an **inverse matrix** of the matrix A. There is precisely one matrix B such that $AB = BA = I$ and it is often denoted A^{-1}.

Example 2.2.2. Let $A = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}$. Then $A^{-1} = \begin{pmatrix} 1.5 & -2.5 \\ -1 & 2 \end{pmatrix}$, as shown below:

$$AA^{-1} = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1.5 & -2.5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

and

$$A^{-1}A = \begin{pmatrix} 1.5 & -2.5 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Definition 2.2.3. An **upper triangular matrix** is a square matrix that has only zeros below the diagonal.

The following definition is taken directly from Edwards and Penney [1, p. 366]

Definition 2.2.4. The number λ is said to be an **eigenvalue** of the $n \times n$ matrix A provided there exists a nonzero vector v such that

$$Av = \lambda v,$$

in which case the vector v is called an **eigenvector** of the matrix A. We also say that the eigenvector v is associated with the eigenvalue λ, or that the eigenvalue λ corresponds to the eigenvector v.

Example 2.2.5. Let $A = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}$. If the vector $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, then

$$Av_1 = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3v_1$$

Therefore v_1 is an eigenvector of A associated with the eigenvalue $\lambda_1 = 3$.

If the vector $v_2 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$, then

$$Av_2 = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \end{pmatrix} = -2 \begin{pmatrix} -3 \\ 2 \end{pmatrix} = -2v_2$$

Therefore v_2 is an eigenvector of A associated with the eigenvalue $\lambda_2 = -2$. In summary, the numbers $\lambda_1 = 3$ and $\lambda_2 = -2$ are both eigenvalues of the matrix A and correspond to the eigenvectors $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$, respectively.

The following definition is taken from Edwards and Penney [1, p. 368].

Definition 2.2.6. The number λ is an eigenvalue of the matrix A if and only if λ satisfies the **characteristic equation** of A,

$$det(A - \lambda I) = 0.$$

Theorem 2.2.7. *If a matrix A is upper triangular then the eigenvalues of A lie on the diagonal.*

Example 2.2.8. Consider the matrix $A = \begin{pmatrix} 1 & 4 & 1 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix}$. To find the eigenvalues of A we solve the characteristic equation $det(A - \lambda I) = 0$ so,

$$det(A - \lambda I) = det \begin{pmatrix} 1-\lambda & 4 & 1 \\ 0 & 2-\lambda & 5 \\ 0 & 0 & 3-\lambda \end{pmatrix} = (1-\lambda) \begin{pmatrix} 2-\lambda & 5 \\ 0 & 3-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda)(3-\lambda) = 0.$$

Therefore the eigenvalues of A are $\lambda = 1, 2, 3$ which are the elements that lie on the diagonal of A.

The following definition is taken from Edwards and Penney [1, p. 451].

Definition 2.2.9. If λ is an eigenvalue of the matrix A, then a **rank** r **generalized eigenvector** associated with λ is a vector v such that

$$(A - \lambda I)^r v = 0 \text{ but } (A - \lambda I)^{r-1} v \neq 0.$$

If $r = 1$ then this simply means that v is an eigenvector associated with λ. Then finding an eigenvector of rank 2 associated with λ means finding v_2 such that $(A-\lambda I)v_2 = v_1$. Furthermore, a **length k chain of generalized eigenvectors based on the eigenvector** v_1 is a set $\{v_1, v_2, ..., v_k\}$ of k generalized eigenvectors such that

$$(A - \lambda I)v_k = v_{k-1},$$
$$(A - \lambda I)v_{k-1} = v_{k-2},$$
$$\vdots$$
$$(A - \lambda I)v_2 = v_1.$$

Example 2.2.10. Let $A = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}$. Because A is upper triangular by Theorem 2.2.7 we know that the eigenvalues of A are $\lambda = 2$ and $\lambda = 2$. Let $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Then we have that

$$Av_1 = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2v_1.$$

Note v_1 is the only eigenvector of A. Since $\lambda = 2$ has multiplicity 2 we need to find the rank 2 generalized eigenvector v_2 such that $(A - 2I)v_2 = v_1$. Let $v_2 = \begin{pmatrix} 0 \\ \frac{1}{5} \end{pmatrix}$. Then we have that

$$(A - 2I)v_2 = \begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = v_1.$$

Therefore $v_2 = \begin{pmatrix} 0 \\ \frac{1}{5} \end{pmatrix}$ is a generalized eigenvector of A associated with the eigenvalue $\lambda = 2$.

The following definitions are taken from Edwards and Penney [1, p. 378].

Definition 2.2.11. The $n \times n$ matrices A and B are called **similar** provided that there exists an invertible matrix P such that

$$B = P^{-1}AP.$$

Definition 2.2.12. the matrix D is **diagonal** if it has zeros everywhere except the diagonal.

Definition 2.2.13. The matrix A is **diagonalizable** if it is similar to a diagonal matrix D; that is, there exists a diagonal matrix D and an invertible matrix P such that

$$A = PDP^{-1}.$$

Theorem 2.2.14. *If a square matrix A is similar to the matrix B then there exists an invertible matrix P such that*

$$A^n = PB^nP^{-1}$$

where n is a positive number.

Proof. Suppose $A = PBP^{-1}$. Then we have

$$A^n = (PBP^{-1})^n$$

$$= (PBP^{-1})(PBP^{-1})...(PBP^{-1})$$

$$= PB(P^{-1}P)B(P^{-1}P)...(P^{-1}P)BP^{-1}$$

$$= PB^nP^{-1}.$$

□

In this project we analyze the matrix T to the power of t where t is a positive integer. However, if T is diagonalizable, we can look at the diagonal matrix D that is similar to T. Diagonal matrices are easy to multiply so finding D^t will simplify our computation. However, most of the matrices we will look at won't be diagonalizable, so we will look at their Jordan Normal Form instead.

The following theorem and definition about Jordan Normal Form are taken directly from Edwards and Penney [1, p. 460].

Theorem 2.2.15. *If the $n \times n$ matrix A has s linearly independent eigenvectors $v_1, v_2, ..., v_s$, then it is similar to a block-diagonal matrix of the* ***Jordan Normal Form***

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_s \end{pmatrix},$$

where each submatrix J_i is a $k \times k$ ***Jordan block*** *of the form*

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix},$$

with λ_i being the eigenvalue of A corresponding to the eigenvector v_i; if $k = 1$, then $J_i = \big(\lambda_i\big)$.

Definition 2.2.16. If the Jordan block J_i is of size $k \times k$, then it corresponds to a length k chain of generalized eigenvectors based on the eigenvector v_i. If all these generalized eigenvectors are arranged as column vectors in proper order corresponding to the order of the Jordan blocks, the result is a nonsingular $n \times n$ matrix Q such that

$$J = Q^{-1}AQ.$$

The block diagonal matrix J such that $A = QJQ^{-1}$ is called the **Jordan Normal Form** of the matrix A and is unique.

Example 2.2.17. Let $A = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}$. From Example 2.2.10 we know that the eigenvalues are both $\lambda = 2$ and that the only eigenvector of A is $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Therefore J is made up of one Jordan block. So

$$J = (J_1) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

We can find the matrix Q by arranging the eigenvectors and generalized eigenvectors of A into a matrix;

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \text{ where } Q^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

Therefore the Jordan Normal Form of A is

$$QJQ^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} = A.$$

2.3 Probability

In this section we review some basic concepts of probability. Our main reference for this section is Schay [6].

A **sample space** S is a set which contains all the possible outcomes of a situation. For example, when flipping a fair coin there are two possible outcomes, heads or tails, so the sample space is $S = \{H, T\}$. The elements of the sample space are called **events**. The probability of event A is denoted by $P(A)$. In the example of flipping a fair coin $P(H) = \frac{1}{2}$. The **conditional probability** of event A given event B is denoted $P(A|B)$. A **random variable** X is a real valued function on a sample space S. The value of X is a numerical value that represents an outcome of an experiment. See Example 2.3.1.

Example 2.3.1. Consider tossing a fair coin twice. The possible outcomes of this event are described by the sample space $S = \{HH, HT, TH, TT\}$. Let X be our random variable that denotes the number of heads obtained. The following table represets the value of X for each outcome in the sample space.

Outcome	HH	HT	TH	TT
X	2	1	1	0

From the table we can see that $P(X = 0) = \frac{1}{4}$, $P(X = 1) = \frac{1}{2}$, and $P(X = 2) = \frac{1}{4}$.

Note that all random variables in this project will be **discrete random variables**, meaning that X only takes on a finite or countably infinite number of values.

The following definition is taken from Schay [6, p. 72].

Definition 2.3.2. For any probability space and any random variable X on it, the function $f : X \to \mathbb{R}$ defined by $f(x) = P(X = x)$, is called the **probability function** of X.

We will also refer to this function as the **probability distribution** of X.

Example 2.3.3. Consider tossing a fair coin twice and let X be the random variable that represents the number of heads obtained as in Example 2.3.1. Then the probability distribution of X is given by

$$f(x) = \begin{cases} \frac{1}{4}, & \text{if } x = 0 \\ \frac{1}{2}, & \text{if } x = 1 \\ \frac{1}{4}, & \text{if } x = 2 \end{cases}.$$

Theorem 2.3.4. ***(The Law of Total Probability)*** *Suppose $B_1, B_2, ..., B_n$ are mutually exclusive events with non-zero probabilities, whose union $B_1 \cup B_2 \cup ... \cup B_n = S$. Then any event A in the sample space S can be written as*

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ... + P(A|B_n)P(B_n)$$

Definition 2.3.5. A **binomial coefficient** describes the number of different ways r elements can be chosen out of a set of n elements and is denoted $\binom{n}{r}$. The formula for the binomial coefficient is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

The following definition is taken directly from Schay [6, p. 74].

Definition 2.3.6. A random variable X is called a **binomial random variable** with parameters n and p, if it has the **binomial distribution** with probability function

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{if } x = 0, 1, 2, ..., n$$

Another way to think of a binomial random variable is that it counts the number of "successes" in n trials. Example 2.3.7 illustrates this.

Example 2.3.7. Suppose there's a jar that contains 3 red marbles and 2 blue marbles. Let X be the number of red marbles obtained when we pick two marbles from the jar with replacement. In this example we are considering picking a red marble as a success. The sample space for this experiment is $S = \{RR, RB, BR, BB\}$. Our parameter p is the probability of a success, the probabiliy of picking a red marble from the jar, so $p = P(R) = \frac{3}{5}$. Our paramere n is the number of trials so $n = 2$ since we are picking 2 marbles from the jar.

The following table illustrates the events in the sample space and the value of X for each event.

Outcome	RR	RB	BR	BB
X	2	1	1	0

Using the binomial distribution formula we can find $P(X = x)$ for each x.

$P(X = 0) = \binom{2}{0}\left(\frac{3}{5}\right)^0\left(1 - \frac{3}{5}\right)^2 = \frac{4}{25}$
$P(X = 1) = \binom{2}{1}\left(\frac{3}{5}\right)^1\left(1 - \frac{3}{5}\right)^1 = \frac{12}{25}$
$P(X = 2) = \binom{2}{2}\left(\frac{3}{5}\right)^2\left(1 - \frac{3}{5}\right)^0 = \frac{9}{25}$

2.4 Markov Chains

In this section we will review Markov chains. Our main references for this section are Hillier and Lieberman [3] and Grinstead and Snell [2].

A **Markov chain** is a type of **stochastic process**, which is a collection of random variables $\{X_t\}$ indexed by $t \in \{0, 1, 2, ...\}$. The random variable X_t represents the state or position of the system at time t. The **states** of the system are expressed as $X_t = 0, 1, 2, ...$

In this project we will only be considering discrete time stochastic processes with a finite number of states.

Definition 2.4.1. A stochastic process is a **Markov Chain** if it has the following Markovian property

$$P(X_{t+1} = j | X_0 = i_0, X_1 = i_1, ..., X_{t-1} = i_{t-1}, X_t = i) = P(X_{t+1} = j | X_t = i).$$

This says, given that the present state is $X_t = i$, the next state in the process X_{t+1} is independent of any past events and depends only on the present state.

These probabilies are called one-step **transition probabilities** and are denoted

$$p_{ij} = P(X_{t+1} = j | X_t = i).$$

This is the probability that in one time step the process will move from state i to state j.

These probabilities can be arranged in an $n \times n$ matrix

$$T = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}.$$

This is called the **transition matrix** of the Markov Chain.

Example 2.4.2. Let the following Markov chain represent a weather system that transitions between cloudy, sunny, and rainy weather. In this Markov chain one day is one time step and the states are the different types of weather C, S, and R.

$$T = \begin{array}{c c} & \begin{array}{ccc} C & S & R \end{array} \\ \begin{array}{c} C \\ S \\ R \end{array} & \begin{pmatrix} .25 & .5 & .25 \\ .45 & .25 & .3 \\ .25 & .25 & .5 \end{pmatrix} \end{array}$$

If it's cloudy today, there's a 50% chance of it being sunny tomorrow, a 25% chance of being cloudy, and a 25% chance of being rainy. If it's sunny today, there's a 45% chance of it being cloudy tomorrow, a 25% chance of it being sunny, and a 30% chance of being rainy. If its rainy today, there's a 25% chance of it being cloudy tomorrow, a 25% chance of being sunny, and a 50% chance of being rainy.

The following theorem is taken from Grinstead and Snell [2, p. 407]. The proof is ours.

Theorem 2.4.3. *Let T be the transition matrix of a Markov chain. The ijth entry $p_{ij}^{(n)}$ of the matrix T^n gives the probability that the Markov chain, starting in state i, will be in state j after n steps.*

Proof. We will use induction on n.

Base Case: If $n = 1$, then $p_{ij}^{(1)}$ is the probability that the Markov chain starts in state i and after 1 time step is in state j. This is true by the definition of transition probabilities.

Induction Hypothesis: Let $n \geq 1$. Assume that the ijth entry $p_{ij}^{(n)}$ of T^n is the probability that the Markov chain has gone from state i to state j after n time steps.

Induction Step: Suppose the induction hypothesis holds. Let m be the total number of states in the Markov chain. Then

$$\begin{aligned} p_{ij}^{(n+1)} &= (T^{n+1})_{ij} \\ &= (T^n \cdot T)_{ij} \\ &= p_{i1}^{(n)} \cdot p_{1j}^{(1)} + p_{i2}^{(n)} \cdot p_{2j}^{(1)} + ... + p_{im}^{(n)} \cdot p_{mj}^{(1)} \\ &= \sum_{k=1}^{m} p_{ik}^{(n)} \cdot p_{kj}^{(1)} \end{aligned}$$

Note, we are taking the dot product of the ith row of T^n and the jth column of T to get the ijth element of T^{n+1}. By the induction hypothesis we know that $p_{ik}^{(n)}$ is the probability that the Markov chain goes from state i to state k in n time steps. Hence $\sum_{k=1}^{m} p_{ik}^{(n)} \cdot p_{kj}^{(1)}$ covers all the possible states k and therefore all possible ways the Markov chain can get from state i to state j. Hence $p_{ij}^{(n+1)}$ is the probability that the Markov chain, starting in state i, will be in state j after $n+1$ time steps.

$\square$

Example 2.4.4. Using the Markov chain from Example 2.4.2, what is the probability that it will be sunny two days from now given that it's cloudy today? By Theorem 2.4.3 we can find this probability by looking at the element $p_{CS}^{(2)}$ of the matrix T^2. So we have,

$$T^2 = \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.45 & 0.25 & 0.3 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} \cdot \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.45 & 0.25 & 0.3 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

$$= \begin{array}{c} \\ C \\ S \\ R \end{array} \begin{array}{c} \begin{array}{ccc} C & S & R \end{array} \\ \begin{pmatrix} 0.35 & 0.3125 & 0.3375 \\ 0.3 & 0.3625 & 0.3375 \\ 0.3 & 0.3125 & 0.3875 \end{pmatrix} \end{array}.$$

Therefore, the probability that it will be sunny two days from now given that it's cloudy today is $p_{CS}^{(2)} = 0.3125$.

The following theorem is taken from Grindstead and Snell [2, p. 409]. The proof is ours.

Theorem 2.4.5. *Let T be the transition matrix of a Markov chain, and let $\boldsymbol{u}$ be the probability vector which represents the starting distribution. Then the probability that the chain is in state i after n steps is the ith entry in the vector*

$$\boldsymbol{u}^{(n)} = \boldsymbol{u}T^n$$

In other words, this theorem allows us to find the probability that a Markov chain is in state i after t time steps. This theorem will be helpful in finding the probability distributions of Markov chains.

Proof. We will use induction on n.

Base Case: Let $n = 1$. Let X_t be the random variable that represent the state of the Markov chain at time t. We define $u_j^{(n)}$ as the jth element of the vector $\mathbf{u}^{(n)}$. So, $u_j^{(1)} = P(X_1 = j)$. By the law of Total Probability we have

$$P(X_1 = j)$$

$$= P(X_1 = j|X_0 = 1)P(X_0 = 1)+P(X_1 = j|X_0 = 2)P(X_0 = 2)+...+P(X_1 = j|X_0 = k)P(X_0 = k)$$

$$= \sum_{i=1}^{k} P(X_1 = j|X_0 = i)P(X_0 = i).$$

Since $p_{ij} = P(X_{t+1} = j|X_t = i)$ we have

$$u_j^{(1)} = p_{1j} \cdot u_1^{(0)} + p_{2j} \cdot u_2^{(0)} + ... + p_{kj} \cdot u_k^{(0)}$$

$$= \sum_{i=1}^{k} p_{ij} u_i^{(0)} = u^{(0)} T_j.$$

Where T_j is the jth column of the transition matrix T.

Therefore $u_j^{(1)} = \mathbf{u}^{(0)} T_j$. So for all $j \in \{1, ..., k\}$ we have $\mathbf{u}^{(1)} = \mathbf{u}^{(0)} T^1$

Induction Hypothesis: Assume $\mathbf{u}^{(n)} = \mathbf{u}^{(0)} T^n$.

Induction Step: Suppose the induction hypothesis holds. Then we have

$$u_j^{(n+1)} = P(X_{n+1} = j) = \sum_{i=1}^{k} P(X_{n+1} = j|X_n = i)P(X_n = i)$$

by the law of Total Probability. So,

$$u_j^{(n+1)} = \mathbf{u}^{(n)} T_j$$

and for all $j \in \{1, ..., k\}$ we have

$$\mathbf{u}^{(n+1)} = \mathbf{u}^{(n)}T.$$

By the induction hypothesis

$$\mathbf{u}^{(n+1)} = \mathbf{u}^{(n)}T = (\mathbf{u}^{(0)}T^n)T = \mathbf{u}^{(0)}T^{n+1}.$$

$\square$

Example 2.4.6. Using the Markov chain from Example 2.4.2, let the vector that represents the starting distribution be $\mathbf{u} = (1/3, 1/3, 1/3)$ that is $\mathbf{u} = (P(X_0 = C) = 1/3, P(X_0 = S) = 1/3, P(X_0 = R) = 1/3)$. Then we can calculate the distribution of the states after 2 days by using Theorem 2.4.5 and the transition matrix T^2 that we found in Example 2.4.4.

$$\begin{aligned}\mathbf{u}^{(2)} = \mathbf{u}T^2 &= (1/3, 1/3, 1/3) \cdot \begin{pmatrix} 0.35 & 0.3125 & 0.3375 \\ 0.3 & 0.3625 & 0.3375 \\ 0.3 & 0.3125 & 0.3875 \end{pmatrix} \\ &= (0.316667, 0.329167, 0.354167).\end{aligned}$$

Therefore, $P(X_2 = C) = 0.316667,\ \ P(X_2 = S) = 0.329167,\ $ and $P(X_2 = R) = 0.354167$.

3
The Markov Chain Model

In this chapter we see how a virus spreads across a graph by setting up guidelines for our probability model. We see two methods for analyzing virus spread on graphs. The first method involves finding individual probabilities and looking for patterns while the second method sets up graphs as Markov chains.

3.1 Guidelines for Our Model

In this section we set up the guidelines for our model and then illustrate them using an example.

1. Each graph has a vertex called the root node. The root node is the initial infected vertex at time $t = 0$.

2. If a vertex v is infected the virus can spread to any uninfected vertex adjacent to v in the next time step with probability p and not spread with probability $(1 - p)$.

3. Once a vertex is infected it cannot be uninfected or re-infected.

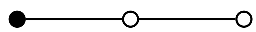
4. The virus spreading to one vertex is independent of the virus spreading to another vertex within the same time step. For example, if a vertex v is infected and is adjacent to two uninfected vertices x and y, both x and y can become infected in the same time step with probability p^2.

5. Throughout this project the symbols we use are the following:

   - t denotes the time step.
   - X_t is our stochastic random variable that represents the total number of infected vertices at time t.

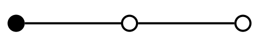
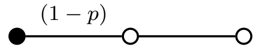
Now let's look at an example of a virus spreading on the graph P_3.

Example 3.1.1. Consider the following path P_3.

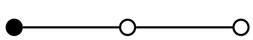
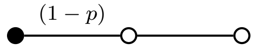
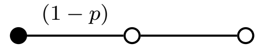


In this graph the root node is at the endpoint. This graph has 3 vertices so the random variable X_t can equal 1, 2, or 3.

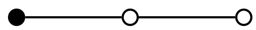
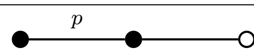
Let's look at the probability that at time step 1 the graph has 1 infected vertex, denoted $P(X_1 = 1)$. This is the probability that after 1 time step the virus didn't spread to any other vertices. The only vertex that the virus could have spread to in one time step is the center vertex since it is the only vertex adjacent to the root node. Therefore $P(X_1 = 1) = (1 - p)$. This probability is illustrated in the following table.

$t = 0$	$t = 1$	Probability
	$(1-p)$ 	$(1-p)$

Now let's look at $P(X_2 = 1)$. This is the probability that after 2 time steps the virus still hasn't spread from the root node. Two time steps gives the virus 2 chances to spread to the adjacent vertex so $P(X_2 = 1) = (1 - p)^2$. The following table illustrates this probability.

$t = 0$	$t = 1$	$t = 2$	Probability
	$(1-p)$ 	$(1-p)$ 	$(1-p)^2$

Now let's look at $P(X_1 = 2)$. This is the probability that in one time step the virus spreads to one more vertex. Since there is only one adjacent vertex this probability is $P(X_1 = 2) = p$. As illustrated in the following table.

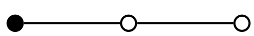
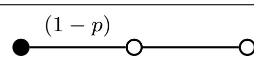
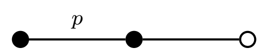
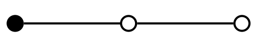
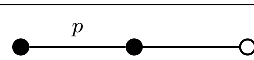
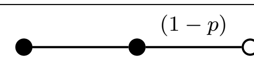
$t = 0$	$t = 1$	Probability
		p

Now let's look at $P(X_2 = 2)$. This is the probability that after 2 times steps the graph has two infected vertices. There are two ways this can happen;

<u>Case 1</u>: The virus doesn't spread in the first time step and then does spread in the second time step.

<u>Case 2</u>: The virus spreads in the first time step and then doesn't spread in the second time step.

The following table illustrates these cases.

Case	$t = 0$	$t = 1$	$t = 2$	Probability
Case 1				$p(1-p)$
Case 2				$p(1-p)$

Adding these cases together we get $P(X_2 = 2) = p(1-p) + p(1-p) = 2p(1-p)$.

Now let's look at $P(X_1 = 3)$. This is the probability that after 1 time step there are 3 infected vertices. The virus can only spread to vertices adjacent to an infected vertex in a time step. So after one time step this graph can have a maximum of 2 vertices infected. Therefore $P(X_1 = 3) = 0$.

Now consider $P(X_2 = 3)$. This is the probability that after two time steps there are 3 infected vertices. This means that the virus has to spread in each time step so $P(X_2 = 3) = p^2$. The following table illustrates this probability.

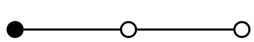
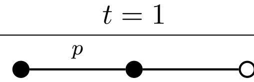
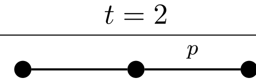
$t = 0$	$t = 1$	$t = 2$	Probability
			p^2

Table 3.1.1 arranges all of these probabilities.

	$P(X_t = 1)$	$P(X_t = 2)$	$P(X_t = 3)$
$t = 0$	1	0	0
$t = 1$	$(1-p)$	p	0
$t = 2$	$(1-p)^2$	$2p(1-p)$	p^2
$t = 3$	$(1-p)^3$	$3p(1-p)^2$	$1 - [(1-p)^3 + 3p(1-p)^2]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t = n$	$(1-p)^n$	$np(1-p)^{n-1}$	$1 - [(1-p)^n + np(1-p)^{n-1}]$

Table 3.1.1. $P(X_t = i)$ for the path P_3.

From this Table 3.1.1 we can see how finding $P(X_t = i)$ for larger values of t helps us see a pattern in the probabilities for each i. Noticing the pattern in each column will lead us to the probability distribution that describes the number of infected vertices at time t for the path P_3.

3.2 Virus Spread using Markov Chains

In this section we look at a different method for finding probability distributions of graphs using Markov chains. We set up a sequence of steps that illustrate our Markov chain method. These steps will be demonstrated using an example.

The Markov Chain Method

Step 1: Compute the one-step transition probabilities p_{ij} and form the transition matrix, T.

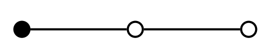
Step 2: Compute the eigenvector matrix Q and find Q^{-1}.

Step 3: Compute the Jordan Form matrix J and find J^t.

Step 4: Solve $T^t = QJ^tQ^{-1}$ to find T^t.

Step 5: Multiply T^t by the starting distribution $\mathbf{u}$ to obtain the probability distribution.

Example 3.2.1. Consider the path P_3 as shown below.



Step 1: Compute the one-step transition probabilities p_{ij} and form the transition matrix, T.

This graph has 3 vertices and therefore 3 states.

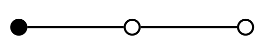
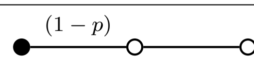
State 1	$X_t = 1$	One infected vertex
State 2	$X_t = 2$	Two infected vertices
State 3	$X_t = 3$	Three infected vertices

So we will get a 3×3 transition matrix of the form

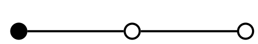
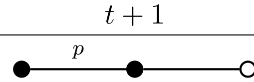
$$T = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}.$$

Recall from Section 2.4 that the transition matrix of a Markov chain contains the one-step transition probabilities denoted p_{ij}. Where $p_{ij} = P(X_{t+1} = j | X_t = i)$ is the probability of moving from state i to state j in one time step. Now let's look at the elements of T for our graph P_3.

$\underline{p_{11} = P(X_{t+1} = 1 | X_t = 1)}$ is the probability that the graph starts in state 1 at time t and in the next time step is still in state 1. This means that in one time step the virus didn't spread from the root node to the adjacent vertex. Therefore $p_{11} = (1-p)$. The following table illustrates this computation.

p_{ij}	t	$t+1$	Probability
p_{11}			$(1-p)$
	state 1	state 1	

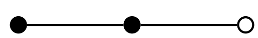
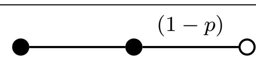
$\underline{p_{12} = P(X_{t+1} = 2 | X_t = 1)}$ is the probability that the graph starts in state 1 at time t and in the next time step is in state 2. This means that in one time step the virus spread from the root node to the middle vertex. So $p_{12} = p$. The following table illustrates this.

p_{ij}	t	$t+1$	Probability
p_{12}			p
	state 1	state 2	

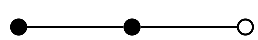
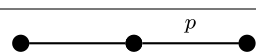
$\underline{p_{13} = P(X_{t+1} = 3 | X_t = 1)}$ is the probability that the graph starts in state 1 at time t and in the next time step is in state 3. This means that the virus must spread from the root node to all of the other vertices in one time step. The virus can only move to the next adjacent vertex in a time step so $p_{13} = 0$.

$p_{21} = P(X_{t+1} = 1|X_t = 2)$ is the probability that the graph starts in state 2 at time t and in the next time step is in state 1. This is the probability that the graph went from having 2 vertices infected to having 1. A vertex cannot be uninfected so $p_{21} = 0$. This will also be the case for p_{31} and p_{32} since the number of infected vertices cannot decrease.

$p_{22} = P(X_{t+1} = 2|X_t = 2)$ is the probability that the graph starts in state 2 at time t and in the next time step is in state 2. This means that the graph has 2 vertices infected at time t and in the next time step still has 2 infected. So $p_{22} = (1 - p)$. The following table illustrates this.

p_{ij}	t	$t+1$	Probability
p_{22}			$(1-p)$
	state 2	state 2	

$p_{23} = P(X_{t+1} = 3|X_t = 2)$ this is the probability that the graph starts in state 2 at time t and in the next time step is in state 3. This probability is similar to p_{12} in that the virus spreads to the only adjacent vertex. So $p_{23} = p$. The following table illustrates this computation.

p_{ij}	t	$t+1$	Probability
p_{23}			p
	state 2	state 3	

$p_{33} = P(X_{t+1} = 3|X_t = 3)$ is the probability that the graph starts in state 3 at time t and in the next time step is still in state 3. Since there are no more vertices to be infected on this graph and no vertex can become uninfected $p_{33} = 1$.

When we arrange these probabilities in the matrix T we get

$$T = \begin{pmatrix} (1-p) & p & 0 \\ 0 & (1-p) & p \\ 0 & 0 & 1 \end{pmatrix}.$$

Step 2: Compute the eigenvector matrix Q and find Q^{-1}.

Now that we have the transition matrix of one-step transition probabilities by Theorem 2.4.3 we can look at T^t to find the t-step probabilities, or the probabilities that the Markov chain moves from state i to state j in t time steps; denoted $p_{ij}^{(t)}$. In order to find T^t we can look at the Jordan Normal Form of T since $T^t = QJ^tQ^{-1}$ (see Theorem 2.2.15 and Definition 2.2.16).

Let's start by finding the matrix Q of T. By Theorem 2.2.7 we know that because T is upper triangular the eigenvalues of T lie on the diagonal. Therefore the eigenvalues are $\lambda = (1-p)$, $\lambda = (1-p)$, and $\lambda = 1$. The matrix Q is made up of the eigenvectors of T so we will start by finding the eigenvector associated with $\lambda = 1$. We want $(T - 1I)v_1 = 0$. So

$$(T-1I)v_1 = \begin{pmatrix} -p & p & 0 \\ 0 & -p & p \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

implies

$$-px_1 + px_2 = 0$$

and

$$-px_2 + px_3 = 0.$$

The solution to these two equations is $x_1 = x_2 = x_3$. Therefore the eigenvector associated with the eigenvalue $\lambda = 1$ is

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Now we will find the eigenvectors associated with the eigenvalue $\lambda = (1-p)$. We want $(T - (1-p)I)v_2 = 0$. So

$$(T-(1-p)I)v_2 = \begin{pmatrix} 0 & p & 0 \\ 0 & 0 & p \\ 0 & 0 & p \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

implies

$$px_2 = 0$$

and

$$px_3 = 0.$$

Therefore $x_2 = x_3 = 0$ and x_1 can be anything. So the only eigenvector associated with $\lambda = (1-p)$ is

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

However, there are two eigenvalues that equal $(1-p)$ so we need to find a generalized eigenvector such that $(T-(1-p)I)v_3 = v_2$. Therefore

$$(T-(1-p)I)v_3 = \begin{pmatrix} 0 & p & 0 \\ 0 & 0 & p \\ 0 & 0 & p \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

implies

$$px_2 = 1$$

and

$$px_3 = 0.$$

So $x_2 = \frac{1}{p}$, $x_3 = 0$, and x_1 can be anything.Therefore the generalized eigenvector based on the eigenvector v_2 is

$$v_3 = \begin{pmatrix} 0 \\ \frac{1}{p} \\ 0 \end{pmatrix}.$$

When we arrange the eigenvectors and generalized eigenvectors of T into a matrix we get

$$Q = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & \frac{1}{p} \\ 1 & 0 & 0 \end{pmatrix} \text{ where } Q^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & p & -p \end{pmatrix}.$$

Step 3: Compute the Jordan Form matrix J and find J^t.

We saw in Step 2 that T has two linearly independent eigenvectors. Therefore the Jordan Form matrix J will be made up of two Jordan blocks by Theorem 2.2.15. T also has a generalized eigenvector based on the eigenvector v_2 which is associated with $\lambda = (1-p)$. Therefore the Jordan blocks are

$$J_1 = (1) \text{ and } J_2 = \begin{pmatrix} (1-p) & 1 \\ 0 & (1-p) \end{pmatrix}.$$

So we have

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-p) & 1 \\ 0 & 0 & (1-p) \end{pmatrix}.$$

Using *Mathematica* we can find J^t,

$$J^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-p)^t & t(1-p)^{t-1} \\ 0 & 0 & (1-p)^t \end{pmatrix}.$$

Step 4: Solve $T^t = QJ^tQ^{-1}$ to find T^t.

$$T^t = QJ^tQ^{-1} = Q = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & \frac{1}{p} \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-p)^t & t(1-p)^{t-1} \\ 0 & 0 & (1-p)^t \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & p & -p \end{pmatrix}$$

$$= \begin{pmatrix} (1-p)^t & tp(1-p)^{t-1} & 1-(1-p)^t - tp(1-p)^{t-1} \\ 0 & (1-p)^t & 1-(1-p)^t \\ 0 & 0 & 1 \end{pmatrix}.$$

Step 5: Multiply T^t by the starting distribution $\mathbf{u}$ to obtain the probability distribution.

Now we can find the probability distribution that describes the number of infected vertices at time t on the graph P_3. The starting distribution of P_3 is $\mathbf{u} = (P(X_0 = 1), P(X_0 = 2), P(X_0 = 3)) = (1, 0, 0)$ since at time 0 there is only one vertex infected, the root node. By Theorem 2.4.5 the vector that describes the probability distribution of this graph is

$$\mathbf{u}^{(t)} = \mathbf{u}T^t = (1,0,0) \cdot \begin{pmatrix} (1-p)^t & tp(1-p)^{t-1} & 1-(1-p)^t - tp(1-p)^{t-1} \\ 0 & (1-p)^t & 1-(1-p)^t \\ 0 & 0 & 1 \end{pmatrix}$$
$$= ((1-p)^t, tp(1-p)^{t-1}, 1-(1-p)^t - tp(1-p)^{t-1}).$$

Note $\mathbf{u}(t) = (P(X_t = 1), P(X_t = 2), P(X_t = 3))$.

Therefore the probability distribution of P_3 is

$$P(X_t = i) = \begin{cases} (1-p)^t, & \text{if } i = 1 \\ tp(1-p)^{t-1}, & \text{if } i = 2 \\ 1-[(1-p)^t + tp(1-p)^{t-1}], & \text{if } i = 3 \end{cases}.$$

Recall from Table 3.1.1 that we found the same proabability distribution.

4
Linear Algebra of the Markov Chain Model

In this chapter we present some theorems that will be useful in understanding the rest of this project. We see a theorem that describes what the eigenvalues look like for the transition matrix of every graph, a theorem that describes two eigenvectors of the transition matrix T of every graph, and theorems that describe what J^t looks like for J that contains Jordan blocks J_i that are 3×3 or smaller.

4.1 Eigenvalues and Eigenvectors of T

Theorem 4.1.1. *The eigenvalues for the Transition Matrix T of a graph with n vertices are 1 and $(1-p)^m$ where $m \in \mathbb{N}$.*

Proof. Let T be the Transition Matrix of a graph with n vertices. All the entries of T are defined as $p_{ij} = P(X_{t+1} = j | X_t = i)$ which is the probability that that graph is in state i and in the next time step is in state j. Consider p_{ij} where $i > j$. This is the probability that the number of infected vertices on the graph decreases. Because a vertex cannot be uninfected $p_{ij} = 0$ when $i > j$. Therefore T is an upper triangular matrix. By Theorem 2.2.7 we know that the eigenvalues of an upper triangular matrix lie on the diagonal. Therefore the eigenvalues of T are all the entries p_{ij} where $i = j$. These are the probabilities that the graph starts with i infected vertices and in the next time step still has i infected vertices. This means that the virus didn't spread from any of the infected vertices. This probability is equal to $(1-p)^m$ where m is the number of edges connected an infected vertex to an uninfected vertex. Note that $p_{nn} = (1-p)^0 = 1$ since there are no more vertices to be infected. □

For example, recall the transition matrix of the path P_3 that we found in Example 3.2.1 as shown below.

$$T = \begin{pmatrix} (1-p) & p & 0 \\ 0 & (1-p) & p \\ 0 & 0 & 1 \end{pmatrix}$$

T is upper triangular so by Theorem 2.2.7 the eigenvalues are $\lambda = (1-p)$ and $\lambda = 1$.

Theorem 4.1.2. *The basis of eigenvectors for the transition matrix T of a graph with n vertices contains the vectors*

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad \textit{and} \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Proof. By Theorem 4.1.1 we know that $\lambda_1 = 1$ is an eigenvalue of every graph. For each row i of a Markov chain we know that the elements of that row add up to 1 since they cover all the possibilities of moving from state i to another state. Therefore when we solve $(T - 1I)$ we are subtracting 1 from each row so the elements of each row now add up to 0. Therefore the first entry of v_2 can be anything. So

$$(T-1I)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = 0 \quad \text{so,} \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ is an eigenvalue of } T.$$

Because T is upper triangular we know that the element p_{11} is an eigenvalue of T by Theorem 2.2.7. So when we find $(T - p_{11}I)v_2 = 0$ we have that the matrix $(T - p_{11}I)$ is a matrix with zeros in the first column. Therefore

$$(T - p_{11}I)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = 0 \quad \text{so,} \quad v_2 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} \text{ is an eigenvalue of } T.$$

□

Recall from Example 3.2.1 that the eigenvector of T associated with the eigenvalue $\lambda = 1$ is

$$v_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}.$$

And the eigenvector associated witht the eigenvalue $\lambda = p_{11} = (1-p)$ is

$$v_2 = \begin{pmatrix}1\\0\\0\end{pmatrix}.$$

4.2 The Jordan Block J_i^t

The following theorems show what the Jordan block J_i^t looks like for J_i that are 3×3 or smaller. From these theorems we can determine what J^t looks like for J that is made up of Jordan blocks that are 3×3 or smaller.

Theorem 4.2.1. *If the Jordan block* $J_i = \begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}$ *then* $J_i^t = \begin{pmatrix}\lambda^t & t\lambda^{t-1}\\ 0 & \lambda^t\end{pmatrix}$ *where t is a positive integer.*

Proof. We will use induction on t.

Base Case: Let $t = 1$. Then $J_i^1 = \begin{pmatrix}\lambda^1 & 1\cdot\lambda^{1-1}\\ 0 & \lambda^1\end{pmatrix} = \begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} = J_i$.

Induction Hypothesis: Assume $J_i^t = \begin{pmatrix}\lambda^t & t\lambda^{t-1}\\ 0 & \lambda^t\end{pmatrix}$ for some $t \geq 1$.

Induction Step: Suppose the induction hypothesis holds. Then

$$\begin{aligned} J_i^{t+1} &= J_i^t \cdot J_i \\ &= \begin{pmatrix}\lambda^t & t\lambda^{t-1}\\ 0 & \lambda^t\end{pmatrix}\cdot\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} \\ &= \begin{pmatrix}\lambda^{t+1} & (t+1)\lambda^{(t+1)-1}\\ 0 & \lambda^{t+1}\end{pmatrix}. \end{aligned}$$

□

Theorem 4.2.2. *If the Jordan block* $J_i = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$ *then,* $J_i^t = \begin{pmatrix} \lambda^t & t\lambda^{t-1} & T_{t-1}\lambda^{t-2} \\ 0 & \lambda^t & t\lambda^{t-1} \\ 0 & 0 & \lambda^t \end{pmatrix}$ *where* t *is a positive integer and where* T_{t-1} *is the* $(t-1)$*th term in the Triangular Number Sequence such that* $T_0 = 0$, $T_1 = 1$, $T_2 = 3$, *etc.*

Proof. We will use induction on t.

Base Case: Let $t = 1$. Then

$$J_i^1 = \begin{pmatrix} \lambda^1 & 1 \cdot \lambda^0 & T_0 \\ 0 & \lambda^1 & 1 \cdot \lambda^0 \\ 0 & 0 & \lambda^1 \end{pmatrix} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = J_i.$$

Induction Hypothesis: Assume $J_i^t = \begin{pmatrix} \lambda^t & t\lambda^{t-1} & T_{t-1}\lambda^{t-2} \\ 0 & \lambda^t & t\lambda^{t-1} \\ 0 & 0 & \lambda^t \end{pmatrix}$.

Induction Step: Then

$$J_i^{t+1} = J_i^t \cdot J_i = \begin{pmatrix} \lambda^t & t\lambda^{t-1} & T_{t-1}\lambda^{t-2} \\ 0 & \lambda^t & t\lambda^{t-1} \\ 0 & 0 & \lambda^t \end{pmatrix} \cdot \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda^{t+1} & (t+1)\lambda^t & T_t\lambda^{t-1} \\ 0 & \lambda^{t+1} & (t+1)\lambda^t \\ 0 & 0 & \lambda^{t+1} \end{pmatrix}.$$

□

Recall from Example 3.2.1 that the matrix J for the path P_3 is made up of two Jordan blocks

$$J_1 = (1) \text{ and } J_2 = \begin{pmatrix} (1-p) & 1 \\ 0 & (1-p) \end{pmatrix}.$$

So,

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix}.$$

Therefore we have that

$$J^t = \begin{pmatrix} J_1^t & 0 \\ 0 & J_2^t \end{pmatrix}$$

where

$$J_1^t = (1^t) = 1 \text{ and } J_2^t = \begin{pmatrix} (1-p)^t & t(1-p)^{t-1} \\ 0 & (1-p)^t \end{pmatrix}.$$

So,

$$J^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-p)^t & t(1-p)^{t-1} \\ 0 & 0 & (1-p)^t \end{pmatrix}.$$

5

Virus Spread on Paths, Stars, Cycles, and Complete Graphs

In this chapter we use Markov chains to study the spread of a virus on paths, stars, cycles, and complete graphs. We find the transition matrix T for n-paths, n-stars, n-cycles, and complete n-graphs. We also find the Jordan Normal Form matrices J and Q for paths, stars, and cycles.

5.1 Virus Spread on Paths

In this section we analyze the spread of a virus on paths where the root node is positioned at one of the endpoints.

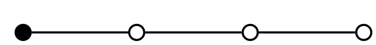
5.1.1 Transition Matrices for the Paths P_3 and P_4

We now look at the transition matrices for the paths P_3 and P_4. These examples will help us formulate the transition matrix for any n-path.

In Example 3.2.1 we found that the transition matrix for the path P_3 is

$$T = \begin{pmatrix} (1-p) & p & 0 \\ 0 & (1-p) & p \\ 0 & 0 & 1 \end{pmatrix}.$$

Example 5.1.1. In this example we will find the transition matrix for the path P_4 as shown below.



This graph has 4 vertices and therefore 4 states; state 1 = 1 vertex infected, state 2 = 2 vertices infected, etc. The following table illustrates the transition probabilities p_{ij} of the transition matrix for the path P_4.

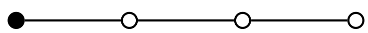
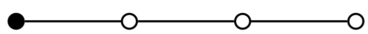
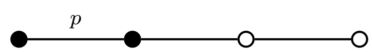
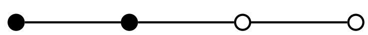
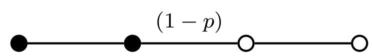
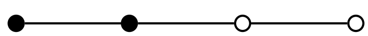
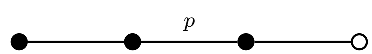
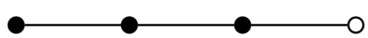
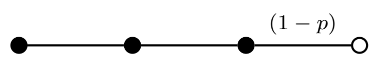
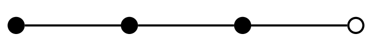
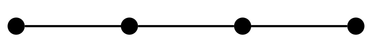
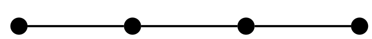
p_{ij}	Time $= t$	Time $= t+1$	$p_{ij} =$
p_{11}			$(1-p)$
p_{12}			p
p_{22}			$(1-p)$
p_{23}			p
p_{33}			$(1-p)$
p_{34}			p
p_{44}			1

Table 5.1.1. Elements of the transition matrix for the path P_4

For example, consider p_{23}. This is the probability that the graph starts in state 2 at time t and in the next time step is in state 3. This means that in one time step the virus spread to one more vertex. There is only one possible vertex for the virus to spread to so this probability is $p_{23} = p$.

Note that the elements below the diagonal (p_{21}, p_{31}, p_{32}, p_{41}, p_{42}, and p_{43}) are all the probabilities where the number of infected vertices decreases. Since a vertex cannot become uninfected these probabilities are 0.

The elements p_{13}, p_{14}, and p_{24} are all the probabilites where the virus must spread to two or more vertices in a time step. On a path the virus can only spread to a maximum of one vertex in a time step so these probabilities are 0.

When we arrange these probabilities into a matrix we get

$$T = \begin{pmatrix} (1-p) & p & 0 & 0 \\ 0 & (1-p) & p & 0 \\ 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

5.1.2 Transition Matrix for an n-Path

Theorem 5.1.2. *The Transition Matrix T for a path with n vertices P_n is the $n \times n$ matrix*

$$T = \begin{pmatrix} (1-p) & p & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & (1-p) & p \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Proof. Let T be the transition matrix for a path P_n. Consider the elements p_{ij} where $i > j$. Note $p_{ij} = P(X_{t+1} = j | X_t = i)$ is the probability that the graph is in state j given that in the previous time step the graph was in state i. When $i > j$, p_{ij} is the probability that the number of infected vertices decreases. Since a vertex cannot become uninfected this probability is 0.

Consider p_{ij} where $i = j \neq n$. This is the probability that the graph stays at i infected vertices during a time step. On a path there is only one vertex that the virus can move to in a time step so when $i = j \neq n$, $p_{ij} = (1-p)$. Consider p_{ij} where $i = j = n$. Then $p_{ij} = 1$ since there are no more vertices for the virus to spread to and no vertex can be uninfected.

Consider p_{ij} where $j = i + 1$. This is the probability that the virus spreads to one more vertex in a time step. On a path there is only one vertex that the virus can move to in a time step so $p_{ij} = p$.

Consider p_{ij} where $j \geq i+2$. This is the probability that the virus spreads to 2 or more vertices in a time step. On a path there is only one possible vertex that can be infected in a time step so $p_{ij} = 0$. □

5.1.3 *The Matrix Q for the Paths P_3 and P_4*

In this section we find the matrix Q for the paths P_3 and P_4. These examples will help us formulate the matrix Q for any n-path.

In Example 3.2.1 we found that the Jordan Form matrix Q for the path P_3 is

$$Q = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & \frac{1}{p} \\ 1 & 0 & 0 \end{pmatrix}$$

Example 5.1.3. In this example we find Q for the path P_4. The matrix Q is made up of the eigenvectors of T. We know that T has four eigenvalues $\lambda = 1$ and three that are $\lambda = (1-p)$.

By Theorem 4.1.2 we know that the eigenvector associated with $\lambda = 1$ is

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

And we know that an eigenvector associated with $\lambda = p_{11} = (1-p)$ is

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

When we solve for v_2 we get

$$(T-(1-p)I)v_2 = \begin{pmatrix} 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \\ 0 & 0 & 0 & -p \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0.$$

This implies

$$px_2 = 0$$

$$px_3 = 0$$

$$px_4 = 0$$

$$-px_4 = 0$$

Therefore $x_2 = x_3 = x_4 = 0$ and x_1 can be anything. So the only eigenvector associated with $\lambda = (1-p)$ is v_2. However, $\lambda = (1-p)$ has multiplicity 3 so we need to find two generalized eigenvectors such that $(T-(1-p)I)v_3 = v_2$ and $(T-(1-p)I)v_4 = v_3$. We will start by finding v_3.

$$(T-(1-p)I)v_3 = \begin{pmatrix} 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \\ 0 & 0 & 0 & p \end{pmatrix} \cdot v_3.$$

Let $v_3 = \begin{pmatrix} 0 \\ \frac{1}{p} \\ 0 \\ 0 \end{pmatrix}$ Then

$$(T-(1-p)I)v_3 = \begin{pmatrix} 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \\ 0 & 0 & 0 & p \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \frac{1}{p} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = v_2.$$

Hence v_3 is a generalized eigenvector of T.

Now we will find v_4.

$$(T-(1-p)I)v_4 = \begin{pmatrix} 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \\ 0 & 0 & 0 & p \end{pmatrix} \cdot v_4.$$

Let $v_4 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{p^2} \\ 0 \end{pmatrix}$. Then

$$(T-(1-p)I)v_4 = \begin{pmatrix} 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \\ 0 & 0 & 0 & p \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ \frac{1}{p^2} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{p} \\ 0 \\ 0 \end{pmatrix} = v_3.$$

When we arrange these eigenvectors and generalized eigenvectors into a matrix we get

$$Q = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & \frac{1}{p} & 0 \\ 1 & 0 & 0 & \frac{1}{p^2} \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

5.1.4 The Matrix Q for an n-Path

Lemma 5.1.4. *Let T be the transition matrix for a path with n vertices P_n. Then T has two distinct eigenvalues $\lambda = 1$ and $\lambda = (1-p)$ where $\lambda = (1-p)$ has multiplicity $n-1$ and the only eigenvector associated with $\lambda = (1-p)$ is*

$$v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Proof. Let T be the transition matrix of a path with n vertices, P_n. The matrix T is upper triangualr so by Theorem 2.2.7 the eigenvalues lie on the diagonal. Recall from Theorem 5.1.2 that the diagonal entries of T are 1 and $(1-p)$. Therefore the eigenvalues are $\lambda = 1$ and $\lambda = (1-p)$ where $(1-p)$ has multiplicity $n-1$. To find the eigenvectors associated with $\lambda = (1-p)$ we solve $(T-(1-p)I)v = 0$. So

$$(T-(1-p)I)v = \begin{pmatrix} 0 & p & 0 & \dots & 0 \\ 0 & 0 & p & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & p \\ 0 & \dots & \dots & 0 & -p \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = 0$$

implies

$$\begin{aligned} px_2 &= 0 \\ px_3 &= 0 \\ &\vdots \\ px_n &= 0 \\ -px_n &= 0. \end{aligned}$$

Therefore $x_2 = x_3 = x_4 = ... = x_n = 0$ and x_1 can be anything. So the only eigenvector associated with $\lambda = (1-p)$ is

$$v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

□

Theorem 5.1.5. *The matrix Q for the transition matrix T of a path with n vertices P_n is the $n \times n$ matrix*

$$Q = \begin{pmatrix} 1 & 1 & 0 & \dots & \dots & 0 \\ 1 & 0 & \frac{1}{p} & \ddots & & \vdots \\ 1 & 0 & 0 & \frac{1}{p^2} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \frac{1}{p^{n-2}} \\ 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Proof. Let T be the transition matrix of the path with n vertices, P_n. The matrix Q is made up of the eigenvectors of T. By Theorem 4.1.2 we know that

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ and } u_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

are eigenvectors of the transition matrix T where v_1 is associated with $\lambda = 1$ and u_1 is associated with $\lambda = p_{11} = (1-p)$. By Lemma 5.1.4 u_1 is the only eigenvector associated with $\lambda = (1-p)$. However, $\lambda = (1-p)$ has multiplicity $n-1$ so there is a length $n-1$ chain of generalized eigenvectors based on the eigenvector u_1. This chain of generalized eigenvectors is a set $\{u_1, u_2, ..., u_{n-1}\}$ such that

$$\begin{gathered}(T-(1-p)I)u_2 = u_1\\(T-(1-p)I)u_3 = u_2\\\vdots\\(T-(1-p)I)u_{n-1} = u_{n-2}.\end{gathered}$$

We will prove what the generalized eigenvectors look like through induction on k.

Base Case: We know that the first eigenvector in the chain is

$$u_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}\frac{1}{p^{1-1}}\\0\\\vdots\\0\end{pmatrix}.$$

Induction Hypothesis: Assume u_k is the kth generalized eigenvector where $k \neq n-1$. Assume

$$u_k = \begin{pmatrix}0\\\vdots\\0\\\frac{1}{p^{k-1}}\\0\\\vdots\\0\end{pmatrix}$$

where $\frac{1}{p^{k-1}}$ is the kth entry in the vector u_k.

Induction Step: Suppose the induction hypothesis holds. Then solving for the next generalized eigenvector in the chain, u_{k+1},means solving

$$(T-(1-p)I)u_{k+1} = \begin{pmatrix} 0 & p & 0 & \dots & 0 \\ 0 & 0 & p & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & p \\ 0 & \dots & \dots & 0 & -p \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{1}{p^{k-1}} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = u_k.$$

Solving for the elements of u_{k+1} we get the following equations

$$0x_1 + px_2 + 0x_3 + ... + 0x_n = 0$$
$$\vdots$$
$$0x_1 + ... + 0x_{k-1} + px_k + 0x_{k+1} + ... + 0x_n = 0$$
$$0x_1 + ... + 0x_k + px_{k+1} + 0x_{k+2} + ... + 0x_n = \frac{1}{p^{k-1}}$$
$$0x_1 + ... + 0x_{k+1} + px_{k+2} + 0x_{k+3} + ... + 0x_n = 0$$
$$\vdots$$
$$0x_1 + ... + 0x_{n-1} + px_n = 0$$

Let $x_{k+1} = \frac{1}{p^k}$ and let every other $x_i = 0$. Then $0x_1 + ... + 0x_k + px_{k+1} + 0x_{k+2} + ... + 0x_n = px_{k+1} = \frac{p}{p^k} = \frac{1}{p^{k-1}}$ and all other equations equal 0. Hence all the equations are satisfied. Therefore

$$u_{k+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{1}{p^k} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $\frac{1}{p^k}$ is the $k+1th$ element of the vector u_{k+1}. □

5.1.5 Jordan Form Matrix J for the Paths P_3 and P_4

In this section we find the matrix J for the paths P_3 and P_4. These examples will help us formulate the matrix J for any n-path.

In Example 3.2.1 we found that the Jordan Form matrix J for the path P_3 is

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (1-p) & 1 \\ 0 & 0 & (1-p) \end{pmatrix}$$

Example 5.1.6. In this example we find the matrix J for the path P_4. In Example 5.1.1 we found that the transition matrix for P_4 is

$$T = \begin{pmatrix} (1-p) & p & 0 & 0 \\ 0 & (1-p) & p & 0 \\ 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

By Theorem 2.2.7 we know that the eigenvalues of T lie on the diagonal. So, T has two distinct eigenvalues $\lambda = 1$ and $\lambda = (1-p)$ where $\lambda = (1-p)$ has multiplicity 3. By Lemma 5.1.4 we know that there is only one eigenvector associated with $\lambda = (1-p)$. Therefore T has two linearly independent eigenvectors, so J is made up of two Jordan blocks by Theorem 2.2.15

$$J_1 = (1) \text{ and } J_2 = \begin{pmatrix} (1-p) & 1 & 0 \\ 0 & (1-p) & 1 \\ 0 & 0 & (1-p) \end{pmatrix}.$$

Therefore we know

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-p) & 1 & 0 \\ 0 & 0 & (1-p) & 1 \\ 0 & 0 & 0 & (1-p) \end{pmatrix}$$

5.1.6 Jordan Form Matrix J for an n-Path

Theorem 5.1.7. *The Jordan Form matrix J for the transition matrix T of a path with n vertices P_n is the $n \times n$ matrix*

$$J = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (1-p) & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & (1-p) \end{pmatrix}.$$

Proof. Let T be the transition matrix for the path P_n. Let J be the Jordan Form matrix of T. We know T is upper triangular so by Theorem 2.2.7 the eigenvalues lie on the diagonal. Therefore by Theorem 5.1.2 the eigenvalues are $\lambda = 1$ and $\lambda = (1-p)$ where $\lambda = (1-p)$ has multiplicity $n-1$. By Lemma 5.1.4 we know that there is only one eigenvector associated with $\lambda = (1-p)$. Therefore T has two linearly independent eigenvectors so J is made up of two Jordan blocks by Theorem 2.2.15.

$$J_1 = (1) \text{ and } J_2 = \begin{pmatrix} (1-p) & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & (1-p) & 1 \\ 0 & \dots & \dots & 0 & (1-p) \end{pmatrix}$$

where J_2 is an $n-1 \times n-1$ matrix. Therefore

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} = J = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (1-p) & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & (1-p) \end{pmatrix}.$$

□

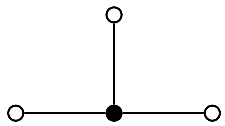
5.2 Virus Spread on Star Graphs

In this section we analyze the spread of a virus on star graphs where the root node is at the center vertex.

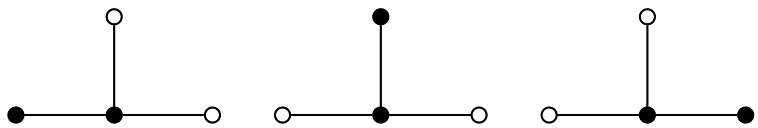
5.2.1 Transition Matrix for the Star Graph S_4

In this section we find the transition matrix for the star graph S_4. This example will help us formulate the transition matrix for any n-star.

Example 5.2.1. In this example we find the transition matrix of the star graph S_4 as shown below.

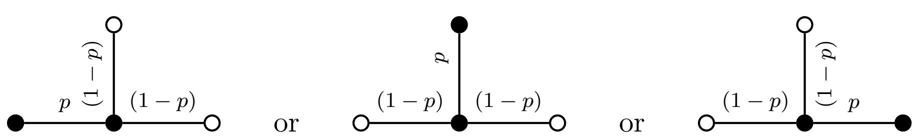


This graph has 4 vertices and therefore 4 states. Consider the state where 2 vertices are infected. This state can be represented three ways, as shown below.



We consider each of these representations the same state since each graph can be rearranged to look like the other. We also get the same transition probabilities no matter which representation we choose. For example, consider the probability p_{22} for each representation. In each one there are two uninfected vertices that the virus could spread to so for each this probability is $(1-p)^2$. Therefore $p_{22} = (1-p)^2$ for the graph S_4. This is also the case for the state where 3 vertices are infected.

Now, let's find the transition probabilities for S_4. Consider p_{12}. This is the probability that the virus spread from the root node to just one of the adjacent vertices. This can happen three ways, as shown below.



Therefore this probability is $p_{12} = 3p(1-p)^2$. Another way to think of this probability is that out of the 3 vertices the the virus can spread to we want to choose 1. If we consider the virus spreading to a vertex a success we can use the binomial distribution to find this probability. Therefore $\binom{3}{1} p^1(1-p)^{3-1} = 3p(1-p)^2$.

We can use the same method to find p_{13}. This is the probability that the virus spreads from the root node to two more vertices. Hence the virus spreads to two out of the three uninfected vertices. Therefore $p_{13} = \binom{3}{2} p^2(1-p)^{3-2} = 3p^2(1-p)$.

Table 5.2.1 illustrates the rest of the transition probabilities for the star graph S_4.

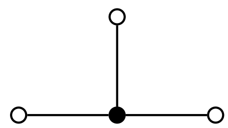
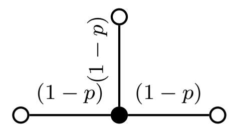
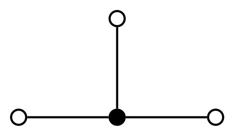
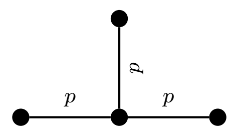
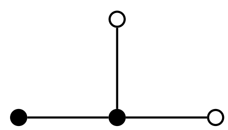
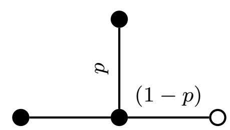
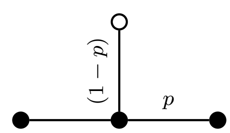
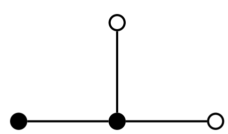
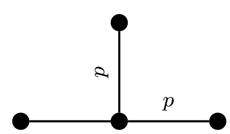
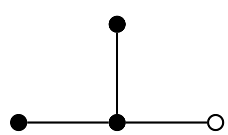
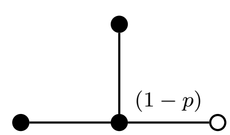
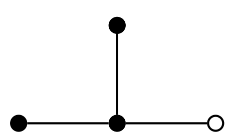
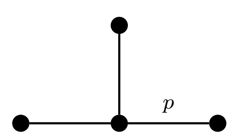
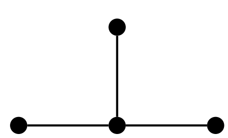
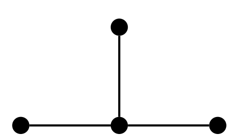
p_{ij}	Time$=t$	Time$=t+1$	$p_{ij} =$
p_{11}			$(1-p)^3$
p_{14}			p^3
p_{23}		 or 	$2p(1-p)$
p_{24}			p^2
p_{33}			$(1-p)$
p_{34}			p
p_{44}			1

Table 5.2.1. Elements of the transition matrix for the star graph S_4

Note that the elements below the diagonal (p_{21}, p_{31}, p_{32}, p_{41}, p_{42}, and p_{43}) are all the elements where the number of infected vertices decreases. Because of the fact that a virus cannot become uninfected, these probabilities are 0.

When we arrange these probabilities into a matrix we get

$$T = \begin{pmatrix} (1-p)^3 & 3p(1-p)^2 & 3p^2(1-p) & p^3 \\ 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

5.2.2 Transition Matrix for an n-Star

Theorem 5.2.2. *The Transition Matrix T for a star graph with n vertices S_n is the $n \times n$ matrix with elements p_{ij} such that*

$$p_{ij} = \begin{cases} \binom{n-i}{j-i} p^{j-i}(1-p)^{n-j}, & \text{if } i \le j \\ 0, & \text{if } i > j \end{cases}$$

Note, the elements p_{ij} of the transition matrix of a star graph with n vertices follow a binomial distribution with parameters $n-i$ and p.

Proof. Let T be the transition matrix of a star graph with n vertices, S_n. Consider p_{ij} where $i > j$. This is the probability that the number of infected vertices on the graph decreases. A vertex cannot become uninfected so this probability is 0. Consider p_{ij} where $i \le j$. This is the probability that the virus spreads to $j-i$ more vertices. Note that on a star graph all the vertices are adjacent to the center vertex (the root node) so the virus can spread to any number of the uninfected vertices in a time step. The number $n-i$ is the number of uninfected vertices left on the graph. Therefore $\binom{n-i}{j-i}$ is the number of ways the virus can spread to $j-i$ more vertices out of the remaining $n-i$ vertices. The probability $p^{j-i}(1-p)^{n-j}$ is the probability that the virus spreads to $j-i$ more vertices and doesn't spread to the remaining $n-j$ vertices. Therefore $p_{ij} = \binom{n-i}{j-i} p^{j-i}(1-p)^{n-j}$ when $i \le j$. □

5.2.3 The Matrix Q for the Star Graph S_3

In this section we find the matrix Q for the star graph S_3. This example will help us formulate the matrix Q for any n-star.

Example 5.2.3. In this example we find Q for the star S_3. The matrix Q is made up of the eigenvectors of the transition matrix of S_3. Using Theorem 5.2.2 we can find the transition matrix for this graph, as shown below.

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 \\ 0 & (1-p) & p \\ 0 & 0 & 1 \end{pmatrix}$$

T is upper triangular so by Theorem 2.2.7 the eigenvalues lie on the diagonal. Therefore the eigenvalues of T are $\lambda = 1$, $\lambda = (1-p)$, and $\lambda = (1-p)^2$. By Theorem 4.1.2 we know that the eigenvector associated with $\lambda = 1$ is

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

and the eigenvector associated with $\lambda = p_{11} = (1-p)^2$ is

$$v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Now, lets find the eigenvector v_3 associated with $\lambda = (1-p)$. We want $(T - (1-p)I)v_3 = 0$. Therefore

$$(T - (1-p)I)v_3 = \begin{pmatrix} -p(1-p) & 2p(1-p) & p^2 \\ 0 & 0 & p \\ 0 & 0 & p \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

which implies $-p(1-p)x_1 + 2p(1-p)x_2 + p^2x_3 = 0$ and $px_3 = 0$. Let $x_1 = 2$, $x_2 = 1$, and $x_3 = 0$. Then both equations are satisfied. Therefore

$$v_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

When we arrange these eigenvectors in a matrix we get

$$Q = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Using *Mathematica* we can find the matrix Q for S_4 and S_5 as well.

<u>Q for S_4:</u>

$$Q = \begin{pmatrix} 1 & 3 & 3 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

<u>Q for S_5:</u>

$$Q = \begin{pmatrix} 1 & 4 & 6 & 4 & 1 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

These matrices help us form the following conjecture about Q for S_n.

5.2.4 The Matrix Q for an n-Star

Conjecture 5.2.4. *The matrix Q for the transition matrix T of a star graph with n vertices S_n is the $n \times n$ matrix*

$$Q = \begin{pmatrix} \binom{n-1}{0} & \binom{n-1}{1} & \cdots & \cdots & \binom{n-1}{n-1} \\ \binom{n-2}{0} & \binom{n-2}{1} & \cdots & \binom{n-2}{n-2} & 0 \\ \vdots & & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & & \vdots \\ \binom{n-n}{0} & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

5.2.5 Jordan Form Matrix J for the Star Graph S_4

In this section we find the matrix J for the star graph S_4. This example will help us formulate the matrix J for any n-star.

Example 5.2.5. In this example we find J for the star S_4. In Example 5.2.1 we found that the transition matrix for S_4 is

$$T = \begin{pmatrix} (1-p)^3 & 3p(1-p)^2 & 3p^2(1-p) & p^3 \\ 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

By Theorem 2.2.7 we know that the eigenvalues of T lie on the diagonal. So, T has 4 distinct eigenvalues $\lambda = 1$, $\lambda = (1-p)$, $\lambda = (1-p)^2$, and $\lambda = (1-p)^3$. Hence there are 4 linearly independent eigenvectors so by Theorem 2.2.15 J is made up of 4 Jordan blocks.

$J_1 = (1), J_2 = ((1-p)), J_3 = ((1-p)^2)$ and $J_4 = ((1-p)^3)$. So

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-p) & 0 & 0 \\ 0 & 0 & (1-p)^2 & 0 \\ 0 & 0 & 0 & (1-p)^3 \end{pmatrix}$$

5.2.6 Jordan Form Matrix J for an n-Star

Theorem 5.2.6. *The Jordan Form J for the transition matrix T of a star graph with n vertices S_n is the $n \times n$ matrix*

$$J = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & (1-p) & \ddots & & \vdots \\ \vdots & \ddots & (1-p)^2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & (1-p)^{n-1} \end{pmatrix}.$$

Note that the matrix J is diagonal for star graphs.

Proof. Let T be the transition matrix for the star graph S_n. We know that T is upper triangular so by Theorem 2.2.7 the eigenvalues of T are all the elements p_{ii}. By Theorem 5.2.2 we know that $p_{ii} = \binom{n-i}{i-i} p^{i-i}(1-p)^{n-i} = (1-p)^{n-i}$. Therefore there are n distinct eigenvalues of the form $\lambda = (1-p)^{n-i}$ where $i \in \{1, 2, ..., n\}$. Therefore there are n linearly independent eigenvectors so J is made up of n Jordan blocks. Hence

$$J = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & (1-p) & \ddots & & \vdots \\ \vdots & \ddots & (1-p)^2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & (1-p)^{n-1} \end{pmatrix}.$$

□

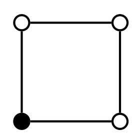
5.3 Virus Spread on Cycles

In this section we analyze the spread of a virus on cycles.

5.3.1 Transition Matrix for the Cycle C_4

In this section we find the transition matrix for the cycle C_4. This example will help us formulate the transition matrix for any n-cycle.

Example 5.3.1. In this example we will find the transition matrix for the cycle C_4 as shown below.



Note on this graph the virus can start from any vertex.

This cycle has 4 vertices and therefore 4 states. So we will get a transition matrix of the form

$$T = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix}.$$

Consider the probability p_{14}. This is the probability that the virus spreads from the root node to all other vertices in one time step. The root node is only adjacent to two vertices so it's impossible for the virus to spread to all 4 vertices in one time step. Therefore $p_{14} = 0$.

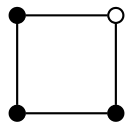
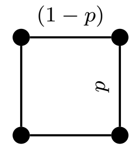
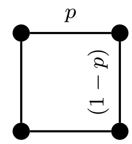
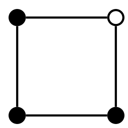
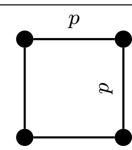
Consider the probability p_{34}. This is the probability that the graph has 3 vertices infected at time t and then spreads to the last vertex in the next time step. The last vertex is adjacent to two infected vertices so this can happen two ways.

Case 1: The virus spreads from one vertex and not the other.

Case 2: The virus spread from both vertices.

The following table illustrates these cases.

Adding these two cases together, we get $p_{34} = 2p(1-p) + p^2$.

Case	Time$= t$	Time$= t+1$	Probability
Case 1		 or 	$2p(1-p)$
Case 2			p^2

Note that the elements below the diagonal (p_{21}, p_{31}, p_{32}, p_{41}, p_{42}, and p_{43}) are all the elements where the number of infected vertices decreases. Because a vertex cannot become uninfected these probabilities are 0. Table 5.3.1 illustrates the rest of the transition probabilities for the cycle C_4.

When we arrange these probabilities into a matrix we get

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 & 0 \\ 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

5.3.2 Transition Matrix for an n-Cycle

Theorem 5.3.2. *The Transition Matrix T for a cycle with n vertices C_n is the $n \times n$ matrix*

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & (1-p)^2 & 2p(1-p) & p^2 \\ \vdots & & & & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

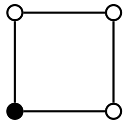
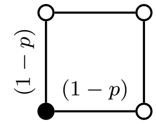
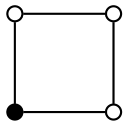
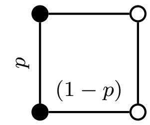
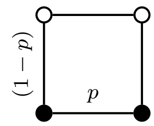
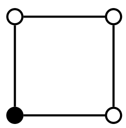
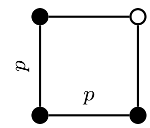
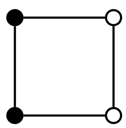
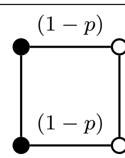
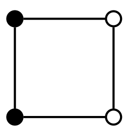
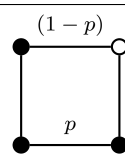
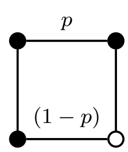
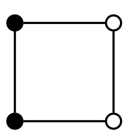
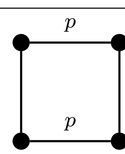
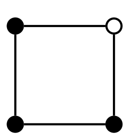
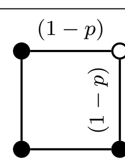
p_{ij}	Time$= t$	Time$= t+1$	$p_{ij} =$
p_{11}			$(1-p)^2$
p_{12}		 or 	$2p(1-p)$
p_{13}			p^2
p_{22}			$(1-p)^2$
p_{23}		 or 	$2p(1-p)$
p_{24}			p^2
p_{33}			$(1-p)^2$

Table 5.3.1. Elements of the transition matrix for the cycle C_4

Proof. Let T be the transition matrix for the cycle graph C_n. Consider p_{ij} where $i > j$. This is the probability that the number of infected vertices decreases after one time step. No vertex can become uninfected so this probability is 0.

Consider p_{ii} or when $i = j$. When $i = j = n$ this is the probability that if every vertex is infected at time t that every vertex will still be infected at time $t + 1$. Because no vertex can become uninfected and there are no more vertices for the virus to spread to so $p_{nn} = 1$. When $i = j \neq n$ this is the probability that the virus doesn't spread to any more vertices in the next time step and that there is at least one uninfected vertex left. On a cycle there are always two different ways the virus can spread when $i \neq n$ so $p_{ii} = (1 - p)^2$.

Consider p_{ij} where $j = i + 1$. This is the probability that the virus spreads to just one more vertex in a time step. This can happen in two different cases.

Case 1: Let $i = n - 1$ and $j = n$. This is the case where there is only one uninfected vertex left on the graph which means that it is adjacent to two infected vertices. This vertex can become infected in two ways. Either the virus spreads from one of the infected vertices and not the other or it spreads from both. This probability is $p_{ij} = 2p(1 - p) + p^2$.

Case 2: Let $i \leq n - 2$ and $j \leq n - 1$. This is the case where there are at least two uninfected vertices left on the graph. This means that the virus can spread to two different vertices in one time step. The probability that the virus only spreads to one of them is $p_{ij} = 2p(1 - p)$.

Consider p_{ij} where $j = i+2$. This is the probability that the virus spreads to two more vertices in one time step. When $j = i + 2$ there are always 2 uninfected vertices adjacent to infected vertices on the graph. Therefore the probability the virus spreads to both of them is $p_{ij} = p^2$.

Consider p_{ij} where $j \geq i + 3$. This is the probability that the virus spreads to 3 or more vertices in a time step. On a cycle the virus can spread to at most two more vertices in a time step. So the probability the virus spreads to three or more vertices in a time step is 0. □

5.3.3 The Matrix Q for the Cycle C_3

In this section we find the matrix Q for the cycle C_3. This example will help us formulate the matrix Q for any n-cycle.

Example 5.3.3. In this example we find the matrix Q for the cycle C_3. By Theorem 5.3.2 we know that the transition matrix for C_3 is

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 \\ 0 & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix Q is made up of the eigenvectors of T which by Theorem 2.2.7 are the elements on the diagonal. So the eigenvalues of T are $\lambda = 1$ and $\lambda = (1-p)^2$ where $(1-p)^2$ has multiplicity 2. By Theorem 4.1.2 we know that the eigenvector associated with $\lambda = 1$ is v_1 and the eigenvector associated with $\lambda = p_{11} = (1-p)^2$ is v_2 where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

When we solve for v_2 we get

$$(T-(1-p)^2I)v_2 = \begin{pmatrix} 0 & 2p(1-p) & p^2 \\ 0 & 0 & 2p(1-p)+p^2 \\ 0 & 0 & 1-(1-p)^2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

This implies

$$2p(1-p)x_2 + p^2x_3 = 0$$
$$[2p(1-p)+p^2]x_3 = 0$$

Therefore $x_3 = 0$, $x_2 = 0$ and x_1 can be anything. So the only eigenvector associated with $\lambda = (1-p)^2$ is v_2. Because $\lambda = (1-p)^2$ has multiplicity 2 we need to find a generalized eigenvector v_3 such that $(T-(1-p)^2I)v_3 = v_2$. Therefore

$$(T-(1-p)^2I)v_3 = \begin{pmatrix} 0 & 2p(1-p) & p^2 \\ 0 & 0 & 2p(1-p)+p^2 \\ 0 & 0 & 1-(1-p)^2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

which implies

$$2p(1-p)x_2 + p^2 x_3 = 1$$

$$[2p(1-p) + p^2]x_3 = 0$$

Therefore $x_1 = x_3 = 0$ and $x_2 = \frac{1}{2p(1-p)}$. Therefore

$$v_3 = \begin{pmatrix} 0 \\ \frac{1}{2p(1-p)} \\ 0 \end{pmatrix}.$$

When we arrange v_1, v_2, and v_3 into a matrix we get

$$Q = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & \frac{1}{2p(1-p)} \\ 1 & 0 & 0 \end{pmatrix}$$

5.3.4 The Matrix Q for an n-Cycle

Lemma 5.3.4. *Let T be the transition matrix for a cycle with n vertices C_n. Then T has two distinct eigenvalues $\lambda = 1$ and $\lambda = (1-p)^2$ where $\lambda = (1-p)^2$ has multiplicity $n-1$ and the only eigenvector associated with $\lambda = (1-p)^2$ is*

$$v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Proof. Let T be the transition matrix of a cycle with n vertices C_n. The matrix T is upper triangualr so by Theorem 2.2.7 the eigenvalues lie on the diagonal. Recall from Theorem 5.3.2 that the diagonal entries of T are 1 and $(1-p)^2$. Therefore the eigenvalues are $\lambda = 1$ and $\lambda = (1-p)^2$ where $(1-p)^2$ has multiplicity $n-1$. To find the eigenvectors associated with $\lambda = (1-p)^2$ we solve $(T - (1-p)^2 I)v = 0$. So

$$(T-(1-p)^2I)v = \begin{pmatrix} 0 & 2p(1-p) & p^2 & 0 & \dots & 0 \\ 0 & 0 & 2p(1-p) & p^2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 2p(1-p) & p^2 \\ \vdots & & & & 0 & 2p(1-p)+p^2 \\ 0 & \dots & \dots & \dots & 0 & 1-(1-p)^2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = 0$$

implies

$$\begin{aligned} 2p(1-p)x_2 + p^2x_3 =& 0 \\ 2p(1-p)x_3 + p^2x_4 =& 0 \\ 2p(1-p)x_4 + p^2x_5 =& 0 \\ \vdots& \\ 2p(1-p)x_{n-1} + p^2x_n =& 0 \\ [2p(1-p)+p^2]x_n =& 0 \\ [1-(1-p)^2]x_n =& 0 \end{aligned}$$

Working backwards through these equations we get $x_n = 0$ which implies $x_{n-1} = 0$ which implies $x_{n-2} = 0$...etc. This process continues through all the equations. Therefore we get $x_2 = x_3 = ... = x_n = 0$ and x_1 can be anything. So the only eigenvector associated with $\lambda = (1-p)^2$ is

$$v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

□

Theorem 5.3.5. *The matrix Q for the transition matrix T of a cycle with n vertices C_n is the $n \times n$ matrix*

$$Q = \left(\begin{array}{cc|cccc} 1 & 1 & 0 & & & \\ 1 & 0 & \frac{1}{2p(1-p)} & & & \\ \vdots & \vdots & 0 & & \cdots & \cdots \\ \vdots & \vdots & \vdots & & & \\ 1 & 0 & 0 & & & \end{array} \right)$$

where the $n-2$ columns to the right of the partition are the elements of a length $n-1$ chain of generalized eigenvectors based on the vector in the second column of Q. The $n-2$ generalized eigenvectors are defined as follows:

Let g_k be the kth generalized eigenvector in the chain such that

$$g_1 = \begin{pmatrix} 0 \\ \frac{1}{2p(1-p)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let x_i be the ith element of the vector g_k and let y_i be the ith element of the previous vector g_{k-1}. Then

$$x_1 = 0, x_n = \frac{y_n}{[1-(1-p)^2]} = \frac{y_{n-1}}{2p(1-p)+p^2}, \text{ and } x_i = \frac{y_{i-1} - p^2 x_{i+1}}{2p(1-p)} \text{ for all } x_i \text{ where } 2 \le i \le n-1.$$

Proof. Let T be the transition matrix of the cycle C_n. The matrix Q contains the eigenvectors of T. We know by Lemma 5.3.4 that T has two distinct eigenvalues $\lambda = 1$ and $\lambda = (1-p)^2$ where $(1-p)^2$ has multiplicity $n-1$. Lemma 5.3.4 also tells us that the only eigenvectors of T are

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where v_1 is associated with $\lambda = 1$ and v_2 is associated with $\lambda = (1-p)^2$. Because $\lambda = (1-p)^2$ has multiplicity $n-1$ and it only corresponds to one eigenvector, there is a length $n-1$ chain of generalized eigenvectors based on v_2 such that

$$\begin{aligned}(T-(1-p)^2I)g_1 =& v_2\\ (T-(1-p)^2I)g_2 =& g_1\\ &\vdots\\ (T-(1-p)^2I)g_{n-2} =& g_{n-3}\end{aligned}$$

Where g_k is the kth generalized eigenvector in the chain. To find g_k we solve $(T-(1-p)^2I)g_k = g_{k-1}$. Let x_i be the ith element in the vector g_k and let y_i be the ith element in the vector g_{k-1}. Then

$$(T-(1-p)^2I)g_k = \begin{pmatrix} 0 & 2p(1-p) & p^2 & 0 & \dots & 0 \\ 0 & 0 & 2p(1-p) & p^2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 2p(1-p) & p^2 \\ \vdots & & & & 0 & 2p(1-p)+p^2 \\ 0 & \dots & \dots & \dots & 0 & 1-(1-p)^2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ \vdots \\ y_n \end{pmatrix}$$

which implies

$$\begin{aligned}2p(1-p)x_2 + p^2x_3 =& y_1\\ 2p(1-p)x_3 + p^2x_4 =& y_2\\ 2p(1-p)x_4 + p^2x_5 =& y_3\\ &\vdots\\ 2p(1-p)x_{n-1} + p^2x_n =& y_{n-2}\\ [2p(1-p) + p^2]x_n =& y_{n-1}\\ [1-(1-p)^2]x_n =& y_n\end{aligned}$$

Therefore x_1 can be anything so let $x_1 = 0$,

$$x_n = \frac{y_n}{[1-(1-p)^2]} = \frac{y_{n-1}}{2p(1-p)+p^2}, \text{ and}$$

$$x_i = \frac{y_{i-1} - p^2 x_{i+1}}{2p(1-p)} \text{ for all } x_i \text{ such that } 2 \le i \le n-1.$$

□

5.3.5 Jordan Form Matrix J for the Cycle C_4

In this section we find the matrix J for the cycle C_4. This example will help us formulate the matrix J for any n-cycle.

Example 5.3.6. In this example we find J for the cycle C_4. In Example 5.3.1 we found that the transition matrix for C_4 is

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 & 0 \\ 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

By Theorem 2.2.7 we know that the eigenvalues of T lie on the diagonal. So, T has two distinct eigenvalues $\lambda = 1$ and $\lambda = (1-p)^2$ where $\lambda = (1-p)^2$ has multiplicity 3. By Lemma 5.3.4 we know that there is only one eigenvector associated with $\lambda = (1-p)^2$. Therefore T has two linearly independent eigenvectors, so J is made up of two Jordan blocks by Theorem 2.2.15

$$J_1 = (1) \text{ and } J_2 = \begin{pmatrix} (1-p)^2 & 1 & 0 \\ 0 & (1-p)^2 & 1 \\ 0 & 0 & (1-p)^2 \end{pmatrix}.$$

Therefore

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-p)^2 & 1 & 0 \\ 0 & 0 & (1-p)^2 & 1 \\ 0 & 0 & 0 & (1-p)^2 \end{pmatrix}$$

5.3.6 Jordan Form Matrix J for an n-Cycle

Theorem 5.3.7. *The Jordan Form J for the transition matrix T of a cycle with n vertices C_n is the $n \times n$ matrix*

$$J = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (1-p)^2 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & (1-p)^2 \end{pmatrix}.$$

Proof. Let T be the transition matrix for the cycle C_n. Let J be the Jordan Form matrix of T. We know T is upper triangular so by Theorem 2.2.7 the eigenvalues lie on the diagonal. Therefore by Theorem 5.3.2 the eigenvalues are $\lambda = 1$ and $\lambda = (1-p)^2$ where $\lambda = (1-p)^2$ has multiplicity $n-1$. By Lemma 5.3.4 we know that there is only one eigenvector associated with $\lambda = (1-p)^2$. Therefore T has two linearly independent eigenvectors so J is made up of two Jordan blocks.

$$J_1 = (1) \text{ and } J_2 = \begin{pmatrix} (1-p)^2 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & (1-p)^2 & 1 \\ 0 & \dots & \dots & 0 & (1-p)^2 \end{pmatrix}$$

where J_2 is an $n-1 \times n-1$ matrix. Therefore

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (1-p)^2 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & (1-p)^2 \end{pmatrix}.$$

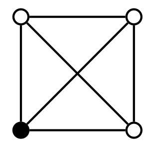
□

5.4 Virus Spread on Complete Graphs

In this section we analyze the spread of a virus on complete graphs.

5.4.1 Virus Spread on the complete graph K_4

In this section we analyze the spread of a virus on the complete graph K_4, as shown below.



Consider the probability p_{34}. This is the probability that the graph has 3 vertices infected at time t and spreads to the last uninfected vertex in the next time step. The last vertex is adjacent to 3 infected vertices so this can happen three ways:

Case 1: The virus spreads from one of the three vertices.

Case 2: The virus spreads from two of the three vertices.

Case 3: The virus spreads from all three vertices.

Table 5.4.1 illustrates these cases. Note that the dashed line indicates the edge that the virus spreads over. Adding these cases together we get $p_{34} = 3p(1-p)^2 + 3p^2(1-p) + p^3$.

Computing the rest of the transition probabilities the same way we get that the transition matrix for the graph K_4 is

$$T = \begin{pmatrix} (1-p)^3 & 3p(1-p)^2 & 3p^2(1-p) & p^3 \\ 0 & (1-p)^4 & 4p(1-p)^3 + 2p^2(1-p)^2 & 4p^2(1-p) + p^4 \\ 0 & 0 & (1-p)^3 & 3p(1-p)^2 + 3p^2(1-p) + p^3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

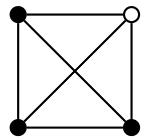
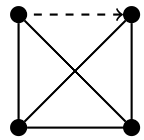
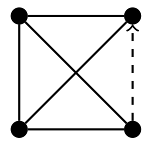
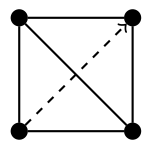
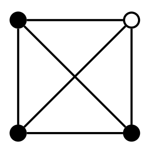
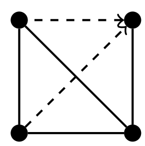
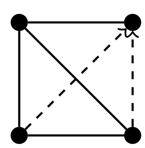
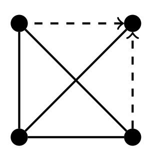
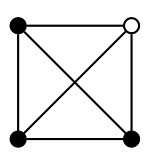
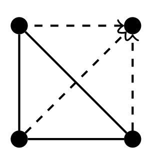
Case	Time= t	Time= $t+1$	Probability
Case 1		 or  or 	$3p(1-p)^2$
Case 2		 or  or 	$3p^2(1-p)$
Case 3			p^3

Table 5.4.1. p_{34} for the complete graph K_4

5.4.2 Transition Matrix for K_n

In Yushan Jiang's senior project *The Analysis of Probabilistic Spread on Complete Graphs* [4] she came up with a theorem that describes the transition matrix for a complete graph with n vertices K_n. The following theorem is taken directly from her project, Jiang [4, p. 33]

Theorem 5.4.1. *Every p_{ij} in the transition matrix P of the Markov chain for K_n follows*

$$p_{ij} = \binom{n-i}{j-i} \cdot [1-(1-p)^i]^{j-i} \cdot [(1-p)^i]^{n-j} 0 \leq i \leq j \leq n.$$

Note that $\binom{n-i}{j-i} = 0$ if $j < i$.

6
Virus Spread on Broom Graphs

In this chapter we introduce a new type of graph called **Broom graphs** and analyze their transition matrices.

6.1 Broom Graphs

Definition 6.1.1. A **broom graph** is the union of a path and another type of simple graph where both graphs are connected by an edge.

We specifically look at **star-brooms** denoted by P_mS_n, **cycle-brooms** denoted by P_mC_n, and **complete-brooms** denoted by P_mK_n where m is the number of vertices in the path and n is the number of vertices in the attached graph.

Figures 6.1.1, 6.1.2, and 6.1.3 illustrate the star-broom graph P_4S_4, the cycle-broom graph P_3C_4, and the complete-broom graph P_3K_4, respectively.

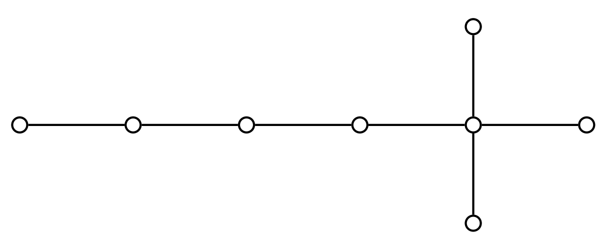


Figure 6.1.1. The star-broom graph P_4S_4

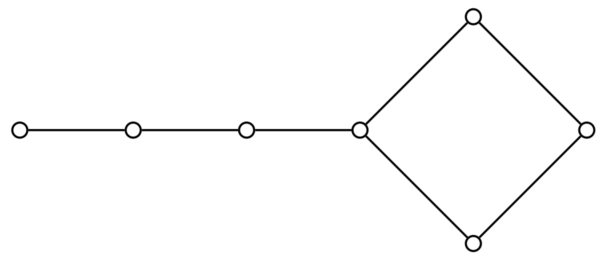


Figure 6.1.2. The cycle-broom graph P_3C_4

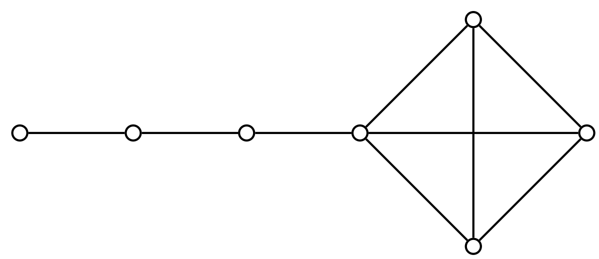


Figure 6.1.3. The complete-broom graph P_3K_4

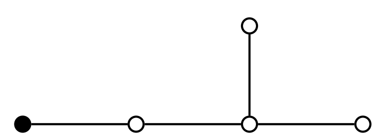
6.2 Virus Spread on Star-Brooms

In this section we analyze the spread of a virus on star-broom graphs where the root node is the starting vertex of the path.

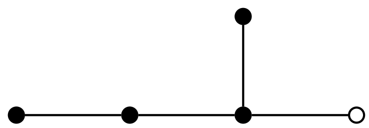
6.2.1 Transition Matrix for the Star-Broom P_2S_3

In this section we look at the transition matrix for the star-broom P_2S_3. This will help us formulate the transition matrix for a star-broom graph with $m + n$ vertices.

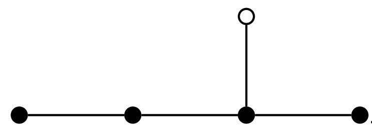
Example 6.2.1. In this example we find the transition matrix for the star-broom graph P_2S_3 as shown below.



Consider the probability p_{44}. This is the probability that the graph starts in state 4 at time t and in the next time step is still in state 4. On the graph P_2S_3 there are two different ways of representing state 4:



and



In both representations there is only one vertex that could be infected in the next time step. Therefore the probability that the graph stays in state 4 is $p_{44} = (1-p)$.

The probability p_{45} is similar except in this case the virus is spreading to the last vertex. Therefore $p_{45} = p$.

The probabilities p_{13}, p_{14}, p_{15}, p_{24}, and p_{25} are all the probabilites where the virus is spreading by more than one vertex across a path. As we saw in Section 5.1, when a virus spreads across a path it cannot spread to more than one vertex in a time step. Therefore these probabilities are 0.

Note that the elements below the diagonal are all the elements where the number of infected vertices decreases. Since a vertex cannot become uninfected these probabilities are also 0.

Table 6.2.1 illustrates the rest of the elements of the transition matrix for P_2S_3

When we arrange these probabilities in a matrix we get

$$T = \begin{pmatrix} (1-p) & p & 0 & 0 & 0 \\ 0 & (1-p) & p & 0 & 0 \\ 0 & 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

6.2.2 Transition Matrix for the Star-Broom P_mS_n

Theorem 6.2.2. *Let T be the transition matrix for a star-broom graph P_mS_n with $m+n$ vertices. Then T is an $m+n \times m+n$ matrix*

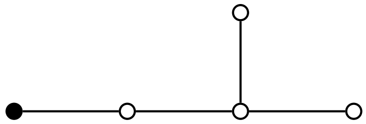
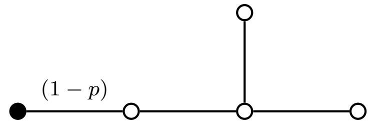
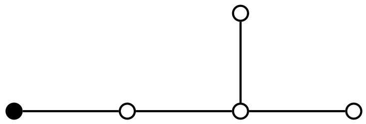
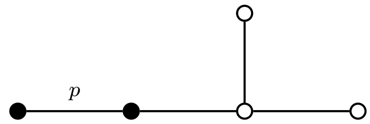
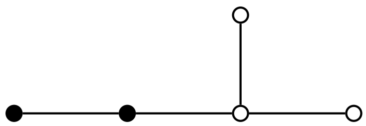
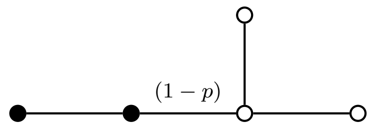
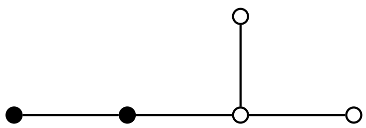
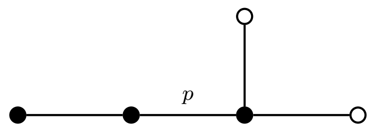
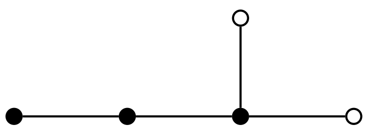
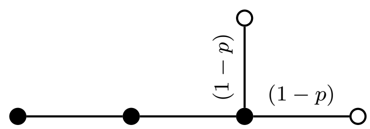
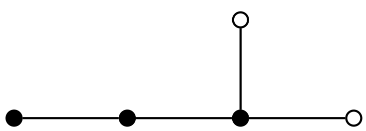
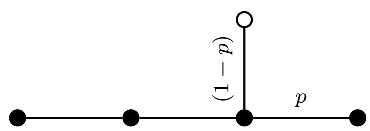
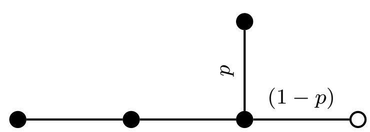
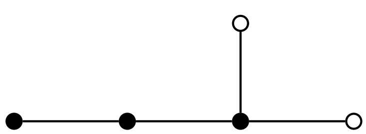
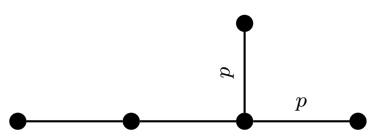
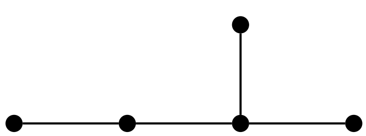
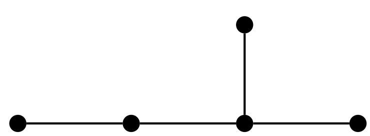
p_{ij}	Time $= t$	Time $= t+1$	$p_{ij} =$
p_{11}			$(1-p)$
p_{12}			p
p_{22}			$(1-p)$
p_{23}			p
p_{33}			$(1-p)^2$
p_{34}		 or 	$2p(1-p)$
p_{35}			p^2
p_{55}			1

Table 6.2.1. Elements of the transition matrix of the graph P_2S_3

$$T = \left(\begin{array}{ccccc|c} (1-p) & p & 0 & \dots & 0 & \\ 0 & \ddots & \ddots & \ddots & \vdots & 0's \\ \vdots & \ddots & \ddots & \ddots & 0 & \\ 0 & \dots & 0 & (1-p) & p & \\ \hline & & & & \multicolumn{2}{|c}{} \\ & 0's & & & \multicolumn{2}{|c}{T \text{ for } S_n} \end{array}\right)$$

where the top left box is an $m \times m+1$ block and the bottom right box is an $n \times n$ block. Note the bottom right $n \times n$ block contains the same elements as the transition matrix for the star S_n. The matrix T has elements such that

$$p_{ij} = \begin{cases} (1-p), & \text{if } i = j \leq m \\ p, & \text{if } j = i+1 \leq m+1 \\ \binom{n-i}{j-i} p^{j-i}(1-p)^{n-j}, & \text{if } i > m \\ 0, & \text{otherwise} \end{cases}$$

Proof. Let T be the transition matrix for the star-broom $P_m S_n$ with the root node at the endpoint of the path. The first m rows are the probabilities of the virus spreading from state i where $i \leq m$. Since there are m vertices on the path portion of the graph these probabilities will be the same as a virus spreading on a path with $m+n$ vertices. So

$$p_{ij} = \begin{cases} (1-p), & \text{if } i = j \leq m \\ p, & \text{if } j = i+1 \leq m+1 \\ 0, & \text{otherwise} \end{cases}$$

Once the virus spreads to $m+1$ vertices the virus has reached the center of the star graph. Every vertex on the path portion has been infected so the virus can only spread to the rest of the star portion of the graph. Therefore the transition probabilities are exactly what they would be for a star with n vertices. Hence

$$p_{ij} = \binom{n-i}{j-i} p^{j-i}(1-p)^{n-j}, \text{ if } i > m.$$

□

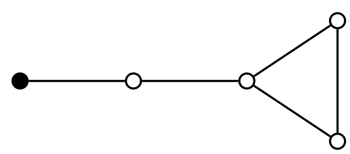
6.3 Virus Spread on Cycle-Brooms

In this section we analyze the spread of a virus on cycle-broom graphs where the root node is the starting vertex of the path.

6.3.1 Transition Matrix for the Cycle-Broom P_2C_3

In this section we look at the transition matrix for the cycle-broom P_2C_3. This will help us formulate the transition matrix for a cycle-broom graph with $m+n$ vertices.

Example 6.3.1. In this example we will find the transition matrix for the cycle-broom graph P_2C_3 as shown below.



Consider the probability p_{45}. This is the probability that the graph starts with 4 infected vertices at time t and spreads to the last vertex in the next time step. The last vertex is adjacent to two infected vertices so there are two different ways this can happen; The virus spreads from one vertex and not the other or it spreads from both. This probability is $p_{45} = 2p(1-p) + p^2$.

Table 6.3.1 illustrates some of the elements of the transition matrix of P_2C_3.

The transition matrix for the graph P_2C_3 is

$$T = \begin{pmatrix} (1-p) & p & 0 & 0 & 0 \\ 0 & (1-p) & p & 0 & 0 \\ 0 & 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & 0 & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

6.3.2 Transition Matrix for the Cycle-Broom P_mC_n

Theorem 6.3.2. *Let T be the transition matrix for a cycle-broom graph P_mC_n with $m+n$ vertices. Then T is an $m+n \times m+n$ matrix*

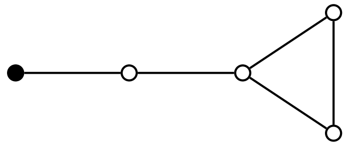
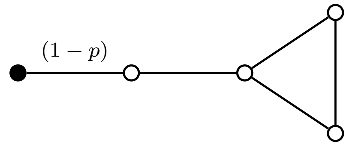
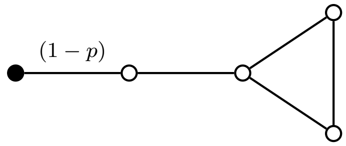
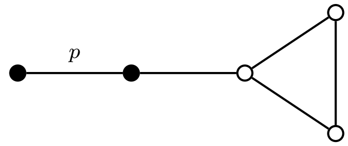
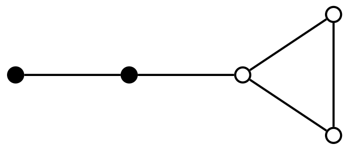
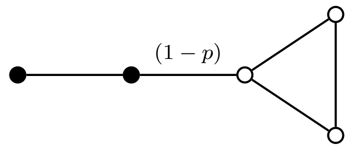
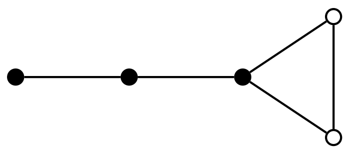
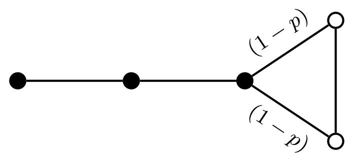
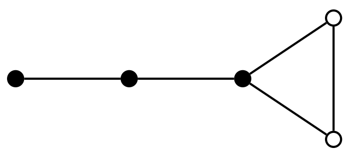
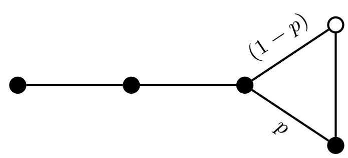
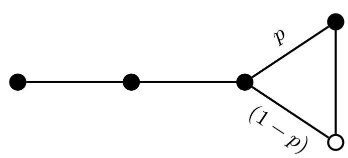
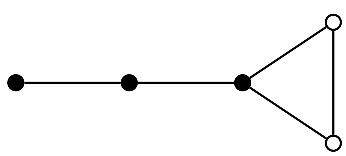
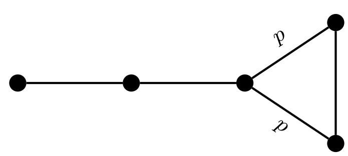
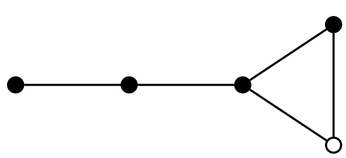
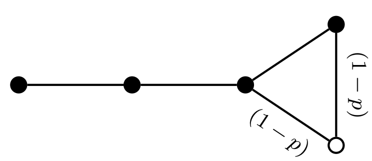
p_{ij}	Time $= t$	Time $= t+1$	$p_{ij} =$
p_{11}			$(1-p)$
p_{12}			p
p_{22}			$(1-p)$
p_{33}			$(1-p)^2$
p_{34}		 or 	$2p(1-p)$
p_{35}			p^2
p_{44}			$(1-p)^2$

Table 6.3.1. Elements of the transition matrix of the graph P_2C_3

$$T = \left(\begin{array}{ccccc|c} (1-p) & p & 0 & \dots & 0 & \\ 0 & \ddots & \ddots & \ddots & \vdots & 0's \\ \vdots & \ddots & \ddots & \ddots & 0 & \\ 0 & \dots & 0 & (1-p) & p & \\ \hline & & 0's & & & \boxed{T \text{ for } C_n} \end{array}\right)$$

where the top left box is an $m \times m+1$ block and the bottom right box is an $n \times n$ block. Note the bottom right $n \times n$ block contains the same elements as the transition matrix for the cycle C_n.

The proof is similar to the proof of Theorem 6.2.2.

6.4 Virus Spread on Complete-Brooms

In this section we analyze the spread of a virus on complete-broom graphs where the root node is the starting vertex of the path.

6.4.1 Transition Matrix for the Complete-Broom P_2K_4

In this section we find the transition matrix for the complete-broom graph P_2K_4 as shown below.

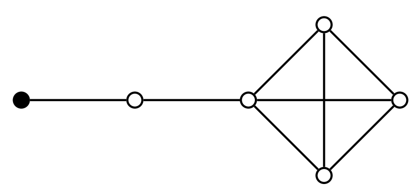


Table 6.4.1 illustrates some of the elements of the transition matrix for the graph P_2K_4. Using the same method we can find the rest of the transition probabilities for P_2K_4. Therefore

$$T = \begin{pmatrix} (1-p) & p & 0 & 0 & 0 & 0 \\ 0 & (1-p) & p & 0 & 0 & 0 \\ 0 & 0 & (1-p)^3 & 3p(1-p)^2 & 3p^2(1-p) & p^3 \\ 0 & 0 & 0 & (1-p)^4 & 4p(1-p)^3 + 2p^2(1-p)^2 & 4p^2(1-p) + p^4 \\ 0 & 0 & 0 & 0 & (1-p)^3 & 1-(1-p)^3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

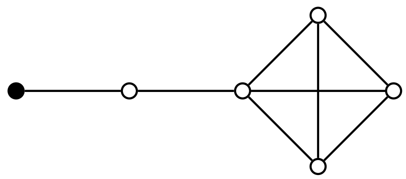
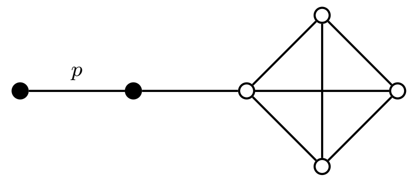
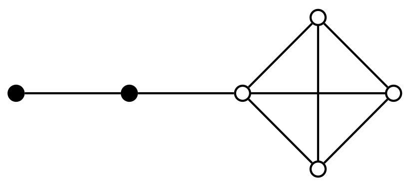
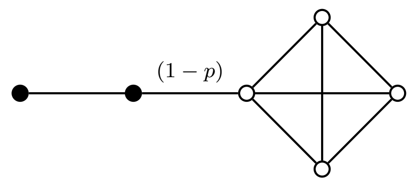
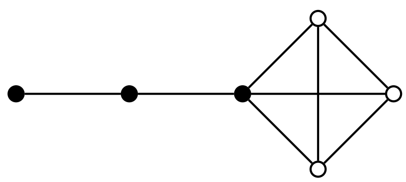
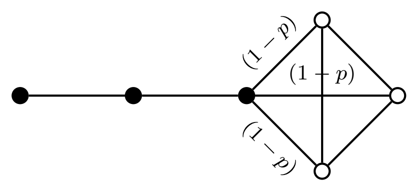
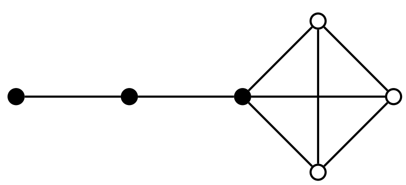
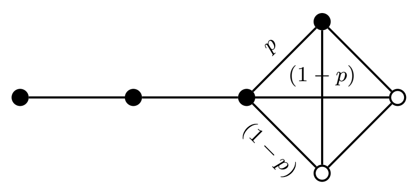
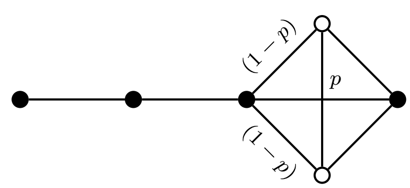
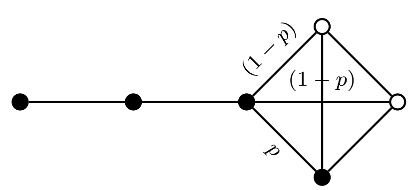
p_{ij}	Time$= t$	Time$= t+1$	Probability
p_{12}			p
p_{22}			$(1-p)$
p_{33}			$(1-p)^3$
p_{34}		 or  or 	$3p(1-p)^2$

Table 6.4.1. Elements of the transition matrix of the graph P_2K_4

6.4.2 Transition Matrix for the Complete-Broom $P_m K_n$

Theorem 6.4.1. *Let T be the transition matrix for a complete-broom graph $P_m K_n$ with $m+n$ vertices. Then T is an $m+n \times m+n$ matrix*

$$T = \left(\begin{array}{ccccc|c} (1-p) & p & 0 & \dots & 0 & \\ 0 & \ddots & \ddots & \ddots & \vdots & 0's \\ \vdots & \ddots & \ddots & \ddots & 0 & \\ 0 & \dots & 0 & (1-p) & p & \\ \hline & & 0's & & \multicolumn{2}{|c}{\boxed{T \text{ for } K_n}} \end{array}\right)$$

where the top left box is an $m \times m+1$ block and the bottom right box is an $n \times n$ block. Note the bottom right $n \times n$ block contains the same elements as the transition matrix for the complete graph K_n.

The proof is similar to the proof of Theorem 6.2.2.

7

Modifying the Markov Chain Model

The Markov chain method that we have been using defines the states as being the number of infected vertices on a graph. For example, state 1 = 1 vertex infected, state 2 = 2 vertices infected, etc. In this chapter we see that this definition of states doesn't work with all graphs. We see an example of a graph that doesn't work and see how to modify our model so that it does.

7.1 Graphs that don't Work with the Markov Chain Method

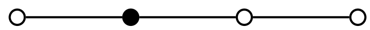
In this section we look at an example of a graph that doesn't work with our current Markov chain model.

Example 7.1.1. In this example we look at the path P_4 where the root node is at the second vertex from the left, as shown below.

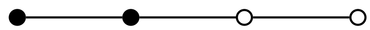
We define the states of this graph the same way we have in previous chapters, state 1 = 1 vertex infected, state 2 = 2 vertices infected, etc.

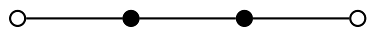
Let's look at how these states are represented on the graph P_4.

State 1:

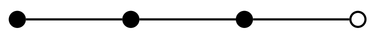


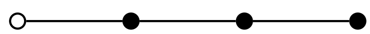
State 2:

(a) 

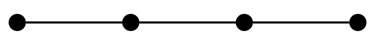
(b) 

State 3:

(a) 

(b) 

State 4:



There are two different ways to represent state 2 and state 3. The graphs that represent state 3 are essentially the same since the probabilities $p_{33} = (1-p)$ and $p_{34} = p$ are the same for both representations.

Now, let's look at the the transition probabilities involving state 2, the second row of the transition matrix $(p_{21}, p_{22}, p_{23}, p_{24})$.

$\underline{p_{21}}$: This is the probability that the graph has 2 infected vertices and then decreases to 1 infected vertex in the next time step. No vertex can become uninfected so this probability is 0.

$\underline{p_{22}}$: This is the probability that the graph has two vertices infected at time t and in the next time step still has 2 vertices infected. Lets consider this probability for both graphs (a) and (b) that represent state 2.

Graph (a): There is only one vertex that the virus could spread to so this probability is $(1-p)$.

Graph (b): There are two vertices that the virus could spread to so this probability is $(1-p)^2$.

Therefore $p_{22} = (1-p) + (1-p)^2$.

$\underline{p_{23}}$: This is the probability that the graph has two vertices infected at time t and in the next time step has 3 vertices infected. Lets consider this probability for both graphs (a) and (b) that represent state 2.

Graph (a): There is only one vertex that the virus could spread to so this probability is p.

Graph (b): There are two vertices that the virus could spread to so this probability is $2p(1-p)$.

$\underline{p_{24}}$: This is the probability that the graph has 2 vertices infected at time t and in the next time step has 4 infected vertices. This probability is only possibel in graph (b) since there are 2 uninfected verties adjacent to an infected vertex. Therefore this probability is p^2.

Note that when we look at the transition matrix for this graph that the elements of row 2 add up to 2.

$$T = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ 0 & (1-p)+(1-p)^2 & p+2p(1-p) & p^2 \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix}$$

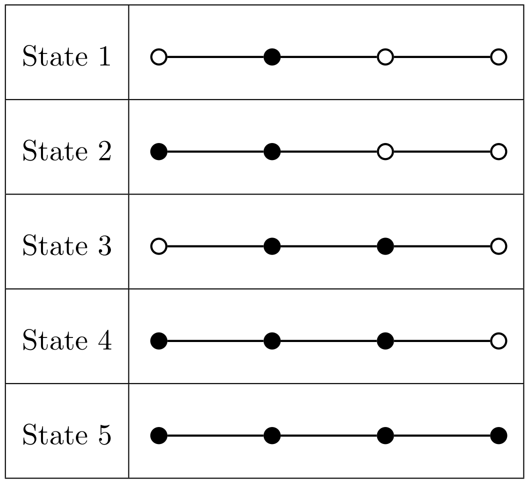
$$(1-p)+(1-p)^2+p+2p(1-p)+p^2 = 2$$

In a Markov chain the elements of each row need to add up to 1 because each row covers all the possibilities. Therefore this graph doesn't work with our current Markov chain model of defining the states as the number of vertices infected.

7.2 Modifying the Markov Chain Method

In this section we see how to modify our current Markov chain method to work with the path P_4 where the root node is at the second vertex.

Instead of defining the states as the number of infected vertiecs we define them as follows:



This graph now has 5 states so it will have a 5×5 transition matrix. Let's look at the transition probabilities for this graph.

Consider the probability p_{23}. This is the probability that the graph is in state 2 at time t and then moves to state 3 in the next time step. In state 2 an endpoint of the graph is infected and in state 3 only the center two vertices are infected. Since the endpoint can't become uninfected this probability is 0. Similarly, $p_{32} = 0$.

Table 7.2.1 illustrates the rest of the transition probabilities for this graph. When we arrange these probabilities into a matrix we get

$$T = \begin{pmatrix} (1-p)^2 & p(1-p) & p(1-p) & p^2 & 0 \\ 0 & (1-p) & 0 & p & 0 \\ 0 & 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Defining the states of a graph this way opens up an area of study for future work. All the graphs analyzed in this project were chosen because they work with our previous Markov chain model of defining the states as the number of infected vertices. By modifying our method we are now able to study the spread of a virus on different types of graphs.

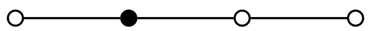
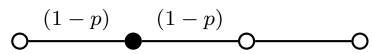
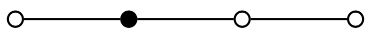
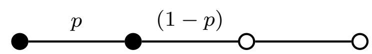
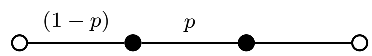
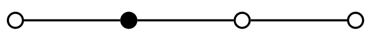
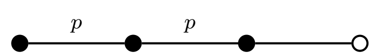
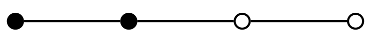
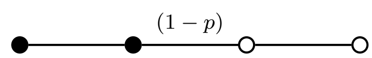
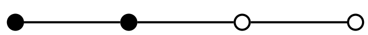
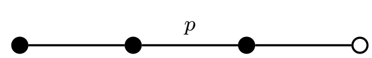
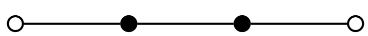
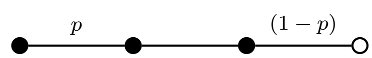
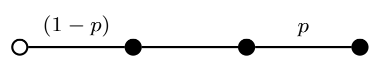
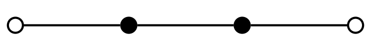
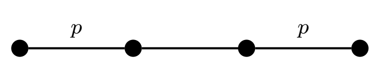
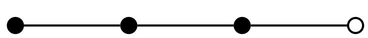
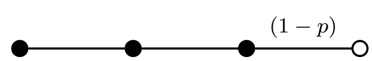
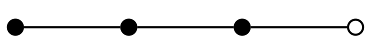
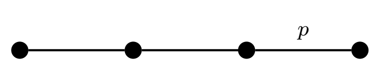
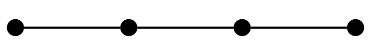
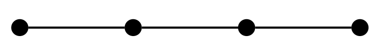
p_{ij}	Time$=t$	Time$=t+1$	Probability
p_{11}			$(1-p)^2$
p_{12}			$p(1-p)$
p_{13}			$p(1-p)$
p_{14}			p^2
p_{22}			$(1-p)$
p_{24}			p
p_{34}		 or 	$2p(1-p)$
p_{35}			p^2
p_{44}			$(1-p)$
p_{45}			p
p_{55}			1

Table 7.2.1. Elements of the transition matrix for the graph P_4 where the root node is at the second vertex

Appendix A
Transition Matricies

A.1 Transition Matrix of an n-Path

$$T = \begin{pmatrix} (1-p) & p & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & (1-p) & p \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

A.2 Transition Matrix of an n-Star

$$p_{ij} = \begin{cases} \binom{n-i}{j-i} p^{j-i}(1-p)^{n-j}, & \text{if } i \leq j \\ 0, & \text{if } i > j \end{cases}$$

A.3 Transition Matrix of an n-Cycle

$$T = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & (1-p)^2 & 2p(1-p) & p^2 \\ \vdots & & & & (1-p)^2 & 2p(1-p)+p^2 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Appendix B

The Eigenvector Matrix Q

B.1 The Matrix Q for n-Paths

$$Q = \begin{pmatrix} 1 & 1 & 0 & \dots & \dots & 0 \\ 1 & 0 & \frac{1}{p} & \ddots & & \vdots \\ 1 & 0 & 0 & \frac{1}{p^2} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \frac{1}{p^{n-2}} \\ 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

B.2 The Matrix Q fot n-Stars

$$Q = \begin{pmatrix} \binom{n-1}{0} & \binom{n-1}{1} & \dots & \dots & \binom{n-1}{n-1} \\ \binom{n-2}{0} & \binom{n-2}{1} & \dots & \binom{n-2}{n-2} & 0 \\ \vdots & & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & & \vdots \\ \binom{n-n}{0} & 0 & \dots & \dots & 0 \end{pmatrix}$$

B.3 The Matrix Q for n-Cycles

$$Q = \left(\begin{array}{cc|ccc} 1 & 1 & 0 & & \\ 1 & 0 & \frac{1}{2p(1-p)} & & \\ \vdots & \vdots & 0 & \dots & \dots \\ \vdots & \vdots & \vdots & & \\ 1 & 0 & 0 & & \end{array} \right)$$

Appendix C
The Jordan Form Matrix J

C.1 The Matrix J for n-Paths

$$J = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (1-p) & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & (1-p) \end{pmatrix}.$$

C.2 The Matrix J fot n-Stars

$$J = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & (1-p) & \ddots & & \vdots \\ \vdots & \ddots & (1-p)^2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & (1-p)^{n-1} \end{pmatrix}.$$

C.3 The Matrix J for n-Cycles

$$J = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (1-p)^2 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & (1-p)^2 \end{pmatrix}.$$

Bibliography

[1] C. Henry Edwards and David E. Penney, *Differential Equations and Linear Algebra*, 3rd ed., Pearson, Upper Saddle River, NJ, 2010.

[2] Charles M. Grinstead and J. Laurie Snell, *Introduction to Probability*, American Mathematical Society, Providence, RI USA, 2003.

[3] Frederick S. Hillier and Gerald J. Leiberman, *Introduction to Operations Research*, 7th ed., McGraw-Hill, New York, NY, 2001.

[4] Yushan Jiang, *The Analysis of Probabilistic Spread on Complete Graphs*, A Senior Project from Bard College, 2014.

[5] Anam Nasim, *Spread of a Virus on Networks: A Probabilistic Approach*, A Senior Project from Bard College, 2013.

[6] Géza Schay, *Introduction to Probability with Statistical Applications*, Birkhäuser Publishing, Boston, 2007.

[7] Robin J. Wilson, *Introduction to Graph Theory*, 5th ed., Prentice hall, Harlow, 2010.