DELFT UNIVERSITY OF TECHNOLOGY

FACULTY OF ELECTRICAL ENGINEERING, MATHEMATICS AND COMPUTER SCIENCE

Malthusian behaviour of the Leslie model for population growth

by

JILL DUFFY

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended publicly on October 12, 2023.

Student number:	5154987	
Thesis committee:	Dr. H. Yoldaş,	TU Delft, supervisor
	Dr. C. Kraaikamp,	TU Delft

TU Delft Delft University of Technology

Abstract

The Leslie model considers age-specific birth and survival rates to describe how a population size and age distribution changes over time. This thesis investigates the long-term dynamics of the Leslie model for population growth, utilizing mathematical theorems such as Perron-Frobenius, Doeblin's theorem, and Branching processes.

The Perron-Frobenius theorem guarantees the existence of a dominant eigenvalue. This dominant eigenvalue and its corresponding eigenvector represent the long-term behaviour of a population; The dominant eigenvalue indicates the long-term population growth, and the corresponding eigenvector indicates the long-term age distribution. Furthermore, the Perron-Frobenius theorem implies that a population asymptotically reaches a stable age distribution that is independent of its initial age structure. Once this stable age distribution is reached, the population continues to grow exponentially, exhibiting Malthusian behaviour. Doeblin's theorem, although not directly applicable to the complete Leslie model, provides valuable insights into the long-term behaviour of Markov chains. As Doeblin's theorem can not be applied to the complete Leslie model, the reproduction process of the Leslie model is formulated as a Branching process. Introducing the Leslie model as a Branching process allows for the consideration of demographic stochasticity. Simulations reveal that for larger populations, the Branching process closely mirrors the Leslie model, while disparities become more pronounced in smaller populations. These results illustrate the impact of probabilistic factors in population dynamics, as well as the strength of the Leslie model for larger populations.

Contents

Chapter 1

Introduction

Population projections are hypothetical scenarios that aim to show developments of the population size and structure. These population projections are based on a set of assumptions for future levels of fertility and mortality. The projections provide information about how the population size and structure would change if the assumptions made regarding fertility and mortality remained true. As a result, population projections help the public, statisticians and policymakers to understand population dynamics [23].

Thomas Malthus was an English economist and is well known for his influential essay "*An Essay on the Principle of Population*," first published in 1798. Malthus's work laid the foundation for what is now known as Malthusian theory or Malthusianism. In this essay, Malthus argued that populations tend to grow at an exponential rate, in the absence of significant checks on its growth, such as food scarcity, disease, or other limitations on resources [20].

This research is focused on the Leslie model for population growth, a commonly used model in population demography. Similar to the Malthusian theory, the Leslie model assumes that a population can grow in an unlimited environment.

The Leslie model is a mathematical framework used to study population dynamics. The Leslie model represents how a population's size and age distribution change over time. By quantifying age-specific reproductive and survival rates, researchers gain insights into the ways different age groups contribute to population growth or decline. Thereby, the Leslie model is a valuable tool for ecologists and demographers to analyze and project population changes, providing insights into how factors such as birth rates, mortality, and age structure influence overall population trends.

The goal of this thesis is to investigate the long-term behaviour within the Leslie model for population growth, utilizing various mathematical theorems. Various mathematical tools allow for complementary insights that can increase our understanding. Chapter 2 provides an in-depth exploration of the Leslie model, offering insights into its historical development, key assumptions, and mathematical foundations. Then, in Chapter 3, the Perron-Frobenius theorem is introduced, a mathematical tool to analyze the Leslie model's long-term population growth dynamics. Chapter 4, introduces Doeblin's theorem and discusses its application to the Leslie model, shedding light on the model's behaviour as a Markov chain. Chapter 5 formulates part of the Leslie model as a branching process. This approach considers demographic stochasticity, and evaluates how it compares to the Leslie model. Finally, Chapter 6 provides a summary of the findings and a discussion of the limitations within this research.

Chapter 2

Leslie model for population growth

2.1 Patrick Holt Leslie

Patrick Holt Leslie was born in 1900 in Scotland. Leslie studied at Oxford University and obtained a bachelor's degree in physiology, in 1921. Due to health problems, he could not continue his studies in physiology and instead turned to statistics. In 1935, Leslie joined the Bureau of Animal Population, a new research centre set up by Charles Elton. The purpose of this research centre was to study the fluctuations of animal populations through field studies and laboratory experiments. Elton suggested to Leslie, that it would be valuable if the mortality and fertility schedules of an organism could be combined into a single expression. Accordingly, Leslie studied and experimented with rodents from Canada, to analyse the patterns of their population growth. Leslie used methods developed by Alfred Lotka, who researched human demography.

In 1945 Leslie published his most famous article in Biometrika, a journal, founded by Galton, Pearson and Weldon in 1901. The article was entitled *On the use of matrices in certain population mathematics*. The article is focused on the growth of the female population of different rats of different age categories. Though Leslie has brought an enormous contribution to modern environmental science, it took 25 years for ecologists to adopt Leslie's matrix population models [6]. After computer science improved, software was available to execute matrix calculations necessary to apply the Leslie model [1].

2.2 Applications of the Leslie model

The Leslie model is frequently used in population dynamics. Population dynamics is the branch of ecology that deals with the variation in time and space of population size and density for one or more species [2]. The Leslie model considers age-specific birth and survival rates to describe how a population size and age distribution changes over time. The Leslie model provides a method for understanding and projecting the patterns of a population, through the analysis of age-specific vital rates and application of matrix algebra. The Leslie model has applications in various fields, including wildlife management, conservation biology and human demographics.

There is a crucial distinction between forecasting and projecting. Whereas a forecast predicts what will happen, a projection describes what would happen, given certain hypotheses. Thus, the Leslie model is used to project a population if the present conditions were to be maintained. Indeed, one of the most powerful ways to examine present conditions is to examine their consequences if they were they to remain as they are [5]. The Leslie model is a method to project the present state of a population into the future.

The model is applied to ecology, to examine present conditions model changes in a population of organisms over a period of time. Since the growth and dynamics of a population directly affect the state of the economy, politics and the environment, population projections are important, as they assist government and researchers to make decisions about the future.

As the Leslie model can demonstrate the consequences of present demographic trends, it allows researchers to examine how plans to manage populations affect the vitality of a population [6]. Thus population projections may indicate that existing trends and policies are likely to lead to undesirable outcomes. For this reason there is often interest in examining which demographic features of a population influence the realized rate of increase or how certain changes in demographic parameters would ultimately influence population growth [24].

2.3 Derivation of the Leslie model

The Leslie model is a discrete, age-structured model of population growth, in which a population is divided into discrete age classes of equal size. The structured model describes how individuals move in time among the defined classes. The model thereby describes the dynamics of the population class distribution and, as a result, the dynamics of the population as a whole.

Given the age distribution of a population on a certain date, the Leslie model provides a method to determine the age distribution of the survivors and descendants of the original population at subsequent time intervals [18, 19].

2.3.1 Assumptions

The Leslie model for population growth is based on various key assumptions.

- The Leslie model requires age specific population numbers, fertility rates and mortality rates. In order to simplify the problem, the model assumes that **the age-specific birth and survival rates remain constant over a period of time** [18, 19]. Though in reality birth and survival rates are more likely to change through time or change depending on the population density.

- Since the Leslie model is a discrete model, it discards all information on the ages of individuals within age classes. Thus, it is assumed that **all individuals in an age group have the same birth and survival rates.**

- In the Leslie model, **a population is considered to be closed to migration.**

- In the Leslie model, **one sex of the population is considered.** This is usually the female component of the population. Therefore, the model assumes there is no shortage of males that could inhibit reproductive potential [24].

- **There is no maximum size that the population can reach**, assuming the population can increase in an unlimited environment.

- **The birth rates, survival rates and initial population distribution are known.**

2.3.2 Constructing the Leslie model

The Leslie model divides age, which is a continuous variable, into distinct classes. Let i represent an age class in a population consisting of m age classes. Then define $P_i(t)$ as the number of individuals at time t in an age class i with $i \in \{1, \ldots, m\}$. For each age class i, there is a constant birth rate b_i and constant survival rate s_i. Birth rate b_i is the average number of offspring produced by individuals in age class i. Survival rate s_i is the rate of individuals from an age class i at time t to the age class $i + 1$ at time $t + 1$.

The survival rate represents the probability that an individual in age class i survives to the subsequent age class. Therefore, the age-specific death rates d_i are defined as $d_i = 1 - s_i$. It is assumed that a unit of time is the same as an age interval.

Since the Leslie model considers only the female sex of the population, the birth rates b_i consist of female offspring and the survival rates s_i represents the survival rates for female offspring. Moreover, $P_i(t)$ represents female individuals at time t in age class i.

In this model, a female population is divided into age classes of the same length. So if the maximum age attained by any women in a female population is denoted by A years and the population is divided into m age classes, then each class has a duration of A/m years. Table 2.1 shows the age classes with their corresponding age ranges for each age class $i \in \{1, \ldots, m\}$.

Age Class	Age Range
1	$[0, \frac{A}{m})$
2	$[\frac{A}{m}, \frac{2A}{m})$
3	$[\frac{2A}{m}, \frac{3A}{m})$
$\vdots$	$\vdots$
$m-1$	$[\frac{(m-2)A}{m}, \frac{(m-1)A}{m})$
m	$[\frac{(m-1)A}{m}$, A]

Table 2.1: Age classes $i \in \{1, \ldots, m\}$ and corresponding ranges.

Processes between two consecutive observation times can be explained using the demographic parameters b_i and s_i. By definition, it is obtained that $b_i \geq 0$ for $i = 1, ..., m$ and $0 < s_i \leq 1$ for $i = 1, ..., m-1$. We defined $P_i(t)$ as the number of females in age group i at time t. Then $P(t)$, the total population at a time t, is defined as

$$P(t) = \sum_{i=1}^{m} P_i(t). \tag{2.1}$$

We define the age distribution vector $P(t)$ at time t as

$$P(t) := \begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ \vdots \\ P_m(t) \end{bmatrix}. \tag{2.2}$$

Then the initial age distribution vector $P(0)$, at time $t = 0$ is defined as

$$P(0) := \begin{bmatrix} P_1(0) \\ P_2(0) \\ P_3(0) \\ \vdots \\ P_m(0) \end{bmatrix}. \tag{2.3}$$

The objective is to project the population from time t to $t+1$. The number of offspring produced by each class can be calculated by multiplying the reproductive rate by the number of females in that particular age class. The sum of all these values gives the total number of offspring produced.

$$P_1(t) = b_1P_1(t-1) + b_2P_2(t-1) + ... + b_mP_m(t-1) = \sum_{i=1}^{m} b_iP_i(t-1). \tag{2.4}$$

The number of females in the second age class at time t, are those in the first age class at time $t-1$, who are still alive at time t. In general, if at time $t = 0$ there are $P_i(0)$ females alive in age group i to $i+1$. Then the survivors of these females will form the group of females in age group $i+1$ to $i+2$ at time $t = 1$. Thus $P_i(1) = s_iP_i(0)$. For the calculation of $P_m(t)$, the population in age class m at time t, s_mP_m is added. This represents the possibility of surviving beyond the last age group of the model. The following equations are obtained:

$$\begin{aligned} P_i(t) &= s_{i-1}P_{i-1}(t-1), \quad i = 2, ..., m-1 \\ P_m(t) &= s_{m-1}P_{m-1}(t-1) + s_mP_m(t-1). \end{aligned} \tag{2.5}$$

From Equation (2.4) and (2.5), we get the following system of equations

$$\begin{cases} P_1(t) = & b_1P_1(t-1) + b_2P_2(t-1) + ... + b_mP_m(t-1) \\ P_2(t) = & s_1P_1(t-1) \\ P_3(t) = & s_2P_2(t-1) \\ \vdots & \\ P_m(t) = & s_{m-1}P_{m-1}(t-1) + s_mP_m(t-1). \end{cases} \tag{2.6}$$

The system of linear equations (2.6) can be rewritten in matrix form as

$$\begin{bmatrix} P_1(t) \\ P_2(t) \\ P_3(t) \\ \vdots \\ P_m(t) \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{m-1} & b_m \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & s_m \end{bmatrix} \begin{bmatrix} P_1(t-1) \\ P_2(t-1) \\ P_3(t-1) \\ \vdots \\ P_m(t-1) \end{bmatrix}. \tag{2.7}$$

In matrix notation,

$$P(t) = \boldsymbol{L}P(t-1), \tag{2.8}$$

where

$$\boldsymbol{L} := \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{m-1} & b_m \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & s_m \end{bmatrix}. \tag{2.9}$$

The Matrix $\boldsymbol{L}$ in Equation (2.8) is called the Leslie Matrix, consisting of m rows and m columns. All elements of the matrix are zero, except those in the first row and the subdiagonal below the principal diagonal of the matrix. The Leslie matrix is a projection matrix and can be described by the addition (2.10) of a Fertility matrix $\boldsymbol{F}$, and Survival matrix $\boldsymbol{S}$. The Fertility matrix describes the reproduction dynamics of a population, consisting of the age specific birth rates of a population. The survival matrix describes the transitions between age classes, consisting of the age specific survival rates of a population:

$$\boldsymbol{L} = \boldsymbol{F} + \boldsymbol{S} \tag{2.10}$$

$$\boldsymbol{F} := \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{m-1} & b_m \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad \boldsymbol{S} := \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & s_m \end{bmatrix}. \tag{2.11}$$

In this research, we will consider $s_m = 0$, meaning that the probability of surviving from age class m is zero. This causes the Leslie model to capture a population as a whole and implies that individuals do not age indefinitely. So the maximum age reached by any female in a population is in age class m. As a result, an adjusted Leslie matrix is obtained as follows:

$$\boldsymbol{L} := \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{m-1} & b_m \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & 0 \end{bmatrix}. \tag{2.12}$$

From Equation (2.8), it can be deduced that

$$\begin{aligned} P(1) &= \boldsymbol{L}P(0), \\ P(2) &= \boldsymbol{L}P(1) = \boldsymbol{L}^2P(0), \\ P(3) &= \boldsymbol{L}P(2) = \boldsymbol{L}^2P(1) = \boldsymbol{L}^3P(0), \\ &\vdots \\ P(k) &= \boldsymbol{L}P(k-1) = \boldsymbol{L}^kP(0). \end{aligned}$$

Thus, if the starting age distribution $P(0)$ and the Leslie matrix $\boldsymbol{L}$ are known, the age distribution of females at any later time can be determined using the following equation:

$$P(k) = \boldsymbol{L}^kP(0). \tag{2.13}$$

2.4 Eigenvalues and eigenvectors of the Leslie matrix

Equation (2.13) gives the distribution and size of a population at any given time, it does not however give an insight into the dynamics of the population growth process. In order to do so, the eigenvalues and eigenvectors of the Leslie matrix (2.12) will be analysed. We will first recall the definitions of eigenvalues and eigenvectors.

Definition 2.1. *[10] Let $\boldsymbol{A}$ be an $n \times n$ matrix. A scalar λ is an eigenvalue of $\boldsymbol{A}$ if there is a non-zero column vector v in $\mathbb{R}^n$ such that $\boldsymbol{A}v = \lambda v$. The vector v is then an eigenvector of, $\boldsymbol{A}$ corresponding to λ. The equation $\boldsymbol{A}v = \lambda v$ can be written as $\boldsymbol{A}v - \lambda \boldsymbol{I}v = 0$, where $\boldsymbol{I}$ is the $n \times n$ identity matrix. This is equivalent to $(\boldsymbol{A} - \lambda\boldsymbol{I})v = 0$. Thus, eigenvector v must be a solution of the homogeneous linear system*

$$(\boldsymbol{A} - \lambda\boldsymbol{I})v = 0. \tag{2.14}$$

An eigenvalue of $\boldsymbol{A}$ *is therefore a scalar* λ *for which system (2.14) has a non-zero solution* v*. The eigenvalues are the solution of the characteristic equation*

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0. \tag{2.15}$$

An eigenpair, denoted as (λ, v) *consists of an eigenvalue* λ *and its corresponding eigenvector* v*.*

2.4.1 Derivation of the characteristic equation of the Leslie model

Leslie aimed to comprehend the system's behaviour over time by seeking a solution $P(t) = r^t v$ [1], that exhibits geometric growth or decline. Geometric growth or decline occurs when a population experiences a consistent proportional increase or decrease in size from one age class to the next age class. Thus, the objective is to determine the age distribution v such that it is consistently multiplied by a constant value, r, at each time step. Hence, the number r and v must satisfy

$$\boldsymbol{L}v = rv. \tag{2.16}$$

The objective is to find the age distribution, v which at each time step is multiplied by a constant r. Such distributions are called stable. Therefore, by Definition 2.1, we are looking for eigenvalue r and its corresponding eigenvector v. By using Equation (2.12), Equation (2.16) can be rewritten as

$$\begin{cases} b_1v_1 + b_2v_2 + \cdots + b_mv_m = rv_1, \\ s_1v_1 = rv_2,\ s_2v_2 = rv_3, \ldots, s_{m-1}v_{m-1} = rv_m. \end{cases} \tag{2.17}$$

Now, $v_2, \ldots, v_m$ can be expressed as

$$v_2 = \frac{s_1v_1}{r}, \quad v_3 = \frac{s_1s_2v_1}{r^2}, \ldots, v_m = \frac{s_1s_2\ldots s_{m-1}v_1}{r^{m-1}}. \tag{2.18}$$

Substituting $v_1, \ldots, v_m$ from Equation (2.18) into the first Equation from (2.17), dividing by v_1 and multiplying by r^{m-1}, Leslie obtained the characteristic equation

$$r^m = b_1r^{m-1} + s_1b_2r^{m-2} + \cdots + s_1s_2\ldots s_{m-1}b_m. \tag{2.19}$$

Dividing by r^m and simplification gives the characteristic equation of a Leslie matrix $\boldsymbol{L}$.

$$1 - b_1r^{-1} - s_1b_2r^{-2} - \cdots - s_1s_2\ldots s_{m-1}b_mr^{-m} = 0. \tag{2.20}$$

Chapter 3

Perron-Frobenius theorem

Chapter 1 provided a description of the Leslie model for population growth, and the key assumptions underlying its formulation. In this chapter, eigenvalues and eigenvectors of the Leslie matrix will be utilized to explore the long-term dynamics of a population, through the implications of the Perron-Frobenius theorem.

3.1 Derivation of the Perron-Frobenius theorem

3.1.1 Perron's theorem

Definition 3.1. *Let $\sigma(\boldsymbol{A})$ denote the set of all eigenvalues λ of a matrix $\boldsymbol{A}$. Then $\rho(\boldsymbol{A})$, the spectral radius of a matrix $\boldsymbol{A}$, is defined as*

$$\rho(\boldsymbol{A}) = max\{|\lambda| : \lambda \in \sigma(\boldsymbol{A})\}. \tag{3.1}$$

Definition 3.2. *A matrix $\boldsymbol{A} \in \mathbb{R}^{m\times m}$ is said to be a non-negative matrix whenever each $a_{ij} \geq 0$. It is said to be a positive matrix whenever each $a_{ij} > 0$.*

Theorem 3.3. Perron's theorem[22] If $\boldsymbol{A}$ is a positive matrix with $r = \rho(\boldsymbol{A})$, then the following statements are true.

- $r > 0$.
- $r \in \sigma(\boldsymbol{A})$, and r is called the Perron root.
- $alg\ mult_{A}(r) = 1$.
- There exists an eigenvector $x > 0$ such that $\boldsymbol{A}x = rx$.
- The Perron vector is the unique vector p defined by

  $$\boldsymbol{A}p = rp, \quad p > 0, \quad and \quad \|p\|_1 = 1,$$

  and, except for positive multiples of p, there are no other non-negative eigenvectors for $\boldsymbol{A}$.

Perron's theorem applies to positive matrices, giving conditions for the existence of a dominant eigenvalue: the eigenvalue with the greatest absolute value. However, since the Leslie matrix (2.12) is a non-negative matrix, Perron's theorem cannot be applied to the Leslie model. Frobenius extended Perron's theorem to non-negative matrices, by incorporating the additional requirement of irreducibility.

Definition 3.4. *The dominant eigenvalue* λ_d *is the eigenvalue with the largest absolute value:* $|\lambda_d| \geq |\lambda|$. *The eigenvalue is called strictly dominant if* $|\lambda_d| > |\lambda|$.

3.1.2 Irreducibility of matrices

Definition 3.5. *[9] [11] Matrix* $\boldsymbol{A}$ *is reducible if there exists a permutation matrix* $\boldsymbol{P}$ *such that*

$$\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{0} & \boldsymbol{Z} \end{bmatrix}, \tag{3.2}$$

where $\boldsymbol{X}$ *and* $\boldsymbol{Z}$ *are both square matrices. Thus, a matrix* $\boldsymbol{A}$ *is reducible if and only if it can be transformed into block upper-triangular form through permutations of rows and columns. If no such* $\boldsymbol{P}$ *exists,* $\boldsymbol{A}$ *is called an irreducible matrix.*

Definition 3.6. *[22] Let* $G(\boldsymbol{A})$ *denote the graph associated with matrix* $\boldsymbol{A}$. *Then* $G(\boldsymbol{A})$ *is defined to be the graph on* n *nodes* $\{N_1, \ldots, N_n\}$ *in which there is a directed edge leading from* N_i *to* N_j *if and only if* $a_{ij} \neq 0$.

Definition 3.7. *[22]* $G(\boldsymbol{A})$ *is called connected if, for each pair of nodes,* (N_i, N_k) *there is a sequence of edges from* N_i *to* N_k .

Lemma 3.8. *[22] A non-negative matrix is irreducible if and only if it is connected.*

Lemma 3.8 implies that the graph associated with the non-negative irreducible matrix contains a path from every node to every other node [4]. In the context of the Leslie matrix for population growth, this then indicates that the life cycle graph, associated with the Leslie matrix, contains a path from every age class to every other age class. This implies that individuals in one age class, have a positive probability to eventually reach all other age classes in the model, through a series of birth and survival rates.

Thus, for a reducible matrix, this implies that the life cycle must contain at least one age class, for which there exists another age class that is impossible to be reached through a series of birth and survival rates. An example of this is a Leslie model for a population that includes post-reproductive age classes; age classes for which $b_i = 0$ and thereby cannot contribute to younger age classes. Hence, in order for the Leslie matrix $\boldsymbol{L}$ in Equation (2.12) to be irreducible, it should hold that $b_m > 0$.

To derive the Perron-Frobenius theorem, the utilization of the following lemmas is necessary.

Lemma 3.9. *[22] If a non-negative* $n \times n$ *matrix* $\boldsymbol{A}$ *is irreducible, then* $(\boldsymbol{I} + \boldsymbol{A})^{n-1} > 0$, *where* $\boldsymbol{I}$ *is the* $n \times n$ *Identity matrix.*

Proof. Let $\boldsymbol{A}$ be a non-negative irreducible matrix of size $n \times n$. Let $a_{ij}^{(k)}$ denote the (i, j)-entry in matrix $\boldsymbol{A}^k$. Since matrix $\boldsymbol{A}$ is irreducible, there exists a path of length k from state i to state j in the directed graph associated with matrix $\boldsymbol{A}$. Thus, there exists $0 \leq k \leq n-1$ such that $a_{ij}^{(k)} > 0$. Now, $(\boldsymbol{I} + \boldsymbol{A})^{n-1}$ can be expanded by the Binomial theorem:

$$(\boldsymbol{I} + \boldsymbol{A})^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} \boldsymbol{I}^{(n-1-k)} \boldsymbol{A}^k = \sum_{k=0}^{n-1} \binom{n-1}{k} \boldsymbol{A}^k.$$

Thus, for each (i, j)- entry in $(I + A)^{n-1}$ this implies

$$\left[(\boldsymbol{I} + \boldsymbol{A})^{n-1}\right]_{ij} = \left[\sum_{k=0}^{n-1} \binom{n-1}{k} \boldsymbol{A}^k\right]_{ij} = \sum_{k=0}^{n-1} \binom{n-1}{k} a_{ij}^{(k)} > 0.$$

Since each (i,j) entry in $(\boldsymbol{I}+\boldsymbol{A})^{n-1}$ is positive, the matrix $(\boldsymbol{I}+\boldsymbol{A})^{n-1} > 0$ is positive. □

Lemma 3.10. *[22] Let matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$. If $\boldsymbol{A} \leq \boldsymbol{B}$, thus $a_{ij} \leq b_{ij}$ for each i and j, then $\rho(\boldsymbol{A}) \leq \rho(\boldsymbol{B})$.*

Lemma 3.11. *[22] Let $\boldsymbol{A}$ be a non-negative matrix of size $n \times n$. The following statements are true*

$$\begin{aligned} &r = \rho(\boldsymbol{A}) \in \sigma(\boldsymbol{A}). \\ &\boldsymbol{A}z = rz \text{ for some nonnegative vector } z \text{ such that } (z)_i \geq 0 \text{ and } z \neq 0 \quad \forall i \in \{1, \cdots n\}. \end{aligned} \tag{3.3}$$

Proof. Let $\boldsymbol{A}$ be a non-negative matrix of size $n \times n$ and let $r = \rho(\boldsymbol{A})$. Consider the sequence of positive matrices $\boldsymbol{A}_k = \boldsymbol{A} + (\frac{1}{k})\boldsymbol{E}$, where $\boldsymbol{E}$ is the $n \times n$ matrix in which each element equals to 1. Then $\boldsymbol{A}_k > 0$. Since $\boldsymbol{A}_k$ is a positive matrix, Perron's theorem for positive matrices 3.3 can be applied. Let $r_k > 0$ denote the Perron root and $p_k > 0$ denote the Perron vector.

Since for all Perron vectors p_k, $\|p_k\|_1 = 1$, the sequence $\{p_k\}_{k=1}^{\infty}$ is bounded. By the Bolzano-Weierstrass theorem, each bounded sequence in $\mathbb{R}^n$, has a convergent subsequence:

$$\lim_{i\to\infty} \{p_{k_i}\}_{i=1}^{\infty} = z, \quad \text{where } z \geq 0, \text{ and } z \neq 0. \tag{3.4}$$

Since $\boldsymbol{A}_1 > \boldsymbol{A}_2 > \cdots > \boldsymbol{A}$, Lemma 3.10 implies that $r_1 \geq r_2 \geq \cdots \geq r$. Therefore, $\{r_k\}_{k=1}^{\infty}$ is a monotonic sequence of positive numbers and is bounded from below by r. Then by the monotone convergence theorem

$$\lim_{k\to\infty} r_k = r^* \text{ exists, and } r^* \geq r. \tag{3.5}$$

Thus, for a subsequence $\{r_{k_i}\}_{i=1}^{\infty}$

$$\lim_{i\to\infty} r_{k_i} = r^* \text{ exists, and } r^* \geq r \tag{3.6}$$

Note that $\lim_{k\to\infty} \boldsymbol{A}_k = \lim_{k\to\infty} \boldsymbol{A} + (\frac{1}{k})\boldsymbol{E} = \boldsymbol{A}$, implying also that $\lim_{i\to\infty} \boldsymbol{A}_{k_i} = \boldsymbol{A}$. Then by Perron's theorem 3.3 and using that the limit of a product is the product of the limits, we obtain

$$\boldsymbol{A}z = \lim_{i\to\infty} \boldsymbol{A}_{k_i} p_{k_i} = \lim_{i\to\infty} r_{k_i} p_{k_i} = r^* z.$$

Thus, by definition of an eigenvalue 2.1 $r^* \in \sigma(\boldsymbol{A})$. From the definition of the spectral radius 3.1, it follows that $r^* \leq r$. Now, since $r^* \leq r$ and $r^* \geq r$, we can conclude that $r^* = r$ and $\boldsymbol{A}z = rz$ with $z \geq 0$ and $z \neq 0$. □

3.1.3 Perron-Frobenius theorem

The Perron-Frobenius theorem states that a nonnegative irreducible matrix has a positive algebraically simple dominant eigenvalue. This eigenvalue corresponds to a unique positive eigenvector.

Theorem 3.12. (Perron-Frobenius theorem) [22] If $\boldsymbol{A}$ is a nonnegative irreducible matrix, then the following statements are true.

- $r = \rho(\boldsymbol{A}) \in \sigma(\boldsymbol{A})$ and $r > 0$.
- $\mathit{alg\ mult}_{\boldsymbol{A}}(r) = 1$.

- There exists an eigenvector $x > 0$ such that $\boldsymbol{A}x = rx$.
- The unique vector defined by

  $$\boldsymbol{A}p = rp, \quad p > 0, \quad \text{and} \quad \|p\|_1 = 1,$$

  is called the Perron vector. There are no non-negative eigenvectors for $\boldsymbol{A}$ except for positive multiples of p, regardless of the eigenvalue.

Proof. Let $\boldsymbol{A}$ be a non-negative irreducible matrix of size $n \times n$. From Lemma 3.11 it follows that $r = \rho(\boldsymbol{A}) \in \sigma(\boldsymbol{A})$.

Now to prove that r has algebraic multiplicity 1, define $\boldsymbol{B} = (\boldsymbol{I} + \boldsymbol{A})^{n-1}$. Then, by Lemma 3.9, matrix $\boldsymbol{B}$ is positive.

Lemma 3.13. *[22] Let f be a polynomial. if $\lambda \in \sigma(\boldsymbol{A})$, then $f(\boldsymbol{A}) \in \sigma(f(\boldsymbol{A}))$. Thus, if λ is an eigenvalue for matrix $\boldsymbol{A}$, then $f(\lambda)$ is an eigenvalue for $f(\boldsymbol{A})$, whenever $f(\boldsymbol{A})$ exists. Furthermore, it holds that alg mult$_{\boldsymbol{A}}(\lambda) =$ alg mult$_{f(\boldsymbol{A})}(f(\lambda))$.*

From this spectral mapping property, it follows that

$$\lambda \in \sigma(\boldsymbol{A}) \leftrightarrow (1+\lambda)^{n-1} \in \sigma(\boldsymbol{B}). \tag{3.7}$$

$$\textit{alg mult}_{\boldsymbol{A}}(\lambda) = \textit{alg mult}_{\mathbf{B}}((1+\lambda)^{n-1}). \tag{3.8}$$

Now, denote $\mu = \rho(\boldsymbol{B})$. Using the definition of the spectral radius 3.1, μ can be expressed as

$$\begin{aligned} \mu &= \max\{|(1+\lambda)^{n-1}| : (1+\lambda)^{n-1} \in \sigma(\boldsymbol{B})\} \\ &= \max\{|(1+\lambda)|^{n-1} : (1+\lambda)^{n-1} \in \sigma(\boldsymbol{B})\} \\ &= \max\{|(1+\lambda)|^{n-1} : \lambda \in \sigma(\boldsymbol{A})\}, \quad \text{by the equivalence relation in equation (3.7) .} \\ &= (\max|(1+\lambda)| : \lambda \in \sigma(\boldsymbol{A}))^{n-1} \\ &= (1+\rho(\boldsymbol{A}))^{n-1} \\ &= (1+r)^{n-1} \end{aligned}$$

Now, since matrix $\boldsymbol{B}$ is positive, Perron's theorem (3.1) can be applied. Therefore, *alg mult*$_{\boldsymbol{B}}(\mu)$ = 1. Thus *alg mult*$_{\boldsymbol{B}}((1+r)^{n-1}) = 1$. Then it follows by Equation (3.8) that *alg mult*$_{\boldsymbol{A}}(r) = 1$.

To show that there exists a positive eigenvector, corresponding with r, note that by Lemma 3.11, $\boldsymbol{A}z = rz$ for some $z \geq 0$ and $z \neq 0$. So there exists a non-negative eigenvector x associated with r, that form the eigenpair (r, x) for matrix $\boldsymbol{A}$. By the spectral mapping property, it then follows that (μ, x) is an eigenpair for matrix $\boldsymbol{B}$. By Perron's theorem for positive matrix $\boldsymbol{B}$, there are no non-negative eigenvectors for matrix $\boldsymbol{B}$, other than the Perron vector and its positive multiples. Therefore, x must be a positive multiple of the Perron vector of $\boldsymbol{B}$ and thus x must be positive.

To show that this vector is unique, let (λ, y) be an eigenpair for $\boldsymbol{A}$, such that $y \geq 0$. Then for the left Perron vector $x > 0$ of $\boldsymbol{A}$, $x^T y > 0$. We obtain that

$$rx^T = x^T \boldsymbol{A}, \tag{3.9}$$

$$rx^T y = x^T \boldsymbol{A} y = x^T \lambda y = \lambda x^T y. \tag{3.10}$$

Since $x^T y > 0$, it follows that $r = \lambda$. Thus, there is no other non-negative eigenvector other than the Perron vector and its positive multiples.

Now it can be proved that $r > 0$. Suppose $r = 0$, then $\boldsymbol{A}x = rx = 0$. Since $\boldsymbol{A} \geq 0$ and $x > 0$, we obtain $\boldsymbol{A}x = 0$, only if $\boldsymbol{A}$ is the Null matrix. However, this gives a contradiction for the assumption that matrix $\boldsymbol{A}$ is irreducible. ☐

3.2 Analysis of Leslie model via the Perron-Frobenius theorem

For irreducible and non-negative Leslie matrix $\boldsymbol{L}$ from (2.12) the Perron-Frobenius theorem 3.12 can be applied to guarantee that there is a dominant eigenvalue, that is the spectral radius $\rho(\boldsymbol{L}) = r$. The application of Perron-Frobenius theorem 3.12 to the Leslie model for population growth, will provide insight into the long-term behaviour and stability of the population. The limiting behaviour of the Leslie model, as a result of Perron-Frobenius theorem, will be demonstrated. Let us assume Leslie matrix $\boldsymbol{L}$ is diagonalizable. Thus, there exist an invertible matrix S and diagonal matrix $\boldsymbol{D}$ such that $\boldsymbol{L} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}$, where $\boldsymbol{D}$ is the diagonal matrix with eigenvalues $\lambda_1, \ldots, \lambda_m$ on its diagonal and matrix S consists of the eigenvectors corresponding to the eigenvalues in matrix $\boldsymbol{D}$. Then $\boldsymbol{L}^k$ can be expressed as follows:

$$\boldsymbol{L}^k = \boldsymbol{S}\boldsymbol{D}^k\boldsymbol{S}^{-1}. \tag{3.11}$$

Now, substituting $\boldsymbol{L}^k$ into Equation (2.13) gives

$$P(k) = \boldsymbol{S}\boldsymbol{D}^k\boldsymbol{S}^{-1}P(0), \tag{3.12}$$

For $\boldsymbol{D}^k$ the following matrix is obtained

$$\boldsymbol{D}^k := \begin{bmatrix} \lambda_1^k & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2^k & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_3^k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_m^k \end{bmatrix}. \tag{3.13}$$

Dividing Equation (3.12) by λ_1^k gives

$$\frac{P(k)}{\lambda_1^k} = \boldsymbol{S} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{\lambda_2^k}{\lambda_1^k} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{\lambda_3^k}{\lambda_1^k} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{\lambda_m^k}{\lambda_1^k} \end{bmatrix} \boldsymbol{S}^{-1}P(0). \tag{3.14}$$

Since λ_1 is the dominant eigenvalue, by definition it holds that $|\lambda_1| > |\lambda_i|$. Then $|\frac{\lambda_i}{\lambda_1} < 1|$ for $i = 1, \ldots, m$. Therefore strict dominance of λ_1 ensures that $\lim_{k \to \infty} \frac{\lambda_i^k}{\lambda_1^k} = 0$ for $i = 2, \ldots, m$. Therefore taking a limit from both sides of (3.14) gives

$$\lim_{k\to\infty}\frac{P(k)}{\lambda_1^k}=\lim_{k\to\infty}\boldsymbol{S}\begin{bmatrix}1 & 0 & 0 & \cdots & 0 & 0\\ 0 & \frac{\lambda_2^k}{\lambda_1^k} & 0 & \cdots & 0 & 0\\ 0 & 0 & \frac{\lambda_3^k}{\lambda_1^k} & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & \frac{\lambda_m^k}{\lambda_1^k}\end{bmatrix}\boldsymbol{S}^{-1}P(0). \tag{3.15}$$

Then using that $\lim_{k\to\infty}\frac{\lambda_i^k}{\lambda_1^k}=0$ for $i=2,\ldots,m$, Equation (3.15) is simplified as follows:

$$\lim_{k\to\infty}\frac{P(k)}{\lambda_1^k}=\boldsymbol{S}\begin{bmatrix}1 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & 0\end{bmatrix}\boldsymbol{S}^{-1}P(0). \tag{3.16}$$

The multiplication $\boldsymbol{S}^{-1}P(0)$ is a vector with constant entries

$$\boldsymbol{S}^{-1}P(0)=\begin{bmatrix}c_1\\ c_2\\ c_3\\ \vdots\\ c_m\end{bmatrix}, \tag{3.17}$$

Using Equation (3.17) we obtain

$$\boldsymbol{S}\begin{bmatrix}1 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & 0\end{bmatrix}\boldsymbol{S}^{-1}P(0)=\boldsymbol{S}\begin{bmatrix}c_1\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix}=c_1v_1. \tag{3.18}$$

Substituting Equation (3.18) into Equation (3.16), the following equation, known as the strong Ergodic theorem [8], is obtained.

$$\lim_{k\to\infty}P(k)=\lim_{k\to\infty}\lambda_1^k c_1 v_1 \tag{3.19}$$

Simplifying (3.19) and denoting the constant c_1 as c gives

$$\lim_{k\to\infty}P(k)=\lim_{k\to\infty}\lambda_1^k c v_1. \tag{3.20}$$

Equation (3.20) remains true for non-diagonalizable Leslie matrices $\boldsymbol{L}$ [13]. For large values of k, (so long-term), Equation (3.20) provides the following approximation

$$P(k)\cong\lambda_1^k c v_1. \tag{3.21}$$

From this approximation, it can be observed that the population vector $P(k)$, thus the total population at time k, will increase exponentially if $\lambda_1 > 1$.

If $\lambda_1 = 1$, Equation (3.21) results in $P(k) \cong cv_1$ for all k. Thus, the population is stable and stationary. Thus, if a population, with Leslie matrix $\boldsymbol{L}$, has a dominant eigenvalue λ_1 that equals 1, the long-term behaviour of the population is stable. This implies that the births exactly compensate the deaths in a population per time interval.

If $\lambda_1 < 1$, the total population $P(k)$ is declining exponentially. Since $\lim_{k\to\infty} \lambda_1^k = 0$, the population will eventually face extinction.

Equation (3.21) yields an equation that provides an approximation of the population at time $k-1$.

$$P(k-1) \cong \lambda_1^{k-1} c v_1. \tag{3.22}$$

Combining Equations (3.21) and (3.22), we obtain

$$P(k) \cong \lambda_1 P(k-1). \tag{3.23}$$

Equation (3.23) indicates that for large time values, each age distribution vector is approximately a scalar multiple of the previous age distribution vector. This scalar multiple λ_1 is the dominant eigenvalue of the Leslie matrix. Thus, the asymptotic growth rate of a population is given by the dominant eigenvalue λ_1.

The dominant eigenvalue and corresponding eigenvector represent the long-term behaviour of a population; the dominant eigenvalue indicates the long-term population growth, and the corresponding eigenvector indicates the long-term age distribution. Hence, the asymptotic growth rate of the population is given by this dominant eigenvalue. Equation (6) implies that the population distribution will remain constant. So, even though the total population can increase or decrease, the age distribution of the population will remain constant. Thus, the proportion of females in each age group will be constant.

The dominant eigenvalue determines the ergodic properties of population growth. A population is said to be ergodic if its eventual behaviour is independent of its initial state [7]. Regardless of the initial population, the population will grow with a rate λ_1 and with a population distribution that is proportional to v_1, the corresponding eigenvector of λ_1.

Thus, the stable population is the age distribution to which a population will evolve to, if its age specific rates of birth and survival were to continue indefinitely. The growth rate λ_1, is defined as the factor that the stable population is multiplied by each year. Therefore, each age class will change by the same constant.

The dominant eigenvalue is the growth rate of a population after the population reaches a stable distribution. It can be concluded that a population described by a Leslie matrix (2.12), thus with constant age specific birth rates and survival rates, asymptotically reaches a stable age distribution that is independent of its initial age structure. Once a population reaches this stable age distribution, the population continues to grow in the rate of the dominant eigenvalue λ_1, while keeping the same age distribution.

3.3 An example 4×4 Leslie model

Consider the 4×4 example Leslie model described by the Leslie matrix in Equation (3.24) and initial age distribution vector at time $t = 0$ in Equation (3.25).

$$\mathbf{L} := \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \end{bmatrix} \tag{3.24}$$

$$P(0) := \begin{bmatrix} P_1(0) \\ P_2(0) \\ P_3(0) \\ P_4(0) \end{bmatrix} = \begin{bmatrix} 100 \\ 200 \\ 200 \\ 300 \end{bmatrix} \tag{3.25}$$

In order to apply Perron-Frobenius theorem 3.12, two conditions must be met: non-negativity and irreducibility. Firstly, the matrix in Equation (3.24) satisfies the non-negativity condition. Secondly, as it is possible to reach every age class from every other age class through a sequence of transitions, the Leslie matrix also fulfills the irreducibility criterion. Consequently, with the Leslie matrix meeting both non-negativity and irreducibility criteria, the Perron-Frobenius theorem implies the existence of a dominant eigenvalue that is both real and positive.

Calculations show that the dominant eigenvalue is 1.744. Once the population reaches a stable distribution, using the result in Equation (6), it follows that the population will grow with a rate of 1.744, since the dominant eigenvalue $\lambda = 1.744 > 1$. The asymptotic growth rate is given by 1.744

Furthermore, by Perron-Frobenius theorem, there corresponds a unique positive eigenvector to this dominant eigenvalue. Once the stable population is reached, future age distributions are scalar multiples. Therefore, the population will grow with a distribution proportional to the eigenvector corresponding to 1.744, which is $v = \begin{bmatrix} 176.808 & 50.691 & 8.720 & 1 \end{bmatrix}^T$. This eigenvector indicates that the first age class will contain the majority of individuals in the stable age distribution. This can be explained by the fact the older age classes exhibit the higher birth rates while having the lower survival rates.

Chapter 4

Doeblin's theorem

To gain more insight into the long-term analysis of the Leslie model, Doeblin's theorem for ergodicity will be investigated in the following chapter. Doeblin's theorem is a mathematical result that provides insight into the long-term behaviour and convergence properties of Markov chains. Markov chains have proven to be important mathematical tools in modelling various processes, including population dynamics. In this chapter, the application of Doeblin's theorem in the context of the Leslie model will be explored.

This chapter will provide a detailed examination of the assumptions and implications of Doeblin's theorem. By analysing the Leslie model, using Doeblin's theorem, we can gain understanding of the stability, convergence and potential patterns that can emerge within population dynamics. This chapter will illustrate the practical utility of Doeblin's theorem, while aiming to enhance our understanding of the Leslie model. Additionally, this allows us to see that different theorems and mathematical tools can provide complementary perspectives, compared to insights gained by Perron-Frobenius theorem.

4.1 Derivation of Doeblin's theorem

Doeblin's theorem facilitates the study of the long-term distribution of Markov chains. For this reason, the following section will introduce key concepts and definitions of Markov chains, fundamental to understanding and deriving Doeblin's theorem.

4.1.1 Markov chains

Markov chains model systems, that are characterized by transitions between different states. A discrete Markov chain is a discrete time stochastic process taking its values in a finite or countably finite set S, called the state space. These states represent the possible conditions of the system being modelled. The system moves from one state to another over discrete time steps. These transitions are governed by transition probabilities, which determine the likelihood of moving from one state to another [12].

Definition 4.1. *[12] Let S be a countable state space and let $\mathbf{X} = (X_n : n \geq 0)$ be a sequence of random variables taking values in S. The sequence $\mathbf{X}$ is called a Markov chain if it satisfies the Markov property*

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) \tag{4.1}$$

for all $n \geq 0$ and all $i_0, i_1, \dots, i_{n+1} \in S$.

A stochastic process is said to have the Markov property if, conditional on its present value, its future is independent of its past [12]. Once in a certain state at a given time, the way in which the system reached that state does not at all affect its future evolution [14].

The transition probabilities of a Markov chain are organized in a transition probability matrix. This matrix captures the probabilities of transitioning from one state to another in a single time step. Each element of the matrix represents the probability of moving from the corresponding row state to the column state. The sum of probabilities in each row of the matrix is equal to 1, reflecting the fact that the system must move to one of the possible states. So, the probabilities in Equation (4.1) can be summarized in a transition probability matrix $\boldsymbol{P}$.

Definition 4.2. *[12] The transition probability matrix $\boldsymbol{P}$ is defined as*

$$\boldsymbol{P} = (p_{i,j} : i, j \in S), \quad \text{given by} \quad p_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i). \tag{4.2}$$

The probability transition matrix is a stochastic matrix as the following holds:

- $p_{i,j} \geq 0$ *for* $i, j \in S$, *and*
- $\sum_{j \in S} p_{i,j} = 1$ *for* $i \in S$ *by the total law of probability, so that $\boldsymbol{P}$ has row sums* 1.

Definition 4.3. *A Markov chain is called time homogeneous, if $\mathbb{P}(X_{n+1} = j | X_n = i) = p_{ij}$, $\forall n \geq 0$. The transition probabilities remain constant and are independent of the specific time step n. So The transition probabilities are not changing as a function of time.*

The initial distribution, denotes as μ, represents the probabilities of starting the Markov chain in each state. It is a probability vector where the entry μ_i corresponds to the probability of beginning in state $i \in S$.

Definition 4.4. *The initial distribution μ is defined as a row vector*

$$\mu = (\mu_i : i \in S) \quad \text{where} \quad \mu_i = \mathbb{P}(X_0 = i). \tag{4.3}$$

To determine the probability distribution of the Markov chain after multiple time steps, we introduce the n-step transition probability matrix $\boldsymbol{P}^{(n)}$. This matrix describes the probabilities of being in various states after n time steps. Each entry $p_{ij}^{(n)}$ represents the probability of transitioning from state i to state j in exactly n steps.

Definition 4.5. *The n-step transition probability matrix $\boldsymbol{P}^{(n)}$ is defined as*

$$\boldsymbol{P}^{(n)} = (p_{ij}^{(n)} : i, j \in S) \quad \text{given by} \quad p_{i,j}^{(n)} = \mathbb{P}(X_{m+n} = j | X_m = i). \tag{4.4}$$

Then it follows that:

$$\mathbb{P}(X_n = j) = (\mu P^n)_j \quad n \geq 0 \quad \text{and} \quad j \in S \tag{4.5}$$

Theorem 4.6. (**Chapman-Kolmogorov equations**) [12] Let $\boldsymbol{P}$ be a transition probability matrix and let $\boldsymbol{P}^{(n)}$ be the n-step transition probability matrix. Then

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}, \quad \text{for } i, j \in S \text{ and } m, n \geq 0. \tag{4.6}$$

Therefore $\boldsymbol{P}^{(m+n)} = \boldsymbol{P}^{(m)} \boldsymbol{P}^{(n)}$. In particular $\boldsymbol{P}^{(2)} = \boldsymbol{P}^{(1)} \boldsymbol{P}^{(1)} = \boldsymbol{P}\boldsymbol{P} = \boldsymbol{P}^2$. Thus $\boldsymbol{P}^{(n)} = \boldsymbol{P}^n$.

The system's distribution in future time steps can be determined by iterative multiplication of the initial distribution by the transition probability matrix. Theorem (4.6) implies the n-step transition probabilities can be determined by calculating the nth power of transition probability matrix $\boldsymbol{P}$.

4.1.2 Convergence of Markov chains

Definition 4.7. *[12] Let* $\mathbf{X}$ *be a Markov chain on state space* S*, with transition probability matrix* $\boldsymbol{P}$*. The vector* $\pi = (\pi_i : i \in S)$ *is called a stationary distribution of the chain if*

- $\pi_i \geq 0$ *for all* $i \in S$*, and* $\sum_{i \in S} \pi_i = 1$,
- $\pi = \pi \boldsymbol{P}$.

Definition 4.8. *A state i is recurrent if, starting from i, the chain returns to i with probability 1. Furthermore, a state i is called positive-recurrent if the mean recurrence time is finite. If the mean recurrence time is infinite, the state i is called null-recurrent.*

Definition 4.9. *[12] The period of state i is given by*

$$d_i = gcd\{n : p_{i,i}(n) > 0\}. \tag{4.7}$$

A state i is called aperiodic if $d_i = 1$*, and periodic if* $d_i > 1$.

By Definition 4.9, a period of a state is defined by the greatest common divisor of the lengths of all possible oaths from this state back to itself. Hence, a state is aperiodic if it does not exhibit a repeating pattern in the number of steps it takes to return to itself.

Theorem 4.10. (Convergence theorem for discrete Markov chains) Consider a Markov chain $\mathbf{X} = (X_n : n \geq 0)$ on state space S. If $\mathbf{X}$ is aperiodic, irreducible, and positive recurrent, then

$$\lim_{n \to \infty} p_{ij} = \pi_j \quad \text{for } i, j \in S, \tag{4.8}$$

where π is the unique stationary distribution of the Markov chain.

The Convergence Theorem for discrete-time Markov chains states that if a Markov chain is aperiodic, irreducible, and positive recurrent, then it converges to a unique steady-state distribution. As the transition probabilities converge to the stationary distribution, it shows that for big enough n, the distribution becomes independent of its initial state.

In the context of the Leslie model, we are interested in a finite discrete Markov chain due to the finite number of age classes involved. In the case of a finite Markov chain, irreducibility implies positive recurrence. By section 3.1.2 it follows that for a finite Markov chain, irreducibility signifies that any state can be reached from any other state within a finite number of steps. This implies that each state can also return to itself within a finite number of steps. Thus, by Definition 4.8, this also implies that all states are positive recurrent. By irreducibility, any state can be reached from any other state in a path. Using that the Markov chain is finite, then there cannot be infinite paths. As irreducibility of a finite Markov chain implies positive recurrence automatically, Theorem 4.10 can be rewritten.

Theorem 4.11. (Convergence theorem for finite discrete Markov chains) Consider a finite Markov chain $\mathbf{X}$ on state space S. If $\mathbf{X}$ is aperiodic and irreducible, then

$$\lim_{n \to \infty} p_{ij} = \pi_j \quad \text{for } i, j \in S, \tag{4.9}$$

where π is the unique stationary distribution of the Markov chain.

The convergence theorem for finite Markov chains, gives insight into the long-term behaviour of a finite discrete Markov chain, by providing conditions for the existence of a unique stationary distribution. Whereas one of these conditions is irreducibility, Doeblin's theorem facilitates long-term analysis of a Markov chain that need not be irreducible.

Both the convergence theorem for discrete Markov chains and Doeblin's theorem provide conditions that imply the existence of a stationary distribution for a Markov chain. Thus, the Convergence Theorem and Doeblin's theorem share the objective of identifying conditions under which a Markov chain converges to a stable, long-term distribution. Despite their similarities, they offer distinct criteria for achieving this convergence, making them complementary mathematical tools for exploring the asymptotic behaviour of Markov chains.

Where Theorem 4.10 requires aperiodicity, irreducibility and positive recurrence, Doeblin's theorem is more flexible in terms of the types of Markov chains it can analyze. As it does not require irreducibility, Doeblin's theorem is applicable to Markov chains that may include an absorbing state. Moreover, Doeblin's theorem provides additional information on the convergence rate to the stationary distribution, which characterizes how quickly the chain approaches its limiting distribution.

4.1.3 Doeblin's theorem

While Doeblin's theorem does not impose the condition of irreducibility, it does necessitate the existence of a state, denoted as j_0, such that from any other state i, there must exist a positive probability of transitioning to state j_0 within a single step. Under this condition, state 'j' will be visited repeatedly over time. Consequently, as time progresses, the influence of the initial distribution of the chain gradually decreases.

Studying the long-term behaviour of a Markov chain, we are interested in $\mu \boldsymbol{P}^n$, the distribution of the Markov chain after n steps, as n goes to infinity. The distribution of the Markov chain after n steps $\mu \boldsymbol{P}^n$ can be expressed as the distribution after $n-m$ steps $\mu \boldsymbol{P}^{n-m}$, followed by m additional steps $\boldsymbol{P}^m$

$$\mu \boldsymbol{P}^n = (\mu \boldsymbol{P}^{n-m}) \boldsymbol{P}^m \tag{4.10}$$

As m increases, the effect of the initial distribution μ on the distribution after n steps becomes less significant compared to the effect of $\boldsymbol{P}^m$, the m-step transition matrix. This is because, as time progresses, the chain gets influenced more by the transitions between states governed by $\boldsymbol{P}^m$. Thus, as the number of steps increases, the distribution of the Markov chain becomes increasingly influenced by the transition matrix $\boldsymbol{P}$ itself, rather than its initial distribution. This implies that $\mu \boldsymbol{P}^n = (\mu \boldsymbol{P}^{n-m}) \boldsymbol{P}^m$ will closely approximate $\mu \mathbf{P}^m$ as m grows significantly. Consequently, it follows by Cauchy's convergence criterion [17] that the limit $\pi = \lim_{n\to\infty} \mu \boldsymbol{P}^n$ exists.

$$\pi = \lim_{n\to\infty} \mu \boldsymbol{P}^{n+1} = \lim_{n\to\infty} (\mu \boldsymbol{P}^n) \boldsymbol{P} = \pi \boldsymbol{P} \tag{4.11}$$

As m grows larger and larger, the distribution $\mu \boldsymbol{P}^n$ approaches a stable distribution, and the influence of the initial distribution diminishes. As a result, the distribution of the Markov chain after a very long time ($n \to \infty$) tends to converge to a fixed distribution by Cauchy's convergence criterion [17], denoted as π. Thus, as n tends to infinity, the difference between $\mu \boldsymbol{P}^n$ and π decreases. To measure this difference, the definition of the variation norm is introduced.

Definition 4.12. *[25] Let $\rho \in \mathbb{R}^S$ be a row vector. Then the variation norm $\|\rho\|_v$ is defined as*

$$\|\rho\|_v = \sum_{i\in S} |\rho_i| \,. \tag{4.12}$$

Doeblin's theorem in states that if, regardless of the initial state, a Markov chain has a positive probability of visiting some fixed state, the chain will eventually reach a stable state. Additionally, it provides a bound on how quickly the chain converges to this stationary distribution. We will apply and prove the theorem for Markov chains that have an immediate, one-time-step, positive probability of visiting a fixed state, as demonstrated in Stroock [25]. To derive Doeblin's theorem, the utilization of the following two lemmas are necessary.

Lemma 4.13. *Let ρ be a row vector, such that $\rho \in \mathbb{R}^S$, and $\|\rho\|_v \leq \infty$. Then,*

$$\sum_{j\in S} (\rho \boldsymbol{P})_j = \sum_{i \in S} (\rho)_i. \tag{4.13}$$

Proof. We have that

$$\begin{aligned}\sum_{j\in S} (\rho \boldsymbol{P})_j &= \sum_{j\in S}\left(\sum_{i\in S} \rho_i \mathbf{P}_{ij}\right) \\ &= \sum_{i\in S}\left(\sum_{j\in S} \rho_i \boldsymbol{P}_{ij}\right) \\ &= \sum_{i\in S}\left(\rho_i \sum_{j\in S} \boldsymbol{P}_{ij}\right) \\ &= \sum_{i\in S} \rho_i, \quad \text{using that } \sum_{j\in S} \boldsymbol{P}_{ij} = 1 \text{ for all } i \in S. \end{aligned} \tag{4.14}$$

□

For the proof in Lemma (4.15) the Kronecker delta function will be used.

Definition 4.14. *The Kronecker Delta function δ_{ij} is defined as*

$$\delta_{ij} = \begin{cases} 1, & \textit{for } i = j \\ 0, & \textit{for } i \neq j \end{cases} \tag{4.15}$$

Lemma 4.15. *Let ρ be a row vector, such that $\rho \in \mathbb{R}^S$, and $\|\rho\|_v \leq \infty$. Let $\boldsymbol{P}$ be a transition probability matrix with the property that, for some state $j_0 \in S$, and $\epsilon > 0$, $\boldsymbol{P}_{ij_0} \geq \epsilon$ for all $i \in S$. Then if $\sum_{i\in S}(\rho)_i = 0$,*

$$\|\rho \boldsymbol{P}^n\|_v \leq (1-\epsilon)^n \|\rho\|_v, \quad \textit{for } n \geq 1. \tag{4.16}$$

Proof. Assume that $\sum_{i\in S} \rho_i = 0$. For the base case, we will show that the inequality holds for $n = 1$. By definition of the variation norm in Equation (4.12),

$$\|\rho \boldsymbol{P}\|_v = \sum_{j\in S} |(\rho \boldsymbol{P})_j|. \tag{4.17}$$

Using the Kronecker delta function in Definition 4.14, we obtain

$$\begin{aligned}
|(\rho \boldsymbol{P})_j| &= \left|\sum_{i\in S} \rho_i \boldsymbol{P}_{ij}\right| \\
&= \left|\sum_{i\in S} \rho_i \boldsymbol{P}_{ij} - \sum_{i\in S} \rho_i \epsilon \delta_{j,j_0}\right|, \quad \text{by the assumption that } \sum_{i\in S} \rho_i = 0. \\
&= \left|\sum_{i\in S} \rho_i (\boldsymbol{P}_{ij} - \epsilon \delta_{j,j_0})\right| \\
&\leq \sum_{i\in S} \left|\rho_i (\boldsymbol{P}_{ij} - \epsilon \delta_{j,j_0})\right| \\
&= \sum_{i\in S} |\rho_i| \left|\boldsymbol{P}_{ij} - \epsilon \delta_{j,j_0}\right| && (4.18) \\
&= \sum_{i\in S} |\rho_i| \left(\boldsymbol{P}_{ij} - \epsilon \delta_{j,j_0}\right), \quad \text{since } \forall i \in S \text{ we have that } \boldsymbol{P}_{ij_0} \geq \epsilon. && (4.19)
\end{aligned}$$

Using the derived bound for $|(\rho \boldsymbol{P})_j|$, in Equation (4.17), Equation (4.16) is obtained for $n = 1$.

$$\begin{aligned}
\|\rho \boldsymbol{P}\|_v &\leq \sum_{j\in S} \left(\sum_{i\in S} |\rho_i| \left(\boldsymbol{P}_{ij} - \epsilon \delta_{j,j_0}\right)\right) \\
&= \sum_{i\in S} |\rho_i| \left(\sum_{j\in S} (\boldsymbol{P}_{ij} - \epsilon \delta_{j,j_0})\right) \\
&= \sum_{i\in S} |\rho_i| \left(\sum_{j\in S} \boldsymbol{P}_{ij} - \sum_{j\in S} \epsilon \delta_{j,j_0})\right) \\
&= (1-\epsilon)\|\rho\|_v && (4.20)
\end{aligned}$$

The last expression follows from the definition of the variation norm 4.12 and using that $\sum_{j\in S} \boldsymbol{P}_{ij} = 1$ for all $i \in S$, by definition of the transition probability matrix $\boldsymbol{P}$ from 4.2.

For the induction hypothesis, suppose that the bound in Equation (4.16) holds for some $n = k \geq 1$ and for all ρ for which $\sum_{i\in S} (\rho)_i = 0$.

$$\|\rho \boldsymbol{P}^k\|_v \leq (1-\epsilon)^k \|\rho\|_v. \tag{4.21}$$

Then for $n = k+1$ we obtain

$$\begin{aligned}
\|\rho \boldsymbol{P}^{k+1}\|_v &= \sum_{j\in S} \left|(\rho \boldsymbol{P}^{k+1})_j\right| \\
&= \sum_{j\in S} \left|((\rho \boldsymbol{P}) \boldsymbol{P}^k)_j\right| \\
&= \sum_{j\in S} \left|(\eta \boldsymbol{P}^k)_j\right|, \quad \text{where } \eta = \rho \boldsymbol{P}. && (4.22)
\end{aligned}$$

Now, note that $\sum_{i\in S} \eta_i = 0$, since

$$\begin{aligned}\sum_{i\in S}\eta_i &= \sum_{i\in S}(\rho\boldsymbol{P})_i \\ &= \sum_{i\in S}\left(\sum_{j\in S}\rho_j\boldsymbol{P}_{ji}\right) \\ &= \sum_{j\in S}\left(\rho_j\sum_{i\in S}\boldsymbol{P}_{ji}\right) \\ &= \sum_{j\in S}\rho_j \quad \text{(Lemma (4.13))} \qquad (4.23) \\ &= 0 \quad \text{by assumption.} \qquad (4.24)\end{aligned}$$

Applying the induction hypothesis (4.21) to Equation (4.22) yields

$$\|\rho\boldsymbol{P}^{k+1}\|_v = \sum_{j\in S}\left|(\eta\boldsymbol{P}^k)_j\right| \le (1-\epsilon)^k\,\|\eta\|_v. \tag{4.25}$$

Using that η is defined as $\eta = \rho\boldsymbol{P}$ together with equation (4.20), the following inequality for $\|\eta\|_v$ can be derived as follows

$$\|\eta\|_v = \|\rho\boldsymbol{P}\|_v \le (1-\epsilon)\,\|\rho\|_v. \tag{4.26}$$

Combining Equations (4.26) and (4.25), Equation (4.16) is obtained for $n = k+1$:

$$\begin{aligned}\|\rho\boldsymbol{P}^{k+1}\|_v &\le (1-\epsilon)^k\,\|\eta\|_v \\ &\le (1-\epsilon)^k\,(1-\epsilon)\,\|\rho\|_v \\ &= (1-\epsilon)^{k+1}\,\|\rho\|_v. \qquad (4.27)\end{aligned}$$

By mathematical induction, if $\sum_{i\in S}\rho_i = 0$, then $\|\rho\boldsymbol{P}^n\|_v \le (1-\epsilon)^n\|\rho\|_v$, for $n \ge 1$. ☐

Theorem 4.16. (Doeblin's theorem) [12] Let $\boldsymbol{P}$ be a transition probability matrix with the property that, for some state $j_0 \in S$, and $\epsilon > 0$, $\boldsymbol{P}_{ij_0} \ge \epsilon$ for all $i \in S$. Then $\boldsymbol{P}$ has a unique stationary probability vector π, $\pi_{j_0} \ge \epsilon$, and, for all initial distributions μ,

$$\|\mu\boldsymbol{P}^n - \pi\|_v \le (1-\epsilon)^n\|\mu-\pi\|_v \le 2(1-\epsilon)^n, \quad n \ge 0. \tag{4.28}$$

Proof. Let $\boldsymbol{P}$ be a transition probability matrix, such that for some state $j_0 \in S$, and $\epsilon > 0$, $(\boldsymbol{P})_{ij_0} \ge \epsilon$ for all $i \in S$. Let ρ be a row vector, such that $\rho \in \mathbb{R}^S$, and $\|\rho\|_v \le \infty$. Then, by Lemma (4.13) and (4.15)

- $\sum_{j\in S}(\rho\boldsymbol{P})_j = \sum_{i\in S}(\rho)_i$
- If $\sum_{i\in S}(\rho)_i = 0$, then $\|\rho\boldsymbol{P}^n\|_v \le (1-\epsilon)^n\|\rho\|_v, \quad$ for $n \ge 1$.

Let μ be a probability vector, thus $\mu \geq 0$ and $\sum_{i \in S} \mu_i = 1$. Denote $\mu_n = \mu \boldsymbol{P}^n$ for all $n \geq 0$. Then for $m \leq n$,

$$\begin{aligned} \|\mu_n - \mu_m\|_v &= \|\mu \boldsymbol{P}^n - \mu \boldsymbol{P}^m\|_v \\ &= \|\mu \mathbf{P}^{n-m} \boldsymbol{P}^m - \mu \boldsymbol{P}^m\|_v \\ &= \|\mu_{n-m} \boldsymbol{P}^m - \mu \boldsymbol{P}^m\|_v \\ &= \|\left(\mu_{n-m} - \mu\right) \boldsymbol{P}^m\|_v \end{aligned} \tag{4.29}$$

Note that since μ_{n-m} and μ are probability vectors, $\sum_{i \in S} \mu_{n-m} = 1$ and $\sum_{i \in S} \mu = 1$. Therefore $\sum_{i \in S} \left(\mu_{n-m} - \mu\right)_i = 0$. Using this result, Lemma 4.15 with $\rho = \mu_{n-m} - \mu$, can be applied to Equation (4.29), to obtain

$$\|\mu_n - \mu_m\|_v \leq (1-\epsilon)^m \|\mu_{n-m} - \mu\|_v. \tag{4.30}$$

By definition of the variation norm 4.12, $\|\mu_{n-m} - \mu\|_v \leq \|\mu_{n-m}\| + \|\mu\|_v = 2$. Therefore Equation (4.30) transforms as follows

$$\|\mu_n - \mu_m\|_v \leq (1-\epsilon)^m \|\mu_{n-m} - \mu\|_v \leq 2(1-\epsilon)^m, \quad \text{for } 1 \leq m < n. \tag{4.31}$$

Thus $\{\mu_n\}_1^\infty$ is Cauchy sequence and therefore convergent [17]. This means that the distance between two consecutive terms becomes arbitrarily small as n increases. Hence there exists a vector π such that the distance between μ_n and π approaches 0 as n increases, thus $\|\mu_n - \pi\| \to 0$.

So the sequence $\{\mu_n\}_1^\infty$ converges to the vector π. Since each μ_n is a probability vector, the limiting vector π must also be a probability vector. Furthermore, the limit is also unique by the uniqueness of limits. This gives the following result

$$\pi = \lim_{n \to \infty} \mu \boldsymbol{P}^{n+1} = \lim_{n \to \infty} (\mu \boldsymbol{P}^n) \boldsymbol{P} = \pi \boldsymbol{P}. \tag{4.32}$$

Equation (4.32) implies that the probability vector π is stationary by Definition 4.7. Using that $\boldsymbol{P}_{ij_0} \geq \epsilon$ for all $i \in S$, Equation (4.32) implies that $(\pi)_{j_0} \geq \epsilon$:

$$(\pi)_{j_0} = \sum_{i \in \mathrm{S}} \pi_i \boldsymbol{P}_{ij_0} \geq \epsilon \sum_{i \in \mathrm{S}} \pi_i = \epsilon. \tag{4.33}$$

Letting μ be an arbitrary probability vector, we can establish the following bound for the distance between the probability distribution after n steps when starting in initial distribution vector μ, and stationary distribution vector π.

$$\|\mu \boldsymbol{P}^n - \pi\|_v = \|(\mu - \pi)\boldsymbol{P}^n\|_v = \|\mu \boldsymbol{P}^n - \pi \boldsymbol{P}^n\|_v \leq (1-\epsilon)^n \|\mu - \pi\|_v \leq 2(1-\epsilon)^n, \text{ for } n \geq 1. \tag{4.34}$$

To derive Equation (4.28), we need to derive the bound for the case $n = 0$, as Equation (4.34) gives the result for $n \geq 1$. For $n = 0$,

$$\|\mu \boldsymbol{P}^n - \pi\|_v = \|\mu \boldsymbol{P}^0 - \pi\|_v = \|\mu \boldsymbol{I} - \pi\|_v = \|\mu - \pi\|_v \leq \|\mu\|_v + \|\pi\|_v = 2. \tag{4.35}$$

Therefore, it has been proven that (4.28) holds for $n \geq 0$. □

Doeblin's theorem states that if there exists a state j_0 such that the probability of transitioning from any state i to this state j_0 is positive and greater or equal than some $\epsilon \geq 0$, then there exists a unique stationary distribution. In addition, the Markov chain converges to that stationary distribution with an exponential speed. As the number of time steps increases, the difference between the chain's distribution and the stationary distribution diminishes exponentially. The rate of convergence is determined by the constant ϵ, with larger values of ϵ indicating faster convergence. Thereby, Doeblin's theorem gives an indication of the convergence speed to its stationary distribution.

4.2 Leslie model as a Markov chain

Doeblin's theorem 4.16, cannot be directly applied to the Leslie model. In order to utilize Doeblin's theorem in the context of the Leslie model for population growth, the Leslie model needs to be formulated as a Markov chain.

Similar to Markov chains, the Leslie model adheres to the Markov property (4.1). As the population distribution at a particular time solely depends on the population distribution in the previous time step. The survival rates, capturing the probabilities of individuals transitioning from one age class to the next, align with the Markov-property inherent in Markov chains. Similarly, the birth rates, determining the offspring of individuals in specific age classes, maintain an independence that resonates with the memory-less characteristic. Indeed, in the Leslie model, the number of offspring produced by an individual in an age class i, does not vary depending on whether an individual in the age class $i - 1$ has had 10 offspring or none. Every individual within a specific age class has an equal expected number of offspring or probability of transitioning to the next age class, regardless of their past reproductive history.

There is a distinction in what the transitions represent in the context of Markov chains versus the Leslie model. In a Markov chain, the transitions typically represent the movement or transitions of an individual or entity from one discrete state to another. Each transition is associated with a probability, and the Markov chain describes how these individual transitions occur over time.

In the Leslie model, the transitions describe how a population as a whole transitions from one age group to another over discrete time intervals. These transitions are driven by age-specific birth and survival rates. The Leslie model does not focus on individual entities but rather on the average behaviour of groups of individuals in different age categories within a population.

Whereas, Markov chain transitions typically relate to individual-level changes between states, Leslie model transitions pertain to group-level changes between age groups based on birth and survival. The Leslie model deals with averages and population dynamics rather than individual-level transitions.

The complexity of formulating the Leslie model as a Markov chain, arises from the fact that the Leslie model involves two distinct processes and thereby contributes in two ways: survival and reproduction. A Markov chain typically models the transitions of individuals or entities from one state to another based on transition probabilities. However, in the Leslie model, the dynamics involve not only the transitions of individuals but also the addition of new individuals through births. These births are not simply transitions of existing individuals but the introduction of entirely new individuals into the population. Incorporating these new entities into a Markov chain is challenging. For example, transition probability t_{12}, refers to the probability that an individual in age class i transitions to age class 2 in a single time step. This transition probability t_{12}, cannot capture *both* the probability that this individual survives to age class 2 and the probability that an individual in age class 1 produces offspring.

The survival process, thus the transitions between age classes of existing individuals can be modelled by a Markov chain. However, using a Markov chain to model the entire Leslie model, thereby also capturing the introduction of offspring, is not directly feasible.

4.2.1 Leslie matrix as Transition matrix

To formulate the Leslie model as a Markov chain, we need to transform the Leslie matrix into a transition probability matrix. By Definition 4.2 this entails that the elements in the matrix represent probabilities and thus must be bounded by 1. Each state must represent a specific age class or population group, and transitions must capture the probability of individuals moving from one age class to another over time while considering both birth and death rates. This again indicates the challenge of incorporating the reproduction process in a Markov chain. As the birth rates, b_i in the Leslie matrix $\boldsymbol{L}$ (2.12), are the average number of offspring produced by individuals in age class i, they do not directly correspond with probabilities of producing offspring.

To formulate the Leslie matrix as a transition probability matrix, only using the survival process, we will use that the Leslie matrix can be written as the addition of the Fertility matrix $\boldsymbol{F}$ and the Survival matrix $\boldsymbol{S}$ (2.10). The Survival matrix $\boldsymbol{S}$ (2.11) can be reinterpreted as a transition matrix, allowing for the application of Doeblin's theorem. For the Leslie matrix we had age classes $i \in \{1, \ldots, m\}$, thus state space $i \in \{1, \ldots, m\}$. For the transition matrix, we will introduce an additional state $m+1$, representing the state of "death". Consequently, the transition matrix will consist of $m+1$ rows and $m+1$ columns, defining the state space $S = \{1, \ldots, m+1\}$.

Adding this extra state 'death' is crucial for satisfying the second property of a stochastic matrix, which states that $\sum_{j \in S} p_{i,j} = 1$ for $i \in S$. In the context of the Leslie model, it ensures that, for each age class, the probabilities of individuals either transitioning to the next age class or die sum up to 1.

The Leslie model incorporates a Leslie matrix denoted as $\boldsymbol{L}$ and an age distribution vector represented by p. Besides a Transition Probability matrix $\boldsymbol{P}$, a Markov chain involves an initial probability vector μ (4.4). This initial probability vector is established by normalizing the age distribution vector p obtained from the Leslie model at time $t = 0$.

4.3 Analysis of the Leslie model via Doeblin's theorem

Doeblin's theorem requires that the transition probability of a Markov chain has some state $j_0 \in S$, and $\epsilon > 0$, $\boldsymbol{P}_{ij_0} \geq \epsilon$ for all $i \in S$. In order to satisfy this property we need the condition that for all $i \in S$, $s_i \neq 1$, so that $1 - s_i \neq 0$, for all $i \in S$.

This condition reflects a realistic scenario, acknowledging that in a population, the chance of survival within an age class is never guaranteed to be 100%. There always exists a small probability of mortality.

4.3.1 An example 4×4 Leslie model

Consider an example Leslie matrix with 4 age classes. Leslie matrix $\boldsymbol{L}$, where $s_i \neq 1$ for all $i \in S$.

$$\boldsymbol{L} := \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \end{bmatrix} \tag{4.36}$$

$$\boldsymbol{L} = \boldsymbol{F} + \boldsymbol{S} \tag{4.37}$$

$$\boldsymbol{F} := \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{S} := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \end{bmatrix} \tag{4.38}$$

For the formulation of the transition probabilities consider for example s_1, the survival rate of age class 1. This signifies the rate of individuals that survive from age class 1 at time t to age class 2 at $t+1$. This corresponds with the probability of transitioning from state 1 to state 2 in the transition probability matrix, and thus $p_{1,2} = s_1$. Conversely, when an individual in age class 1 does not survive to age class 2, it dies and consequently transitions to the death state $m+1 = 5$ with probability $1 - s_1$.

Let $\boldsymbol{T}$ be a transition matrix corresponding to the matrix $\boldsymbol{S}$, where states $i = 1,2,3,4$ are the four age classes and $i = 5$ represents the state of death. Then Survival matrix $\boldsymbol{S}$ (4.38) can be represented as a 5×5 transition matrix $\boldsymbol{T}$ and is defined as

$$\boldsymbol{T} := \begin{bmatrix} t_{11} & t_{21} & t_{31} & t_{41} & t_{51} \\ t_{12} & t_{22} & t_{32} & t_{42} & t_{52} \\ t_{13} & t_{23} & t_{33} & t_{43} & t_{53} \\ t_{14} & t_{24} & t_{34} & t_{44} & t_{54} \\ t_{15} & t_{25} & t_{35} & t_{45} & t_{55} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ s_1 & 0 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 & 0 \\ 0 & 0 & s_3 & 0 & 0 \\ 1-s_1 & 1-s_2 & 1-s_3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0 \\ 0.5 & 0.7 & 0.8 & 1 & 1 \end{bmatrix},$$

where

$$t_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i). \tag{4.39}$$

The transition matrix $\boldsymbol{T} = (t_{i,j})$ is a column stochastic matrix satisfying

- $0 \leq t_{i,j}$ for $i, j \in S = \{1,2,3,4,5\}$, and
- $\sum_{j \in S} t_{i,j} = 1$ for $i \in S$.

For purpose of notation in the context of Markov chains, we denote the transpose of matrix $\boldsymbol{T}$ as $\boldsymbol{P}$, ensuring that matrix $\boldsymbol{P}$ is row stochastic.

$$\boldsymbol{P} = \boldsymbol{T}^T = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} & t_{15} \\ t_{21} & t_{22} & t_{23} & t_{24} & t_{25} \\ t_{31} & t_{32} & t_{33} & t_{34} & t_{35} \\ t_{41} & t_{42} & t_{43} & t_{44} & t_{45} \\ t_{51} & t_{52} & t_{53} & t_{54} & t_{55} \end{bmatrix} = \begin{bmatrix} 0 & s_1 & 0 & 0 & 1-s_1 \\ 0 & 0 & s_2 & 0 & 1-s_2 \\ 0 & 0 & 0 & s_3 & 1-s_3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0.3 & 0 & 0.7 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.40}$$

Computations for matrix $\boldsymbol{P}$ give that

$$\boldsymbol{P}^2 = \begin{bmatrix} 0 & 0 & 0.15 & 0 & 0.85 \\ 0 & 0 & 0 & 0.06 & 0.94 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{P}^3 = \begin{bmatrix} 0 & 0 & 0 & 0.03 & 0.97 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{P}^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.41}$$

Equation (4.41) shows that the population transitions into state 5, the death state. The computations indicate that, after four time steps, the probability of being in state 5 reaches 1. Thus, the population converges to the population distribution vector $\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

If the age distribution vector at time $t = 0$ (2.3) is given by

$$P(0) := \begin{bmatrix} P_1(0) \\ P_2(0) \\ P_3(0) \\ P_4(0) \\ P_5(0) \end{bmatrix} = \begin{bmatrix} 100 \\ 200 \\ 200 \\ 300 \\ 0 \end{bmatrix},$$

where $P_5(0) = 0$ refers to the fact that there are no dead people at time $t = 0$. Then by normalizing $P(0)$, the initial distribution vector μ can be defined as $\mu = \begin{bmatrix} \frac{1}{8} & \frac{2}{8} & \frac{2}{8} & \frac{3}{8} & 0 \end{bmatrix}$.

The probability vector $\pi = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ serves as a stationary distribution, denoted by π. According to Definition (4.7), it should hold that $\pi = \pi\boldsymbol{P}$. Applying this to the transition probability matrix $\boldsymbol{P}$ in Equation (4.40), it can be verified that the vector $\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ indeed is a stationary distribution.

$$\pi\boldsymbol{P} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0.3 & 0 & 0.7 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \pi \tag{4.42}$$

Using Doeblin's theorem, the uniqueness of the stationary probability vector π is given. We will apply Doeblin's theorem 4.16 to the transition probability matrix $\boldsymbol{P}$ specified in Equation (4.40). Note that the requirement that for some state $j_0 \in S$ and $\epsilon \geq 0$, $\boldsymbol{P}_{ij_0} \geq \epsilon$ for all $i \in S$ is met, since this condition is fulfilled by state $j_0 = 5$. Irrespective of the age class an individual belongs to (states 1,2, 3, 4, or 5), there is always a positive probability of transitioning to state 5. There exists a positive probability of death for individuals in any age class.

Let ϵ be the minimum of p_{i5} for $i \in S$. So, for the transition probability matrix $\boldsymbol{P}$ in Equation (4.40), $\epsilon = 0.5$. Then indeed $\boldsymbol{P}_{i5} \geq 0.5$ for all $i \in S$. Hence, Doeblin's theorem implies that the transition probability matrix $\boldsymbol{P}$, has a unique stationary probability distribution vector π, such that $\pi_{j_0} = \pi_5 \geq 0.5$. Additionally, for all initial distributions $\|\mu\boldsymbol{P}^n - \pi\|_v \leq (1-\epsilon)^n \|\mu - \pi\|_v \leq 2(1-\epsilon)^n$ for $n \geq 0$.

Note that for the previously determined stationary distribution (4.42), it holds that $\pi_5 = 1 \geq 0.5$. By Doeblin's theorem we can assert the uniqueness of this stationary distribution $\pi = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ and that with $\epsilon = 0.5$ we obtain the following

$$\|\mu\boldsymbol{P}^n - \pi\|_v \leq 2(1-\epsilon)^n = 2(1-0.5)^n = 2(0.5)^n \text{ for } n \geq 0.$$

As n increases, $(0.5)^n$ decreases exponentially, resulting in

$$\|\mu\boldsymbol{P}^n - \pi\|_v \to 0, \quad \text{which implies that } \mu\boldsymbol{P}^n \cong \pi.$$

Note that computations of the transition probability matrix $\boldsymbol{P}$ in Equation (4.41) give a more specific result than Doeblin's theorem as it indicates that the population converges to the stationary probability vector $\pi = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ at $n = 4$.

4.3.2 The generalized Leslie model

In the subsequent analysis, Doeblin's theorem will be applied to a generalized Leslie model. Again $\boldsymbol{L} = \boldsymbol{F} + \boldsymbol{S}$, where the Survival matrix $\boldsymbol{S}$ will be used to formulate a Markov chain. Consider the general

Leslie matrix $\boldsymbol{L}$ (2.12) consisting of m states in the Leslie matrix. Adding state $m+1$ to representing death, the Survival matrix $\mathbf{S}$ can be formulated as a $(m+1) \times (m+1)$ transition probability matrix $\boldsymbol{T}$.

$$\boldsymbol{T} = \begin{bmatrix} t_{11} & t_{21} & t_{31} & \cdots & t_{m-1,1} & t_{m,1} & t_{m+1,1} \\ t_{12} & t_{22} & t_{32} & \cdots & t_{m-1,2} & t_{m,2} & t_{m+1,2} \\ t_{13} & t_{23} & t_{33} & \cdots & t_{m-1,3} & t_{m,3} & t_{m+1,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ t_{1m} & t_{2m} & t_{3m} & \cdots & t_{m-1,m} & t_{m,m} & t_{m+1,m} \\ t_{1,m+1} & t_{2,m+1} & t_{3,m+1} & \cdots & t_{m-1,m+1} & t_{m,m+1} & t_{m+1,m+1} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ s_1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & 0 & 0 \\ 1-s_1 & 1-s_2 & 1-s_3 & \cdots & 1-s_{m-1} & 1 & 1 \end{bmatrix} \tag{4.43}$$

Again, for the purpose of notation in the context of Markov chains, we denote the transpose of matrix $\boldsymbol{T}$ as $\boldsymbol{P}$, ensuring that matrix $\boldsymbol{P}$ is row stochastic.

$$\boldsymbol{P} = \boldsymbol{T}^T = \begin{bmatrix} 0 & s_1 & 0 & 0 & \cdots & 0 & 1-s_1 \\ 0 & 0 & s_2 & 0 & \cdots & 0 & 1-s_2 \\ 0 & 0 & 0 & s_3 & \cdots & 0 & 1-s_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & s_{m-1} & 1-s_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.44}$$

In general $p_{i,i+1} = s_i$ and $p_{i,m+1} = 1 - s_i$ for $i = 1, \ldots m$. For the last row, state $m+1$, element $p_{m+1,m+1} = 1$ and all other elements are zero. Since if an individual is deceased, thus in state $m+1$, it will remain in this state with probability 1.

Appendix A includes some computations for $\boldsymbol{P}^k$, where $k \geq 2$. From these computations, an expression for $\boldsymbol{P}^k$ with a general k can be formulated. We will prove this expression with the use of mathematical induction.

Lemma 4.17. *Let $\boldsymbol{P}$ be a transition probability matrix defined as in Equation (4.44). Then for $1 \leq n \leq m$, the n-step transition probability matrix, $\boldsymbol{P}^n = (p_{ij}^{(n)})$ is defined as*

$$\boldsymbol{P}^n := \begin{bmatrix} 0 & \cdots & 0 & p^{(n)}_{1,n+1} & 0 & 0 & \cdots & 0 & 1-p^{(n)}_{1,n+1} \\ 0 & \cdots & 0 & 0 & p^{(n)}_{2,n+2} & 0 & \cdots & 0 & 1-p^{(n)}_{2,n+2} \\ 0 & \cdots & 0 & 0 & 0 & p^{(n)}_{3,n+3} & \cdots & 0 & 1-p^{(n)}_{3,n+3} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & p^{(n)}_{m-n,m} & 1-p^{(n)}_{m-n,m} \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{4.45}$$

For $1 \leq n \leq m-1$ *it holds that* $p^{(n)}_{i,n+i} = s_i s_{i+1} \cdots s_{i+n-1}$ *and* $\boldsymbol{P}^n$ *can be expressed as*

$$\boldsymbol{P}^n := \begin{bmatrix} 0 & \cdots & 0 & s_1 \cdots s_n & 0 & 0 & \cdots & 0 & 1-s_1 \cdots s_n \\ 0 & \cdots & 0 & 0 & s_2 \cdots s_{n+1} & 0 & \cdots & 0 & 1-s_2 \cdots s_{n+1} \\ 0 & \cdots & 0 & 0 & 0 & s_3 \cdots s_{n+2} & \cdots & 0 & 1-s_3 \cdots s_{n+2} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & s_{m-n} \cdots s_{m-1} & 1-s_{m-n} \cdots s_{m-1} \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{4.46}$$

Moreover, for $n = m$*,* $\boldsymbol{P}^n$ *satisfies*

$$\boldsymbol{P}^m = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{4.47}$$

Proof. Using that $p_{m-n,m} = p_{m-1,m}$ for $n = 1$, Equation (4.45) and (4.46) yield the following result,

$$\boldsymbol{P} = \begin{bmatrix} 0 & p^{(1)}_{12} & 0 & 0 & \cdots & 0 & 1-p^{(1)}_{12} \\ 0 & 0 & p^{(1)}_{23} & 0 & \cdots & 0 & 1-p^{(1)}_{23} \\ 0 & 0 & 0 & p^{(1)}_{34} & \cdots & 0 & 1-p_{34}(1) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & p^{(1)}_{m-1,m} & 1-p^{(1)}_{m-1,m} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & s_1 & 0 & 0 & \cdots & 0 & 1-s_1 \\ 0 & 0 & s_2 & 0 & \cdots & 0 & 1-s_2 \\ 0 & 0 & 0 & s_3 & \cdots & 0 & 1-s_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & s_{m-1} & 1-s_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Since this matrix corresponds with the expression in Equation (4.44), the base case holds.

For the induction hypothesis, suppose that the expression holds for some $n = k$ with $1 \leq k < m-1$. Then

$$\boldsymbol{P}^{k+1} = \boldsymbol{P}^k \boldsymbol{P} =$$

$$\begin{bmatrix}
0 & \cdots & 0 & p^{(k)}_{1,k+1} & 0 & 0 & \cdots & 0 & 1-p^{(k)}_{1,k+1} \\
0 & \cdots & 0 & 0 & p^{(k)}_{2,k+2} & 0 & \cdots & 0 & 1-p^{(k)}_{2,k+2} \\
0 & \cdots & 0 & 0 & 0 & p^{(k)}_{3,k+3} & \cdots & 0 & 1-p^{(k)}_{3,k+3} \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & p^{(k)}_{m-k,m} & 1-p^{(k)}_{m-k,m} \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
\vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}
\begin{bmatrix}
0 & s_1 & 0 & 0 & \cdots & 0 & 1-s_1 \\
0 & 0 & s_2 & 0 & \cdots & 0 & 1-s_2 \\
0 & 0 & 0 & s_3 & \cdots & 0 & 1-s_3 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & s_{m-1} & 1-s_{m-1} \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}$$

$$= \begin{bmatrix}
0 & \cdots & 0 & p^{(k)}_{1,k+1}s_{k+1} & 0 & 0 & \cdots & 0 & 1-p^{(k)}_{1,k+1}s_{k+1} \\
0 & \cdots & 0 & 0 & p^{(k)}_{2,k+2}s_{k+2} & 0 & \cdots & 0 & 1-p^{(k)}_{2,k+2}s_{k+2} \\
0 & \cdots & 0 & 0 & 0 & p^{(k)}_{3,k+3}s_{k+3} & \cdots & 0 & 1-p^{(k)}_{3,k+3}s_{k+3} \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & p^{(k)}_{m-k,m}s_m & 1-p^{(k)}_{m-k,m}s_m \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
\vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{bmatrix} \tag{4.48}$$

Note in the above equation that since $s_m = 0$, we obtain that $p^{(k)}_{m-k,m}s_m = 0$ and thus equivalently

$1 - p^{(k)}_{m-k,m} s_m = 1$. As a result, the matrix in Equation (4.48) can be computed as

$$= \begin{bmatrix} 0 & \cdots & 0 & p^{(k)}_{1,k+2} & 0 & 0 & \cdots & 0 & 1 - p^{(k)}_{1,k+2} \\ 0 & \cdots & 0 & 0 & p^{(k)}_{2,k+3} & 0 & \cdots & 0 & 1 - p^{(k)}_{2,k+3} \\ 0 & \cdots & 0 & 0 & 0 & p^{(k)}_{3,k+4} & \cdots & 0 & 1 - p^{(k)}_{3,k+4} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & p^{(k)}_{m-k-1,m} & 1 - p^{(k)}_{m-k-1,m} \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.49}$$

Equation (4.49) gives the expression for $\boldsymbol{P}^n$ with $n = k+1$ in (4.45). Furthermore, for $1 \leq k < m-1$,

$$p^{(k+1)}_{i,k+1+i} = p^{(k)}_{i,k+i} p^{(1)}_{i,k+1} = p^{(k)}_{i,k+i} s_{i+k} = s_i s_{i+1} \cdots s_{i+k-1}, \tag{4.50}$$

which aligns with the expression $p^{(n)}_{i,n+i} = s_i s_{i+1} \cdots s_{i+n-1}$ for $1 \leq n \leq m-1$, with $n = k+1$.

By mathematical induction, the expression in Equation (4.45) and (4.46) hold for $1 \leq n \leq m-1$.

For $n = m-1$ we find that

$$\boldsymbol{P}^{m-1} = \begin{bmatrix} 0 & \cdots & p_{1,m} & 1 - p_{1,m} \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.51}$$

Then it follows that $\boldsymbol{P}^m = \boldsymbol{P}\boldsymbol{P}^{m-1}$, resulting in the following computation

$$\boldsymbol{P}^m = \boldsymbol{P}\boldsymbol{P}^{m-1} = \begin{bmatrix} 0 & s_1 & 0 & \cdots & 0 & 1 - s_1 \\ 0 & 0 & s_2 & \cdots & 0 & 1 - s_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & 1 - s_{m-1} \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \cdots & p_{1,m} & 1 - p_{1,m} \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.52}$$

$$= \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.53}$$

For $n \geq m$ we have that $\boldsymbol{P}^n = \boldsymbol{P}^{n-m}\boldsymbol{P}^m$.

$$\boldsymbol{P}^m\boldsymbol{P} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & s_1 & 0 & \cdots & 0 & 1-s_1 \\ 0 & 0 & s_2 & \cdots & 0 & 1-s_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{m-1} & 1-s_{m-1} \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{4.54}$$

Hence, for all $n \geq m$, it becomes evident that the transition probability matrix remains stable, and the entire population has transitioned to state $m+1$, signifying mortality, with a probability of 1. ☐

The expression for the n-step transition probability matrix holds a logical interpretation. It represents the probability of transitioning from age class i to age class $i+n$ in n-time steps. To achieve this transition, an individual must have survived through all intermediate age classes between i and $i+n$ represented by the survival rates s_i.

Now that we have formulated an expression for $\boldsymbol{P}^n$, we can apply Doeblin's theorem, similar to the 4×4 Leslie model. The stationary distribution vector is denoted as $\pi = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$, with j_0 representing the death state, corresponding to state $m+1$. Doeblin's theorem implies that $\|\mu\boldsymbol{P}^n - \pi\|_v \leq 2(1-\epsilon)^n$ for $n \geq 0$, and as $n \to \infty$, $\mu\boldsymbol{P}^n \cong \pi$.

Consider a transition probability matrix $\boldsymbol{P}$ defined as in (4.44). For state $j_0 = m+1$ and $\epsilon = \max\{p_{i,m+1} : i \in S\}$, it is guaranteed that $\boldsymbol{P}_{i,m+1} \geq \epsilon$ for all $i \in S$. Consequently, by Doeblin's theorem, $\boldsymbol{P}$ has a unique stationary probability vector π, with $\pi_{m+1} \geq \epsilon$. For all initial distributions μ, the inequality $\|\mu\boldsymbol{P}^n - \pi\|_v \leq 2(1-\epsilon)^n$ holds for $n \leq 0$, and thus, as $n \to \infty$, $\mu\boldsymbol{P}^n \cong \pi$.

While Doeblin's theorem offers insights into the long-term behaviour as $n \to \infty$, we have established a more specific result for this case. Specifically, it is not only that the distribution converges to the stable distribution π as n approaches infinity, but it does so in no more than m steps. This outcome aligns with expectations: with $m+1$ states, individuals in age class 1 have transitioned to state $m+1$, and the same progression applies to all older individuals. This model studies a population that ages without any new reproduction. As a result, extinction occurs after a time span equivalent to the length of the age classes has passed. The following proposition proves this using induction.

Proposition 4.18. *Let $\boldsymbol{P}$ be a transition probability matrix defined as in (4.44) on state space $S = \{1, \cdots, m+1\}$. Then for all initial distribution vectors μ and unique stationary distribution vector $\pi = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$, it holds that $\mu\boldsymbol{P}^n = \pi$, for $n \geq m$.*

Proof. Let the initial distribution vector μ be denoted as $\mu = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_{m+1} \end{bmatrix}$.
For the base case $n = m$,

$$\mu\boldsymbol{P}^m = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_{m+1} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = \pi \tag{4.55}$$

Using that μ is a probability vector and therefore $\sum_{i \in S} \mu_i = 1$. The computation in Equation (4.55) yields the unique distribution vector π. Hence, the base case holds.

For the Induction Hypothesis, assume that the proposition holds for some $k \geq m$. Then,

$$\mu \boldsymbol{P}^{k+1} = \mu \boldsymbol{P}^{k} \boldsymbol{P} = \pi \boldsymbol{P} = \pi \tag{4.56}$$

By mathematical induction, the proposition holds for all $n \geq m$. $\square$

In this particular scenario, the application of Doeblin's theorem provides insights into a situation where the transition matrix is solely characterized by the survival rates of a Leslie model. This specific case offers a straightforward illustration, and the outcomes align with our prior expectations. As the population ages without any new births, it naturally converges to the population distribution vector $\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$, signifying the ultimate extinction of the entire population. Therefore, while this case may not reveal unexpected dynamics, it serves as a valuable example of how Doeblin's theorem can validate anticipated outcomes in the context of the Leslie model as a Markov chain.

Chapter 5

The Leslie model as a Branching process

Formulating the entire Leslie model as a Markov chain proved challenging due to the inherent complexity of including the reproductive process. As it is limiting to focus solely on the Survival matrix $\boldsymbol{S}$ within the Leslie model, this chapter introduces an alternative approach to gain insights into the asymptotic behaviour of the complete Leslie model. Representing part of the Leslie model as a branching process not only complements the insights derived from Perron-Frobenius theorem but also has the potential to provide a more realistic representation of population dynamics, as it accounts for demographic stochasticity.

5.1 Branching process

Branching processes were invented by biologist Francis Galton and mathematician Henry Watson, who used them to study the extinction of family names. Lotka was the first to apply branching processes to demographic data, calculating the extinction probabilities for family names in the United States. The idea of connecting branching processes to matrix population models was first explored by Pollard [5].

The theory of branching processes is an area of mathematics that describes situations in which an entity exists for a time period and then may be replaced by zero, one, two or more entities of a similar or different type [16]. The Galton-Watson process is the oldest and simplest branching process, in which each individual produces offspring and then dies after one time step. In a branching process, each member of the nth generation produces a number of offspring that is independent of past generations and independent of the number of offspring produced by other members of the same nth generation. As branching processes are examples of Markov chains, this independence aspect of the branching process relates to the Markov property (4.1).

Definition 5.1. ***(The Galton-Watson Branching process)*** *[12] Starting with a single individual at time $n = 0$, every individual in the branching process lives exactly one unit of time, then produces offspring and dies. The number of offspring produced by an individual is determined by a discrete random variable Y following a probability distribution. Thus, the family-sizes of individuals are random variables Y. These are independent random variables are identically distributed.*

Denote Z_n for the population size at time n, and label the individuals at time n as $1, 2, 3, \ldots, Z_n$. Then $Y_1, \ldots, Y_{Z_n}$ denote the family sizes of the individuals. The number of offspring at time $n+1$, Z_{n+1}, is equal to the total number *of offspring of the individuals* $1, 2, 3, \ldots, Z_n$. *Thus,*

$$Z_{n+1} = \sum_{i=1}^{Z_n} Y_i \tag{5.1}$$

Then the branching process is $\{Z_0, Z_1, Z_2, \ldots\} = \{Z_n : n \in \mathbb{N} \cup 0\}$

5.1.1 Demographic Stochasticity

In the Leslie model, where vital rates like survival and birth rates are assumed to be constant over time, the projection of future populations is deterministic. This means that given a Leslie matrix with a survival probability of, for instance $s_{ij} = 0.5$, exactly 50% of individuals in age class i are expected to transition to age class j. This transition is deterministic, as it is precisely defined by the Leslie matrix [5].

Demographic stochasticity is the term used to describe the inherent randomness in demographic processes, like births and deaths, when applied to individuals within a population.When considering demographic stochasticity, the projections become probabilistic. Branching processes incorporate demographic stochasticity by modelling demographic events as random variables instead of deterministic values. Instead of a fixed 50% transition rate for each individual, the transition is treated as a random variable. In that case, each individual in age class i has an independent 50% chance of transitioning to age class j. The introduction of randomness, achieved by applying a probability to each individual, leads to variability in the number of individuals transitioning from age class i to j. In cases where birth rates are non-zero, even when survival rates are set to zero, and the initial population is non-zero, according to the Leslie model, there is no possibility of the population extinction in the immediate time step. Considering demographic stochasticity in Branching processes, there is always a small chance, however minimal, that the population might face extinction in the subsequent time step.

Though demographic stochasticity introduces randomness into population dynamics, its impact depends on the size of the population. In large populations, the effects of demographic stochasticity average out, by the law of large numbers. As a result, the overall population dynamics appear more deterministic. The random fluctuations in transition numbers for individual members cancel each other out, and the population's behaviour aligns closely with the Leslie model projections. Conversely, in small populations, demographic stochasticity plays a more prominent role. With fewer individuals, the random variability, introduced by probability distributions of the vital rates, has a more significant impact on the population's dynamics. This can lead to greater unpredictability and deviation from Leslie model projection. Therefore, for small populations, a model that takes demographic stochasticity into account can provide a more accurate or at least realistic representation of population dynamics compared to a population projection determined by the Leslie model. Branching processes allow for an analysis of the effects of demographic stochasticity while considering the probabilistic nature of reproduction and survival.

5.2 Leslie model as a Branching process

The Galton-Watson branching process is often used to describe the population growth of certain organisms, such as bacteria. However, when dealing with populations of organisms that age and progress through distinct life stages, like humans or many animals, the branching process model needs to be adapted to account for this. So we need a more complex branching process to incorporate the process of aging. So a branching process where individuals progress through age classes, reproduce based on age-specific fertility rates, and die based on age-specific survival rates, all within the same time step. This involves creating a branching process for each individual within the initial population.

5.2.1 Reproduction: birth rates to birth probabilities

The Fertility matrix $\boldsymbol{F}$ in the Leslie model contains expected births per individual, not probabilities. Converting these expectations into probability distributions necessitates specifying the appropriate distribution parameters. Since different age classes have distinct birth rates, a unique probability distribution is needed for each age class.

At each time step, individuals in each age class contribute offspring according to their age-specific reproductive rates b_i. To incorporate the stochastic nature of reproduction, we need to select a suitable probability distribution. The elements in the Fertility matrix F are not probabilities but expected births per individual. Turning these expectations into probability distributions requires extra information. This information could come from data, detailing the likelihood of producing $0, 1, 2, \ldots$ offspring. When such data is not available, reproduction will be modelled by specifying a distribution appropriate to the life history of the species [5].

Using the expected births b_i, we can determine the parameter for this specified probability distribution. Assume that the number of offspring produced by an individual in age class i follows a random distribution with a mean of b_i. This aligns with the definition of the birth rate in the Leslie model, representing the average number of offspring produced by individuals in age class i. As each age class has different birth rates, we will have different probability distributions for each age class.

For example, denote the random variable representing the number of offspring produced by an individual in age class i as Y_i. As indicated by the Leslie matrix, the expected or average number of offspring from an individual in age class i is b_i. An appropriate choice of probability distribution for this scenario is the Poisson distribution, characterized by the parameter λ [3]:

$$\mathbb{P}[Y = x] = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \ldots \tag{5.2}$$

The mean of a Poisson distribution equals λ, and thus for this particular distribution we have that $\lambda = b_i$. As a result, this leads to the formulation of the probability mass function for the Poisson distribution, with b_i as its parameter:

$$\mathbb{P}[X = x] = \frac{e^{-b_i}b_i^x}{x!} \quad x = 0, 1, 2, \ldots \tag{5.3}$$

5.2.2 Aging: Incorporating age-specific survival rates.

At each time step, individuals move from one age class to another based on their age-specific survival probabilities s_i in the Leslie model. To introduce demographic stochasticity into the survival rates, the survival probabilities can be modelled using a Bernoulli distribution with probability s_i. In this representation, *each* individual either survives to the next age class or does not. This way, the survival probabilities s_i can be applied to each individual in each age class to determine which individuals survive to the next generation as they move to the next age class. Repeating this process for each individual within the population will simulate the progression of the entire population over time. For simplification, this iterative process, can be replaced by introducing the Binomial distribution with parameter s_{i-1} and the population size of age class $i - 1$ in the preceding time step.

5.2.3 Simulation of the Branching process

At each time step, the process begins with individuals in each age class represented in a population vector. For each individual within an age class, the number of offspring they will produce is determined independently using the Poisson distribution. This Poisson distribution has a mean offspring defined by the birth rate for that age class. This process of determining offspring is repeated for each individual within every age class, considering their specific birth rates. Then, the offspring from all individuals within each age class are summed. This total represents the offspring produced in the current time step for the entire population.

Next, for each age class, the number of individuals that will survive to the next time step is determined using a Binomial distribution. This process is repeated for each age class, determining how many individuals from each age class will survive and move to the next age class. The result is an age vector representing the population distribution in the next time step.

The above process is repeated for each time step in the simulation. To project the population at a specific time, as $n = 10$, the process starts at $n = 0$ and iteratively performs these reproduction and aging steps for each time step until reaching $n = 10$. Appendix C includes a simulation of the $\times 4$ Leslie model formulated as a branching process in Python. This code calculates and displays the population growth and distribution over time using the branching process approach. In addition, Appendix B includes Python code, that provides a simulation using the Leslie model approach to allow for comparison. In the codes for both Appendix B and C, the birth rates, death rates, initial populations and desired time steps, can be modified.

5.3 Analysis of the Leslie model in comparison to Branching process

5.3.1 An example 4×4 Leslie model

We will consider the 4×4 example Leslie model as in the previous chapters with the Leslie matrix in Equation (3.24) and initial age distribution vector at time $t = 0$ in Equation (3.25).

$$\mathbf{L} := \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \end{bmatrix}, \quad P(0) := \begin{bmatrix} P_1(0) \\ P_2(0) \\ P_3(0) \\ P_4(0) \end{bmatrix} = \begin{bmatrix} 100 \\ 200 \\ 200 \\ 300 \end{bmatrix} \tag{5.4}$$

Simulation depicted in Figure 5.1 and 5.2 illustrate the exponential growth of the Leslie model, with the total population demonstrating a growth rate of $1,744$ as it gradually approaches stability. This outcome aligns with the findings from the Perron-Frobenius theorem.

The computations show that from time step 6 and onwards, the population growth stabilizes and has a consistent growth rate for each age class of 1.744. While from the Perron-Frobenius theorem we determined that the population growth would eventually stabilize with a growth rate of 1.744, it did not offer insights into the number of time steps required for this convergence to occur. Additionally, as anticipated by the Perron-Frobenius theorem, we observe that regardless of any adjustments made to the initial population, the population ultimately converges to the same stable distribution

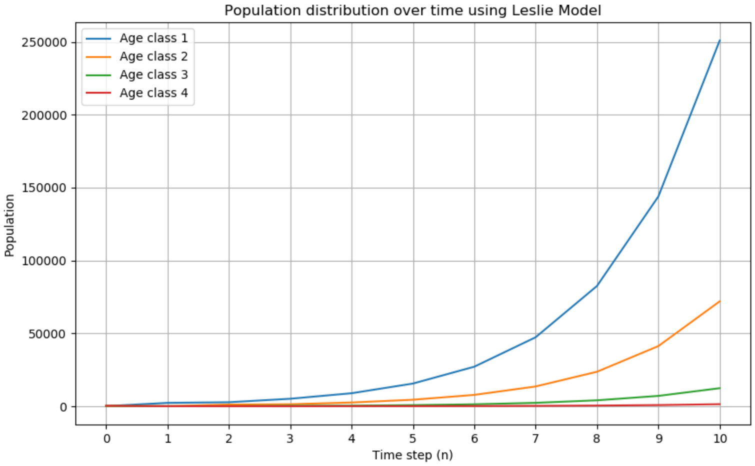

Figure 5.1: Simulation of population distribution over time for example 4×4 Leslie model (5.4), initial population $P(0) = \begin{bmatrix} 100 & 200 & 200 & 300 \end{bmatrix}^T$.

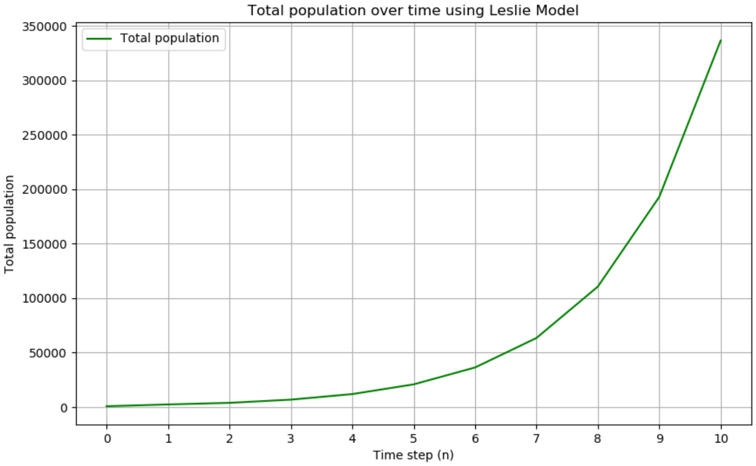

Figure 5.2: Simulation of total population over time for example 4 × 4 Leslie model (5.4), initial population $P(0) = \begin{bmatrix} 100 & 200 & 200 & 300 \end{bmatrix}^T$.

To compare the Leslie model with the Branching process, we need to take into account that the simulations within the Branching process vary each run, as they depend on probability distributions to determine outcomes. Consequently, each simulation run presents a unique trajectory, sometimes closely resembling the deterministic Leslie model and at other times deviating from it.

Figure 5.3 and 5.4 illustrate an example simulation for both the Leslie model and Branching process, with a very small initial population $P(0) = \begin{bmatrix} 1 & 2 & 2 & 3 \end{bmatrix}^T$. The simulation using the Branching process deviates from the exponential growth in the Leslie model. Unlike the constant growth factor in the Leslie model, it shows how the population also decreases at time step $n = 9$. Additionally, the calculated growth rate factors exhibit fluctuations at each time step, reflecting the dynamic nature of the Branching process and the influence of demographic stochasticity on population dynamics. This shows how demographic stochasticity influences the projection of a small population.

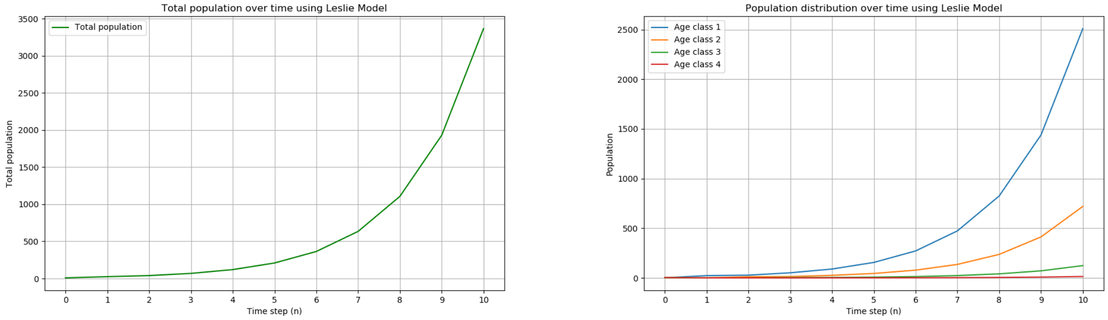

Figure 5.3: Example 4 × 4 Leslie model: total population and population distribution over time with initial population vector $P(0) = \begin{bmatrix} 1 & 2 & 2 & 3 \end{bmatrix}^T$.

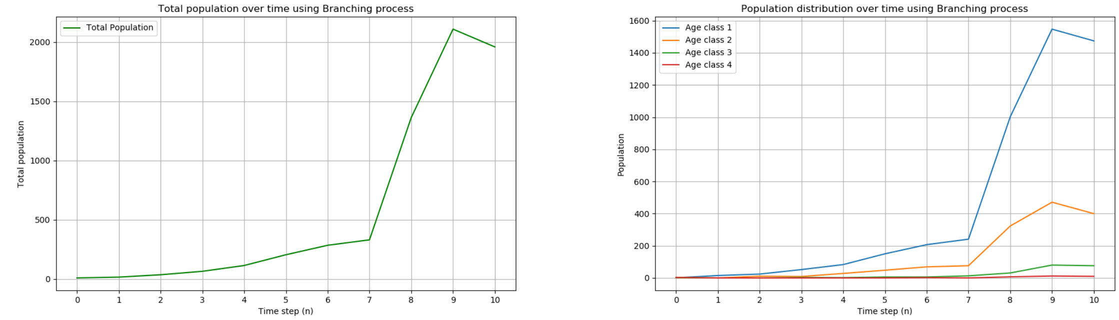

Figure 5.4: A simulation of example 4×4 Leslie model as Branching process: total population and population distribution over time with initial population vector $P(0) = \begin{bmatrix} 1 & 2 & 2 & 3 \end{bmatrix}^T$.

Figure 5.5 and 5.6 illustrate an example simulation for both the Leslie model and Branching process, with a very small initial population $P(0) = \begin{bmatrix} 1000 & 2000 & 2000 & 3000 \end{bmatrix}^T$. Comparing both simulations indicates that the influence of demographic stochasticity diminishes as the population size increases. The total population at $n = 10$ using the Leslie model results in 3366834.3 and the total population at $n = 10$ by this Branching process sample simulation equals 3399919.

During this sample run, computations reveal that starting from time step 5, the growth rates among different age classes exhibit fluctuations within the range of approximately 1.72 to 1.77. These values closely align with the expected growth rate of 1.744 derived from the Leslie model. This observation shows that the sensitivity of population dynamics to demographic stochasticity decreases as population size increases. It indicates the convergence of the Branching process toward the behaviour predicted by the deterministic Leslie model as the population size increases.

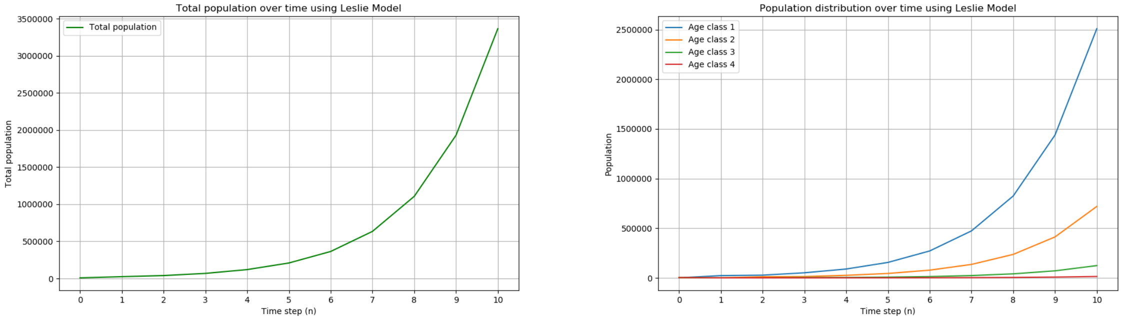

Figure 5.5: Example 4×4 Leslie model: total population and population distribution over time with initial population vector $P(0) = \begin{bmatrix} 1000 & 2000 & 2000 & 3000 \end{bmatrix}^T$

.

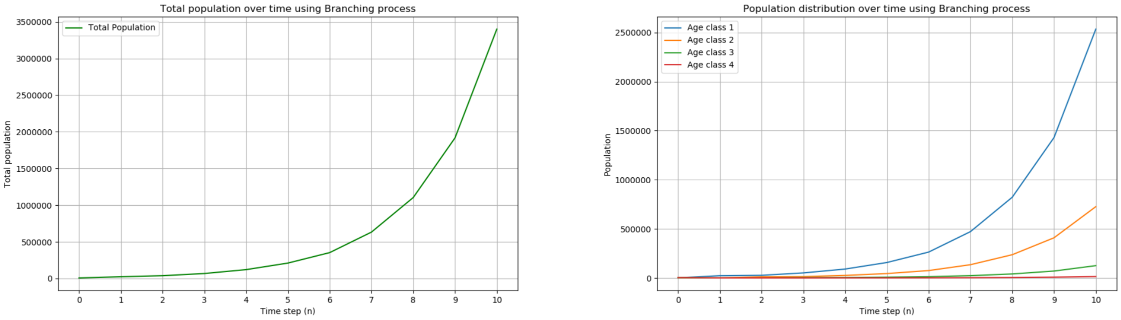

Figure 5.6: Example 4 × 4 Leslie model as Branching process: total population and population distribution over time with initial population vector $P(0) = \begin{bmatrix} 1000 & 2000 & 2000 & 3000 \end{bmatrix}^T$

Table 5.1 demonstrates the percentage range between simulations of the Branching process compared to those of the Leslie model for different population sizes. As population sizes increase, it becomes clear that the Leslie model closely approximates the outcomes predicted by the Branching process, despite the Leslie model's lack of consideration for demographic stochasticity. It highlights the power of the Leslie model, particularly when applied to larger populations.

Additionally, this observation shows the effectiveness of the Perron-Frobenius theorem. It highlights how a relatively straightforward computation of an eigenvector can yield highly accurate growth factor predictions, avoiding the need for extensive calculations within the Branching process framework. Particularly for larger populations, where the computational demands of the Branching process can be substantial, this theorem offers an efficient and reliable alternative for estimating population growth rates.

Initial population vector	**Deviation total population at $n = 10$**
$P(0) = \begin{bmatrix} 1 & 2 & 2 & 3 \end{bmatrix}^T$	$5 - 35\%$
$P(0) = \begin{bmatrix} 10 & 20 & 20 & 30 \end{bmatrix}^T$	$5 - 15\%$
$P(0) = \begin{bmatrix} 100 & 200 & 200 & 300 \end{bmatrix}^T$	$0 - 5\%$
$P(0) = \begin{bmatrix} 1000 & 2000 & 2000 & 3000 \end{bmatrix}^T$	$0 - 2\%$
$P(0) = \begin{bmatrix} 10000 & 20000 & 20000 & 30000 \end{bmatrix}^T$	$0 - 0.5\%$
$P(0) = \begin{bmatrix} 100000 & 200000 & 200000 & 300000 \end{bmatrix}^T$	$0 - 0.2\%$

Table 5.1: Deviation of total population at $n = 10$ Branching process simulation compared to original Leslie model in percentage, for initial population distribution vectors

In this chapter, we explored both a deterministic Leslie model and a stochastic Branching process to simulate population growth under diverse initial conditions. While the Leslie model provides a constant and predictable trajectory of population growth, the Branching process introduces stochasticity. In the Branching process, birth and survival rates are determined by random Poisson or binomial distributions in each iteration, leading to inherent variability from run to run. This stochasticity reflects the real-world uncertainty in demographic processes. As a consequence of this randomness, when starting with small initial populations, computations showed that the future population and growth patterns can exhibit variability across different simulations. Sometimes, the results closely resemble the predictions of the deterministic Leslie model, aligning with its constant outcomes. However, in other runs, the Branching process has revealed more unpredictable growth projections.This variability underlines the importance of considering demographic stochasticity in population projections when dealing with small populations. Thus, the simulations demonstrate the need to account for probabilistic factors when analysing and predicting population trends for small populations.

Deterministic models like the Leslie model tend to provide overly optimistic predictions for populations, especially for small populations [15] [21]. They may underestimate the risk of extinction because they do not consider the possibility of demographic stochasticity and random events that could possibly lead a population to extinction. Therefore, in scenarios concerning small populations, such as those encountered in conservation biology, where populations near extinction are often analysed, the deterministic Leslie model may not provide as accurate or realistic results as the Branching process model.

Chapter 6

Conclusion and Discussion

The aim of this thesis was to investigate the long-term behaviour within the Leslie model for population growth, utilizing various mathematical theorems. We investigated how Perron-Frobenius theorem, Doeblin's theorem and Branching processes can be applied to the Leslie model.

The Perron-Frobenius theorem guaranteed the existence of a dominant eigenvalue for an irreducible Leslie matrix. By diagonalization of the Leslie matrix $\boldsymbol{L}$, we derived the following equation for the population at time k, for large values of k

$$P(k) \cong \lambda_1^k c v_1.$$

From this approximation we observed that for a dominant eigenvalue $\lambda_1 > 1$, the population at time k will increase exponentially. If $\lambda_1 < 1$, the population faces extinction, and for $\lambda_1 = 1$, the population is stable and stationary, indicating that the births exactly compensate the deaths. Furthermore, we derived equation

$$P(k) \cong \lambda_1 P(k-1).$$

This implied that each age distribution vector is approximately a scalar multiple of the previous age distribution vector, illustrating that the population distribution will remain constant. This scalar multiple λ_1 is the dominant eigenvalue of the Leslie matrix. Thus, the asymptotic growth rate of a population is given by the dominant eigenvalue λ_1. The dominant eigenvalue and corresponding eigenvector thereby represent the long-term behaviour of a population; the dominant eigenvalue indicates the long-term population growth, and the corresponding eigenvector indicates the long-term age distribution. Furthermore, this implied that a population described by a Leslie matrix, asymptotically reaches a stable age distribution that is independent of its initial age structure.

The practical utility of Doeblin's theorem applied to the Leslie model was investigated. Doeblin's theorem provided insight into the long-term behaviour of and convergence properties of Markov chains. However, due to the complexities of describing both the survival rates s_i and birth rates b_i within the transition probabilities of the corresponding transition probability matrix, it was not feasible to formulate the complete Leslie model as Markov chain. Doeblin's theorem could be applied to a simplified version of the Leslie model, which exclusively considered the Survival matrix. An additional state was introduced to represent "death". This "death" state was chosen to be state j_0 in Doeblin's theorem. In this context, the application of Doeblin's theorem provided a realistic confirmation that the population would ultimately converge to a stationary population distribution, specifically represented as $\pi = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$, where the final element signifies that everyone within the population has passed away. Moreover, it was observed that in this specific scenario involving a population divided into m age classes, the convergence to this state π would occur over the course of m time steps.

In order to include the Fertility matrix in the analysis of the Leslie model, we then formulated the

Fertility matrix as a Branching process. The birth expectations were converted into birth probabilities through the use of a Poisson distribution and survival probabilities were modelled with a Binomial distribution. Various initial population sizes were tested in simulations for both the Leslie model and the Branching process model. The results of these simulations revealed some important insights into the behaviour of these models under different conditions. For larger populations, we observed that the Branching process closely resembled the Leslie model. The growth rates exhibited relatively minor fluctuations. In other words, when dealing with substantial population sizes, the Branching process and the Leslie model produced comparable results, indicating that the stochastic elements introduced by the Branching process did not significantly affect the overall population dynamics in these cases. Conversely, for smaller populations, a more significant difference appeared between the Branching process and the Leslie model. This difference became more pronounced as population size decreased. The introduction of demographic stochasticity shed light on the importance of considering probabilistic factors in the context of small populations, where deterministic models potentially fall short in capturing the intricacies of real-world dynamics. This emphasized the strength of the Leslie model, as well as the utility of the Perron-Frobenius theorem for larger populations.

One of the fundamental assumptions of the Leslie model is that it considers a closed population, which means it does not account for immigration or emigration. This assumption limits its applicability to real-world scenarios. Consequently, in situations where populations are highly volatile due to emigration, the Leslie model may provide less accurate predictions and a limited understanding of population dynamics. Furthermore, the Leslie model describes a population under the assumption of constant demographic parameters, which may not hold in all real-world scenarios. Future research could incorporate time-varying parameters to capture more dynamic population dynamics. Moreover, real-world populations are influenced by various other factors besides the birth rates and survival rates. Integrating additional ecological variables and environmental influences into the population models could lead to a more realistic model. Another constraint in our investigation relates to the exclusive focus on a single gender within the Leslie model. In many species, gender-based differences in birth and survival rates, significantly influence population dynamics. Therefore, expanding the Leslie model by including both the male and female sex could create more realistic population projections. Additionally, it is important to acknowledge that effectively applying the Leslie model necessitates precise estimations of age-specific birth and survival rates, which can be challenging to obtain.

The Leslie model, with its assumption of an unlimited environment, demonstrates that populations tend to reach a stable age distribution and can exhibit exponential growth, resembling Malthusian behaviour. Though Malthus claimed a population can grow exponentially, he also argued that the Earth's ability to provide subsistence, like food, increases at a slower linear rate. This fundamental imbalance implies that population growth can surpass the Earth's capacity to sustain it, leading to "checks" such as famine, disease, or war, as expressed in Malthus's famous quote, *"The power of population is indefinitely greater than the power in the earth to produce subsistence for man."* [20].

Bibliography

[1] Nicolas Bacaër. *A short history of mathematical population dynamics*. Vol. 618. Springer, 2011.

[2] Michael Begon, John L Harper, Colin R. Townsend, et al. *Ecology. Individuals, populations and communities.* Blackwell scientific publications, 1986.

[3] Michael Begon, Martin Mortimer, and David J. Thompson. *Population ecology: a unified study of animals and plants*. John Wiley & Sons, 2009.

[4] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

[5] Hal Caswell. "Life cycle models for plants". In: *Lectures on mathematics in the life sciences* 18 (1986), pp. 171–233.

[6] Hal Caswell. "Matrix population models: construction, analysis, and interpretation. 2nd edn Sinauer Associates". In: *Inc., Sunderland, MA* (2001).

[7] Joel E. Cohen. "Comparative statics and stochastic dynamics of age-structured populations". In: *Theoretical population biology* 16.2 (1979), pp. 159–171.

[8] Joel E. Cohen. "Ergodic theorems in demography". In: *Bulletin of the American Mathematical Society* 1 (1979), pp. 275–295.

[9] Jim Michael Cushing. *An introduction to structured population dynamics*. SIAM, 1998.

[10] J.B. Fraleigh and R. Beauregard. *Linear Algebra: Pearson New International Edition*. Pearson Education Limited. ISBN: 9781292042725.

[11] Feliks Rouminovich Gantmacher and Joel Lee Brenner. *Applications of the Theory of Matrices*. Courier Corporation, 2005.

[12] Geoffrey Grimmett and Dominic J.A. Welsh. *Probability: an introduction*. Oxford University Press, 2014.

[13] John Impagliazzo. *Deterministic aspects of mathematical demography: An investigation of the stable theory of population including an analysis of the population statistics of Denmark*. Vol. 13. Springer Science & Business Media, 2012.

[14] M. Iosifescu. *Finite Markov processes and their applications*. Wiley Sons, 1980.

[15] Bruce E. Kendall and Marion E. Wittmann. "A stochastic model for annual reproductive success". In: *The American Naturalist* 175.4 (2010), pp. 461–468.

[16] Axelrod Kimmel. *Branching Processes in Biology*. Springer, 2002.

[17] Steven R. Lay. *Analysis with an introduction to proof*. Pearson, 2014.

[18] Patrick H. Leslie. "On the use of matrices in certain population mathematics". In: *Biometrika* 33.3 (1945), pp. 183–212.

[19] Patrick H. Leslie. "Some further notes on the use of matrices in population mathematics". In: *Biometrika* 35.3/4 (1948), pp. 213–245.

[20] Thomas Robert Malthus and Donald Winch. *Malthus:'An Essay on the Principle of Population'*. Cambridge university press, 1992.

[21] Brett A. Melbourne and Alan Hastings. "Extinction risk depends strongly on factors contributing to stochasticity". In: *Nature* 454.7200 (2008), pp. 100–103.

[22] Carl D. Meyer. *Matrix analysis and applied linear algebra*. Vol. 188. Siam, 2023.

[23] *Population Projections - Population and Demography - Eurostat*. URL: https://ec.europa.eu/eurostat/web/population-demography/population-projections.

[24] John R. Skalski, Kristin E. Ryding, and Joshua Millspaugh. *Wildlife demography: analysis of sex, age, and count data*. Elsevier, 2010.

[25] Daniel W. Stroock. *An introduction to Markov processes*. Vol. 230. Springer Science & Business Media, 2013.

Appendix A

$$P := \begin{bmatrix} 0 & s_1 & 0 & 0 & \cdots & 0 & 1-s_1 \\ 0 & 0 & s_2 & 0 & \cdots & 0 & 1-s_2 \\ 0 & 0 & 0 & s_3 & \cdots & 0 & 1-s_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & s_{m-1} & 1-s_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$P^2 = \begin{bmatrix} 0 & 0 & s_1 s_2 & 0 & 0 & \cdots & 0 & 1+s_1(1-s_2)-s_1 \\ 0 & 0 & 0 & s_2 s_3 & 0 & \cdots & 0 & 1+s_2(1-s_3)-s_2 \\ 0 & 0 & 0 & 0 & s_3 s_4 & \cdots & 0 & 1+s_3(1-s_4)-s_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & s_{m-2}s_{m-1} & 1+s_{m-2}(1-s_{m-1})-s_{m-2} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$P^2 = \begin{bmatrix} 0 & 0 & s_1 s_2 & 0 & 0 & \cdots & 0 & 1-s_1 s_2 \\ 0 & 0 & 0 & s_2 s_3 & 0 & \cdots & 0 & 1-s_2 s_3 \\ 0 & 0 & 0 & 0 & s_3 s_4 & \cdots & 0 & 1-s_3 s_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & s_{m-2}s_{m-1} & 1-s_{m-1}s_{m-2} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$P^3 = \begin{bmatrix} 0 & 0 & 0 & s_1 s_2 s_3 & 0 & \cdots & 0 & 1+s_1 s_2(1-s_3)+s_1(1-s_2)-s_1 \\ 0 & 0 & 0 & 0 & s_2 s_3 s_4 & \cdots & 0 & 1+s_2 s_3(1-s_4)+s_2(1-s_3)-s_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & s_{m-3}s_{m-2}s_{m-1} & 1+s_{m-3}s_{m-2}(1-s_{m-1})+s_{m-3}(1-s_{m-2})-s_{m-3} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1+s_{m-2}s_{m-1}+s_{m-2}(1-s_{m-1})-s_{m-2} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$P^3 = \begin{bmatrix} 0 & 0 & 0 & s_1s_2s_3 & 0 & \cdots & 0 & 1 - s_1s_2s_3 \\ 0 & 0 & 0 & 0 & s_2s_3s_4 & \cdots & 0 & 1 - s_2s_3s_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & s_{m-3}s_{m-2}s_{m-1} & 1 - s_{m-3}s_{m-2}s_{m-1} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

$$P^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & s_1s_2s_3s_4 & \cdots & 0 & 1 - s_1s_2s_3s_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & s_{m-4}s_{m-3}s_{m-2}s_{m-1} & 1 - s_{m-4}s_{m-3}s_{m-2}s_{m-1} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Appendix B: Python code Leslie model

```python
import random
import numpy as np
import matplotlib.pyplot as plt

# Define birth rates and survival rates
b1 = 1
b2 = 2
b3 = 3
b4 = 4

s1 = 0.5
s2 = 0.3
s3 = 0.2
s4 = 0

# Define the initial population in each age class
p1 = 100
p2 = 100
p3 = 100
p4 = 100

# Create the Leslie matrix and Initial population vector
L = np.array([[b1, b2, b3, b4],
              [s1, 0, 0, 0],
              [0, s2, 0, 0],
              [0, 0, s3, 0]])

p = np.array([[p1],
              [p2],
              [p3],
              [p4]], dtype=np)

# Lists of birth rates and survival rates
list_birth_rates = [b1, b2, b3, b4]
list_survival_rates = [s1, s2, s3, s4]
list_initial_population = p.tolist()

print("Leslie matrix:", L)
print("initial population", p)
print("Birth Rates:", list_birth_rates)

# Computation population after n time steps and corresponding growth rate using original Leslie model
n = 10
```

```python

total_population_list = []
population_vectors = []
growth_rates = []
age_class_growth_rates = []  # List to store age class-specific growth rates

for n in range(n + 1):
    population_at_n = np.linalg.matrix_power(L, n).dot(p)
    total_population = np.sum(population_at_n)
    total_population_list.append(total_population)
    population_vectors.append(population_at_n)
    print(f"Population at n={n}: {population_at_n}")
    print(f"The total population after {n} time steps using the Leslie model is", total_population)

    if n > 0:
        growth_rate = ((total_population) / total_population_list[n - 1])
        growth_rates.append(growth_rate)
        print(f"The total population has grown with growth rate: {growth_rate:.4f}")

        # Calculate age class-specific growth rates
        age_class_growth_rate = (population_at_n / population_vectors[n - 1])
        age_class_growth_rates.append(age_class_growth_rate.tolist())


# Print age class-specific growth rates for each time step in the Leslie model
for time_step, rates in enumerate(age_class_growth_rates):
    print(f"At time step {time_step}:")
    for i, rate in enumerate(rates):
        print(f"  Age class {i + 1}: {rate[0]:.4f}")

# Plot the total population over time using the Leslie model
plt.figure(figsize=(10, 6))
plt.plot(np.arange(n + 1), total_population_list, label="Total population", color="Black")

plt.xlabel("Time step (n)")
plt.ylabel("Total population")
plt.legend()
plt.xticks(range(0, n + 1))
plt.title("Total population over time using Leslie Model")
plt.grid(True)
plt.show()

# Plot the distributions over time using the Leslie model
population_vectors = np.array(population_vectors)

# Plot the population distribution over time
plt.figure(figsize=(10, 6))
age_classes = np.arange(len(list_birth_rates))
for i in range(population_vectors.shape[1]):
```

```
    plt.plot(np.arange(n+1), population_vectors[:, i], label=f"Age {i + 1}")

plt.xlabel("Time step (n)")
plt.ylabel("Population")
plt.legend(["Age class 1", "Age class 2", "Age class 3", "Age class 4"])
plt.xticks(range(0, n + 1))
plt.title("Population distribution over time using Leslie Model")
plt.grid(True)
plt.show()
```

Appendix C: Python code Leslie model as a Branching Process

```
import random
import numpy as np
import matplotlib.pyplot as plt

#Define birth rates and survival rates
b1 = 1
b2 = 2
b3 = 3
b4 = 4

s1 = 0.5
s2 = 0.4
s3 = 0.3
s4 = 0

#Define the initial population in each age class
p1=100
p2=200
p3=200
p4=300

#Create the Leslie matrix using the defined rates
L = np.array([[b1, b2, b3, b4],
              [s1, 0, 0, 0],
              [0, s2, 0, 0],
              [0, 0, s3, 0]])

#Create the Initial population vector
p = np.array([[p1],
              [p2],
              [p3],
              [p4]])

#Create lists of birth rates and survival rates
list_birth_rates= [b1, b2, b3, b4]
list_survival_rates= [s1, s2, s3, s4]
list_initial_population= p.tolist()

print("Leslie matrix:", L)
print("initial population", p)

```

```
n=10

#Formulating the Branching process
def Population(n):
    generation_vector = p
    generation = 0

    while generation < n:
        generation_vector = BranchingStep(generation_vector)
        generation += 1
    return generation_vector

def BranchingStep(generation_vector):
    offspring=0
    for i in range(0,len(list_birth_rates)):
        y= list_birth_rates[i]
        individual=0
        while individual < generation_vector[i]:
            newOffspring = np.random.poisson(y)
            offspring += newOffspring
            individual += 1

    generationNew= np.zeros(len(list_birth_rates))
    generationNew[0]=offspring
    for i in range (1, len(list_birth_rates)):
        generationNew[i]= np.random.binomial(generation_vector[i-1],
            list_survival_rates[i-1])
    return generationNew

total_population_list_branching = []
population_vectors_branching = []
growth_rates_branching = []
age_class_growth_rates_branching=[]

# Calculate and store the total population and distributions for both
    models
for n in range(n + 1):
    population_at_n_branching = Population(n)
    total_population_branching = np.sum(population_at_n_branching)
    total_population_list_branching.append(total_population_branching)
    population_vectors_branching.append(population_at_n_branching.
        flatten())
    print(f"Population at n={n}: {population_at_n_branching}")
    print(f"The total population after {n} time steps using the
        Branching process is", total_population_branching)

    if n > 0:
        growth_rate_branching = ((total_population_branching) /
            total_population_list_branching[n - 1])
        growth_rates_branching.append(growth_rate_branching)
        print(f"(The total population has grown with growth Rate: {
            growth_rate_branching:.4f})")

        age_class_growth_rate_branching= (population_at_n_branching /
            population_vectors_branching[n-1])
```

```
        age_class_growth_rates_branching.append(
           age_class_growth_rate_branching.tolist())


# Print age class-specific growth rates for each time step
for time_step, rates in enumerate(age_class_growth_rates_branching):
    print(f"At time step {time_step}:")
    for i, rate in enumerate(rates):
        print(f"  Age class {i + 1}: {rate:.4f}")


#plot the total population over time using Branching process
plt.figure(figsize=(10, 6))
plt.plot(np.arange(n + 1), total_population_list_branching, label="Total
    Population", color="green")

plt.xlabel("Time step (n)")
plt.ylabel("Total population")
plt.legend()
plt.xticks(range(0, n + 1))
plt.title("Total population over time using Branching process")
plt.grid(True)
plt.show()

#plot the distributions over time using Branching process
population_vectors_branching = np.array(population_vectors_branching)

plt.figure(figsize=(10, 6))
age_classes = np.arange(len(list_birth_rates))
for i in range(population_vectors_branching.shape[1]):
    plt.plot(np.arange(n + 1), population_vectors_branching[:, i], label
       =f"Branching Model (Age {i + 1})")

plt.xlabel("Time step (n)")
plt.ylabel("Population")
plt.legend(["Age class 1", "Age class 2", "Age class 3", "Age class 4"])
plt.xticks(range(0, n + 1))
plt.title("Population distribution over time using Branching process")
plt.grid(True)
plt.show()
```