

Matrix Models in Population Dynamics: Exploring Structure and Applications

ARYAMAN BAHL¹, SIDDHARTH AGARWAL², SUJAL KESHRI³

¹CND, International Institute of Information Technology, Hyderabad

²CSE, International Institute of Information Technology, Hyderabad

³ CSE, International Institute of Information Technology, Hyderabad

Abstract

This paper provides an overview of matrix models in population dynamics, focusing on population transitions, the Leslie matrix model, the SIR model, applications, and extensions. It discusses key findings and emphasizes the usefulness of matrix models in understanding and predicting population dynamics. The paper explores sensitivity analysis to examine the impact of birth rates, death rates, and migration on populations. It also addresses the effects of age-specific fertility and mortality rates on population growth and the prediction of population structure using stable age distribution. Extensions include stage-structured models, density-dependent models, and metapopulation models. The paper concludes by summarizing insights gained from matrix models and their importance in studying population dynamics.

1 Introduction

Population dynamics is a vital field of study with applications in ecology, demography, epidemiology, and conservation biology. This paper provides an introduction to the fundamental concepts and models used in population dynamics, focusing on matrix models. It explores the construction and interpretation of population transition matrices. The paper discusses the calculation of important population parameters and the prediction of future population sizes using matrix multiplication. Additionally, it examines the implications of the SIR model in the field of epidemiology. The practical applications of matrix models, their extensions, and alternative approaches are also discussed, offering valuable insights into the field of population dynamics.

1.1 Matrix Models in Population Dynamics

Matrix models provide a powerful framework for analyzing population dynamics. They allow us to capture the complex interactions between birth rates, death rates, immigration, and emigration. By representing these dynamics in a matrix format, we can systematically study how populations change over time.

1.2 Construction and Interpretation of Population Transition Matrices

One of the key components of matrix models is the population transition matrix. This matrix describes the probabilities of individuals transitioning from one state (e.g., age or stage class) to another over a specified time interval. The paper explores the construction of these matrices and delves into their interpretation within the context of population dynamics.

1.3 Calculation of Important Population Parameters

Matrix models enable us to calculate crucial population parameters, such as the net reproductive rate (R_0), which represents the average number of offspring produced by an individual throughout its lifetime. These parameters provide insights into population growth, stability, and potential for future change.

1.4 Prediction of Future Population Sizes

Using matrix multiplication, it becomes possible to predict future population sizes based on the initial population distribution and the population transition matrix. This predictive capability is valuable for assessing population trends and planning appropriate conservation or management strategies.

1.5 Implications of the SIR Model in Epidemiology

The paper also discusses the SIR (Susceptible-Infectious-Recovered) model, which is widely used in epidemiology to understand and control the spread of infectious diseases within populations. The SIR model incorporates concepts of population dynamics, disease transmission, and recovery, providing valuable insights for public health interventions.

1.6 Practical Applications of Matrix Models

The practical applications of matrix models in population dynamics are explored. Sensitivity analysis allows us to assess the impact of changes in birth rates, death rates, and migration on population dynamics. The effects of age-specific fertility and mortality rates on population growth can be investigated. Stable age distribution, a concept derived from matrix models, helps predict population structure and age composition.

1.7 Extensions and Alternative Approaches

Matrix models serve as a foundation for more complex population models. The paper discusses stage-structured models that account for different life stages, density-dependent models that consider population size effects, and metapopulation models that

study interactions among multiple subpopulations. Furthermore, the role of stochasticity and uncertainty in population dynamics is examined. Alternative modeling approaches beyond matrix models are also considered.

2 Problem Statement

2.1 Markov Model

One problem lies in the assumption of stationarity in Markov chain models. Markov chain models assume that the probabilities of transitioning between states remain constant over time. However, ecological systems often exhibit non-stationary dynamics, such as population fluctuations, environmental changes, and disturbances. Failing to account for these non-stationary dynamics can lead to inaccurate predictions and hinder the understanding of population dynamics.

2.2 Leslie Model

In the case of Leslie models, which are commonly used to describe age-structured populations, a challenge arises from the assumption of time-invariant vital rates. The model assumes that birth rates, survival rates, and reproductive rates remain constant over time. However, in many real-world populations, these vital rates can vary due to factors such as environmental conditions, resource availability, and interspecific interactions. Ignoring temporal variability in vital rates can result in unrealistic population projections and misinformed management decisions.

2.3 SIR Model

SIR models, widely used to study infectious disease dynamics, encounter challenges related to the assumptions of homogeneous mixing and constant transmission rates. These models assume that individuals within a population have equal chances of interacting and infecting each other, disregarding heterogeneities in contact patterns and transmission dynamics. In reality, social structures, spatial heterogeneity, and individual behaviors can significantly influence disease transmission. Neglecting these complexities can lead to inaccurate epidemic predictions and inadequate public health interventions.

2.4 Food Web

In the context of food webs, the application of Markov chain models faces challenges associated with the assumption of a closed system and the complexity of trophic interactions. Markov chain models often assume a closed food web, neglecting the influence of external factors such as immigration, emigration, and the addition or removal of species. Additionally, capturing the intricate trophic interactions, including indirect effects and feedback loops, presents a challenge due to the high dimensionality and

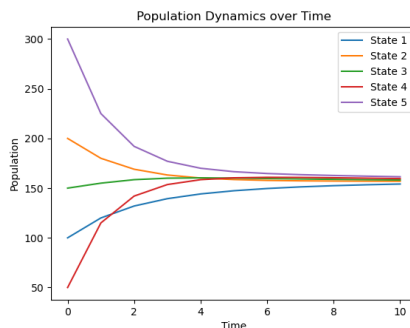
nonlinear nature of food webs. Failing to account for these complexities can limit the predictive power and ecological insights provided by food web models.

3 Matrix Models in Population Dynamics

Matrix models are powerful tools used in population dynamics to study the dynamics of structured populations. In a matrix model, population characteristics, such as age, size, or stage, are organized into discrete classes. The population transition from one class to another is represented by a matrix, where each element represents the probability or rate of transition. These models allow for the analysis of population growth, age or stage structure, and the effects of vital rates (birth, death, and migration) on population dynamics. Matrix models provide insights into population stability, extinction risk, and the impacts of environmental changes. They are widely used in ecology, conservation biology, and demography to inform conservation strategies, wildlife management, and understanding population responses to environmental disturbances.

3.1 Population Dynamics over time

Population dynamics studies changes in population size and composition over time. Markov chains are useful for modeling population dynamics, representing transitions between different states or classes. Transition probabilities in the chain indicate the likelihood of moving between states. Markov chain models enable the analysis of population growth, stability, and the effects of factors on population structure. They aid in predicting future trends and understanding the impacts of interventions. Markov chains form the basis for complex models incorporating birth/death rates, migration, and environment. They support decision-making, policy formulation, and addressing population management challenges in ecology, epidemiology, economics, and social sciences.



3.2 Population Growth varying with Time,Space,Stability graph

3.2.1 Time

The rate of change in population size can be calculated by multiplying the population vector by the matrix. This gives a new vector that represents the population size at a later time.

$$\mathbf{x}(t+1) = A\mathbf{x}(t) \quad (1)$$

For example, let's say that the population vector is $\mathbf{x}=(100,200,300)$, and the matrix is $A=(1.05,0.02,0.01)$. This means that the birth rate is 5 percent, the death rate is 2 percent, and the immigration rate is 1 percent.

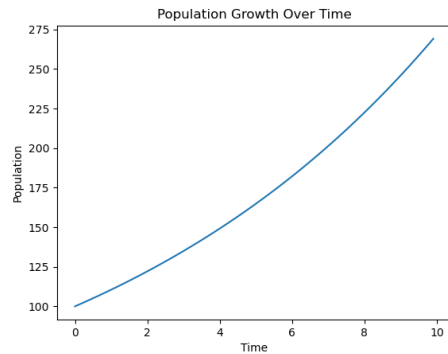
If we multiply the population vector by the matrix, we get the new vector $\mathbf{x}(t+1)=(105,204,303)$. This means that the population size will increase to 105 in the next time period.

This is just a simple example, but it illustrates how population size can be modeled using linear algebra. This can be used to predict how population size will change over time, and to help policymakers make decisions about how to manage population growth. Here are some additional points about how population size depends on time with respect to linear algebra:

The rate of population growth is not always constant. It can change over time due to factors such as war, famine, disease, and technological advances.

The population size of a species can also be affected by the population size of other species. For example, the population size of predators can affect the population size of prey.

Linear algebra can be used to model the interactions between different species, and to predict how these interactions will affect population sizes over time.



3.2.2 Space

Sure, here is a short note on how population size depends on space with respect to linear algebra: In linear algebra, a vector can be used to represent a population. The vector's components can represent the population's size, density, and distribution in space. The population size can be calculated by taking the sum of the components of the vector. The population density can be calculated by dividing the population size by the area of the space. The population distribution can be calculated by finding the average value of the components of the vector.

The population size can depend on space in a number of ways. For example, the population size may be larger in areas with more resources, such as food and water. The population size may also be larger in areas with a more favorable climate. The population size may also be larger in areas with a shorter distance to other populations.

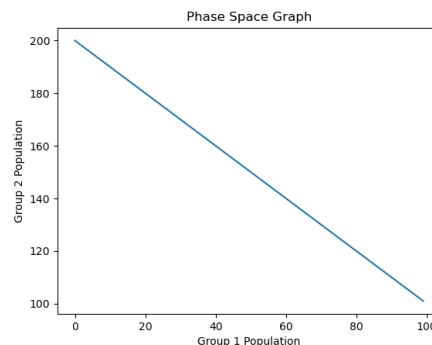
Linear algebra can be used to model the relationship between population size and space. This can be used to predict how population size will change over time, or to identify factors that influence population size.

Here are some examples of how linear algebra can be used to study population size:

A researcher can use linear algebra to model the relationship between population size and the availability of food. This could be used to predict how a population might change in size if there is a drought or a flood.

A conservationist can use linear algebra to model the relationship between population size and the amount of habitat available. This could be used to identify areas where conservation efforts are needed to protect a population from extinction.

A city planner can use linear algebra to model the relationship between population size and the availability of housing. This could be used to plan for future population growth and to ensure that there is enough housing available for everyone.



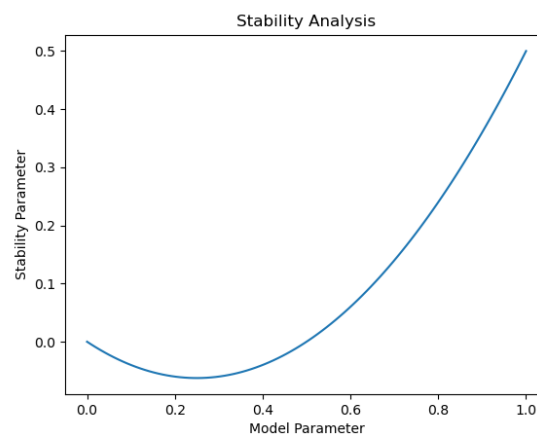
3.2.3 Stability

In addition to diversity, population size can also affect stability through the concept of carrying capacity. Carrying capacity is the maximum population size that an environment can sustain over time. If a population exceeds its carrying capacity, it will begin to decline. This can happen due to a variety of factors, such as resource depletion or environmental degradation. Larger populations are more likely to exceed their carrying capacity than smaller populations, as they put more stress on the environment.

The relationship between population size and stability can be seen in a number of different contexts. For example, in ecology, it is known that larger populations are more likely to be stable than smaller populations. This is because larger populations have more genetic diversity, which can help them to adapt to changes in their environment. In economics, it is also known that larger economies are more stable than smaller economies. This is because larger economies have more diversified industries, which can help them to weather economic downturns.

Here are some additional examples of how population size can affect stability:

In a political system, a larger population can make it more difficult for one group to gain control and maintain power. This is because there are more people with different interests and opinions, which makes it more difficult for any one group to dominate. In a business, a larger customer base can make the business more stable. This is because there are more people who are buying the company's products or services, which provides the company with a steady stream of income. In a community, a larger population can make it more difficult for crime to occur. This is because there are more people who are watching out for each other and reporting crimes to the police.



4 Markov Chains

A Markov chain or Markov process is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. Informally, this may be thought of as, "What happens next depends only on the state of affairs now." A countably infinite sequence, in which the chain moves state at discrete time steps, gives a discrete-time Markov chain (DTMC). A continuous-time process is called a continuous-time Markov chain (CTMC). It is named after the Russian mathematician Andrey Markov.

4.1 Types Of Markov Chain

The system's state space and time parameter index need to be specified. The following table gives an overview of the different instances of Markov processes for different levels of state space generality and for discrete time v. continuous time:

Time	Countable State Space	Continuous or General State Space
Discrete	Markov Chain on a countable space	Markov chain on a measurable state space
Continuous	Continuous time Markov space	Any continuous stochastic process with the Markov property

4.2 Prerequisites for Markov Chain

4.2.1 Stochastic process

A stochastic (or random) process is a sequence of experiments for which the outcome at any stage depends on chance. In a simple model, there are a finite number of possible outcomes, referred to as states, and the process occurs in discrete time.

Let S denote the set of states. Then, the stochastic process is a sequence s_0, s_1, s_2, \dots , where all $s_n \in S$ depend on chance.

4.2.2 Bernoulli scheme

A Bernoulli scheme is a sequence of independent random events. In the sequence s_0, s_1, s_2, \dots , any outcome s_n is independent of the others.

For any integer $n \geq 0$, we have a probability distribution $p(n)$ on S . This means that each state $s \in S$ is assigned a value $p(n)_s \geq 0$ such that $\sum_{s \in S} p(n)_s = 1$. Then, the probability of the event $s_n = s$ is $p(n)_s$.

The Bernoulli scheme is called stationary if the probability distributions $p(n)$ do not depend on n .

4.3 Markov Chain

A Markov chain is a stochastic process with discrete time such that the probability of the next outcome depends only on the previous outcome.

Let $S = \{1, 2, \dots, k\}$. The Markov chain is determined by transition probabilities $p(t)_{ij}$, $1 \leq i, j \leq k$, $t \geq 0$, and by the initial probability distribution q_i , $1 \leq i \leq k$. Here, q_i is the probability of the event $s_0 = i$, and $p(t)_{ij}$ is the conditional probability of the event $s_{t+1} = j$ provided that $s_t = i$. By construction, we have $p(t)_{ij}, q_i \geq 0$, $\sum_i q_i = 1$, and $\sum_j p(t)_{ij} = 1$.

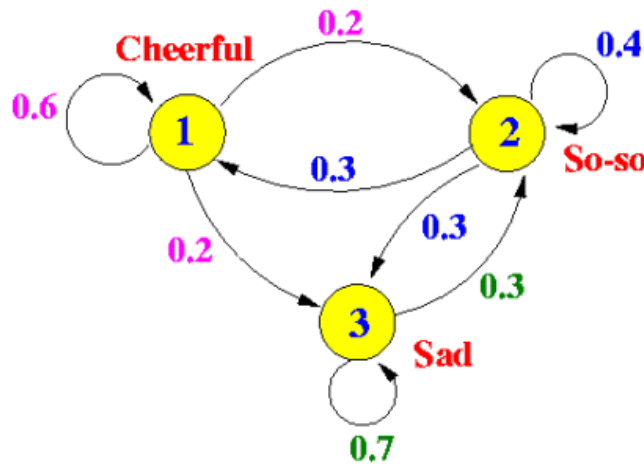
We shall assume that the Markov chain is time-independent, i.e., transition probabilities do not depend on time: $p(t)_{ij} = p_{ij}$. Then, a Markov chain on $S = \{1, 2, \dots, k\}$ is determined by a probability vector $x_0 = (q_1, q_2, \dots, q_k) \in R^k$ and a $k \times k$ transition matrix $P = (p_{ij})$. The entries in each row of P add up to 1.

Let s_0, s_1, s_2, \dots be the Markov chain. Then the vector x_0 determines the probability distribution of the initial state s_0 .

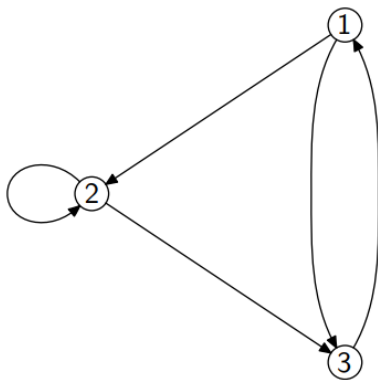
4.3.1 Visualization:

We can visualize Markov chains with directed graphs.

- Vertices of the graph represent the states (entries of a vector).
- Arrows of the graph represent the transitions and their probability.
- We can write the transition probabilities into a stochastic matrix P .



4.3.2 Taking Example of a Random Walk:



$$TransitionMatrix = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{pmatrix}$$

4.3.3 Very primitive weather model:

Two states: "sunny" (1) and "rainy" (2). Transition matrix: $P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$. Suppose that $\mathbf{x}_0 = (1, 0)$ (sunny weather initially).

Make a long-term weather prediction. The probability distribution of weather for day n is given by the vector $\mathbf{x}_n = Q^n \mathbf{x}_0$, where $Q = P$.

To compute Q^n , we need to diagonalize the matrix $Q = \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix}$. $\det(Q - \lambda I) =$

$$\begin{vmatrix} 0.9 - \lambda & 0.5 \\ 0.1 & 0.5 - \lambda \end{vmatrix} = \lambda^2 - 1.4\lambda + 0.4 = (\lambda - 1)(\lambda - 0.4). \text{ Two eigenvalues: } \lambda_1 = 1,$$

$$\lambda_2 = 0.4. (Q - I)\mathbf{v} = 0 \iff \begin{pmatrix} -0.1 & 0.5 \\ 0.1 & -0.5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff (x, y) = t(5, 1), t \in R.$$

$$(Q - 0.4I)\mathbf{v} = 0 \iff \begin{pmatrix} 0.5 & 0.5 \\ 0.1 & 0.1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff (x, y) = t(-1, 1), t \in R.$$

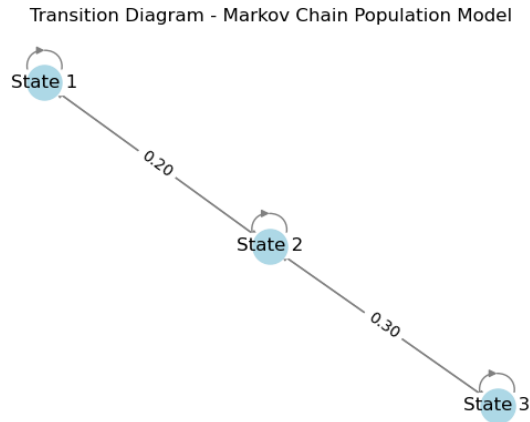
$$\mathbf{v}_1 = \begin{pmatrix} 5 \\ 1 \end{pmatrix} \text{ and } \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \text{ are eigenvectors of } Q \text{ belonging to eigenvalues 1 and}$$

$$0.4, \text{ respectively. } \mathbf{x}_0 = \alpha \mathbf{v}_1 + \beta \mathbf{v}_2 \iff \begin{cases} 5\alpha - \beta = 1 \\ \alpha + \beta = 0 \end{cases} \iff \begin{cases} \alpha = \frac{1}{6} \\ \beta = -\frac{1}{6} \end{cases} \text{ Now}$$

$$\mathbf{x}_n = Q^n \mathbf{x}_0 = Q^n(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha \mathbf{v}_1 + (0.4)^n \beta \mathbf{v}_2, \text{ which converges to the vector } \alpha \mathbf{v}_1 = \begin{pmatrix} \frac{5}{6} \\ \frac{1}{6} \end{pmatrix}$$

as $n \rightarrow \infty$. The vector $\mathbf{x}_\infty = \begin{pmatrix} 5 \\ 6 \\ 1 \\ 6 \end{pmatrix}$ gives the limit distribution. Also, it is a steady-state vector.

4.3.4 Representation of population growth using Markovs chain



5 Leslie Model

The Leslie matrix is a discrete, age-structured model of population growth that is very popular in population ecology named after Patrick H. Leslie .The Leslie matrix (also called the Leslie model) is one of the most well-known ways to describe the growth of populations (and their projected age distribution), in which a population is closed to migration, growing in an unlimited environment, and where only one sex, usually the female, is considered. The Leslie matrix is used in ecology to model the changes in a population of organisms over a period of time. In a Leslie model, the population is divided into groups based on age classes.The Leslie matrix is a square matrix with the same number of rows and columns as the population vector has elements. The (i,j)th cell in the matrix indicates how many individuals will be in the age class i at the next time step for each individual in stage j. At each time step, the population vector is multiplied by the Leslie matrix to generate the population vector for the subsequent time step.

5.1 Leslie Matrix Construction

To build a matrix, the following information must be known from the population:

- n_x : the count of individuals (n) of each age class x
- s_x : the fraction of individuals that survives from age class x to age class $x + 1$
- f_x : fecundity, the per capita average number of female offspring reaching n_0 born from mother of age class x . More precisely, it can be viewed as the number of offspring produced at the next age class $x + 1$ weighted by the probability of reaching the next age class. Therefore, $f_x = s_x b_{x+1}$

From the observations that n_0 at time $t + 1$ is simply the sum of all offspring born from the previous time step and that the organisms surviving to time $t + 1$ are the organisms at time t surviving at probability s_x , one gets $n_{x+1} = s_x n_x$. This implies the following matrix representation:

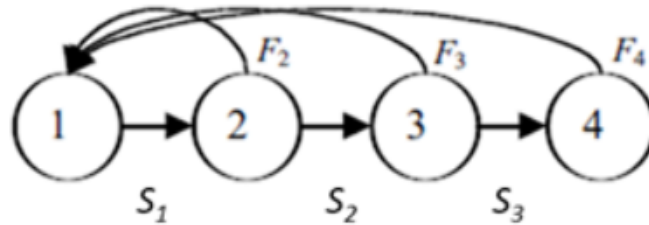
$$\begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{\omega-1} \end{bmatrix}_{t+1} = \begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_{\omega-2} & f_{\omega-1} \\ s_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{\omega-2} & 0 \end{bmatrix} \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{\omega-1} \end{bmatrix}_t$$

where ω is the maximum age attainable in the population.

The Leslie model is very similar to a discrete-time Markov chain. The main difference is that in a Markov model, one would have $f_x + s_x = 1$ for each x , while the Leslie model may have these sums greater or less than 1.

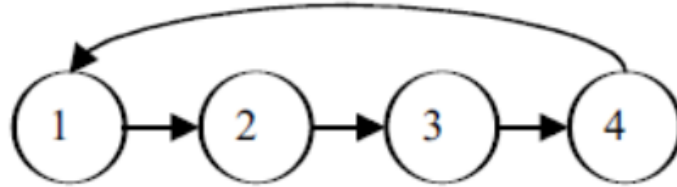
5.2 Determining Stable Age Distribution

Whether we are dealing with age classes or ages, individuals are grouped into discrete classes that are of equal duration for modeling purposes. A typical life cycle of a population with age-class structure is:



The life cycle of a population with age-class structure is depicted using circles to represent different age classes. In this example, we focus on a population with four age

classes. The horizontal arrows connecting the circles represent survival probabilities (S_x), indicating the likelihood of an individual at age x surviving to age $x + 1$. Notably, there is no arrow from age four to age five, signifying a 0 probability of surviving to the fifth age class. The curved arrows at the top represent births, with all arrows leading to age 1 as newborns enter the first age class. The diagram illustrates that individuals aged 2, 3, and 4 are capable of reproduction, while those in age class 1 do not reproduce. If only age class 4 individuals reproduced, the diagram would require modification.



5.3 Population Projection Matrix

For example, with 3 age classes (ages 0-2, females only) the Leslie matrix would take the form:

$$A = \begin{pmatrix} F_0 & F_1 & F_2 \\ S_0 & 0 & 0 \\ 0 & S_1 & 0 \end{pmatrix}$$

F_x = age specific fecundity

S_x = age specific survival rate

Leslie Matrix (A) replaces λ in the shift from a simple exponential growth curve to an age-structured population growth model (note, this is still exponential after shifting)

$$N_{t+1} = N_t \lambda^t, \quad N_t = N_0 \lambda^t \quad (1)$$

where N is the number of individuals present in the population, and t is a time interval of interest. This equation says that the size of a population at time $t + 1$ is equal to the size of the population at time t multiplied by a constant, λ . When $\lambda = 1$, the population will remain constant in size over time. When $\lambda < 1$, the population declines geometrically, and when $\lambda > 1$, the population increases geometrically. Although geometric growth models have been used to describe population growth, like all models they come with a set of assumptions.

5.3.1 Calculating Lambda using the Leslie Matrix Model

The major goal of the matrix model is to compute λ , the finite rate of increase in Equation 1, for a population with age structure. In our matrix model, we can compute the time-specific growth rate as λ_t . The value of λ_t can be computed as:

$$\lambda_t = N_{t+1}/N_t \quad (2)$$

This time-specific growth rate is not necessarily the same λ in Equation 1. To determine N_t and N_{t+1} , we need to count individuals at some standardized time period over time. We will make two assumptions in our computations. First, we will assume that the time step between N_t and N_{t+1} is one year, and that age classes are defined by yearly intervals. This should be easy to grasp, since humans typically measure time in years and celebrate birthdays annually. Second, we will assume for this exercise that our population censuses are completed once a year, immediately after individuals breed (a post-breeding census). The number of individuals in the population in a census at time $t + 1$ will depend on how many individuals of each age class were in the population at time t , as well as the birth and survival probabilities for each age class.

6 SIR Model

The SIR model is a widely used mathematical model in epidemiology that describes the spread of infectious diseases within a population. It divides the population into three groups: Susceptible(S), Infected (I), and Recovered (R)(or Immune).

The model can be divided into 3 categories:

- **Static or Stochastic SIR Model:** The stochastic SIR model incorporates randomness in the transmission and recovery processes of the individuals in the population, allowing for probabilistic outcomes. It accounts for the inherent variability and unpredictability in the spread of infectious diseases by considering random events and individual-level interactions.
- **Linear SIR Model:** The linear SIR model assumes a linear relationship between the rates of infection, recovery, and susceptible individuals. It is a simplified representation of disease dynamics, where the rates remain constant over time. This model does not capture complex interactions or variations in the population and is typically used for theoretical analysis or as a baseline for more sophisticated models.
- **Non-Linear SIR Model:** The non-linear SIR model introduces non-linear functions to represent the rates of infection, recovery, and susceptible individuals. This model allows for more realistic and flexible representations of disease dynamics, considering factors such as population size, limited resources, and varying contact rates.

For purposes of this project we will use Stochastic SIR Model for all further explanation since Linear Algebra plays a significant role here.

The model assumes that individuals can transition between these groups over time and that the population remains constant. To analyze the dynamics of the SIR model using linear algebra, we can represent the population in terms of vectors and matrices. Let's denote the population at a given time as a column vector:

$$T = \begin{pmatrix} S \\ I \\ R \end{pmatrix}$$

where S, I, and R represent the number of susceptible, infected, and recovered individuals, respectively.

Next, we define a matrix, usually referred to as the transition matrix, which captures the transition rates between the different groups. The transition matrix is typically denoted by T and has the form:

$$T = \begin{pmatrix} S \rightarrow S & S \rightarrow I & S \rightarrow R \\ I \rightarrow S & I \rightarrow I & I \rightarrow R \\ R \rightarrow S & R \rightarrow I & R \rightarrow R \end{pmatrix}$$

where the entries represent the transition rates from one group to another.

For example, $S \rightarrow I$ represents the rate at which susceptible individuals become infected, $I \rightarrow R$ represents the rate at which infected individuals recover, and so on.

So, if S_n, I_n, R_n represents the number of susceptible people, infected people, recovered or immune people at the end of n weeks, we can say ,

$$\begin{pmatrix} S_{n+1} \\ I_{n+1} \\ R_{n+1} \end{pmatrix} = T \cdot \begin{pmatrix} S_n \\ I_n \\ R_n \end{pmatrix}$$

Studying the eigen vector of the largest eigen value T^T matrix gives us valuable information about the virus. To check the real world application of the above formula and how accurate and precise it is, we must compare its results by running a simulation.

6.1 Simulation of Virus Spread

6.1.1 Markov Model

Microsimulation models have a clear advantage when considering infectious diseases and vaccination studies. The likelihood of contracting an infection depends on the number of people already infected.

However, Markov models depict health state progression over multiple time cycles with fixed health state transition probabilities. These probabilities are the elements of the stochastic transition matrix T .

Advantages of using Markov model are:

- Markov models are relatively simple to understand and implement. They provide a clear and intuitive framework for modeling systems with probabilistic transitions between states.
- The Markov property, which assumes that the future state depends only on the current state and is independent of the past history, can be advantageous in certain situations. It simplifies the modeling process by allowing focus on the immediate preceding state rather than considering the entire history. This property makes Markov models computationally efficient and particularly useful when memory of distant past events is not necessary or relevant.
- Markov models offer interpretability, allowing users to understand and interpret the underlying dynamics of a system. By examining the transition probabilities between states, one can gain insights into the behavior and dependencies within the system. Thus they are easily interpretable. This interpretability can be valuable for decision-making, process optimization, and understanding complex systems.

However, there are some disadvantages too:

- Markov models are relatively simple to understand and implement. They provide a clear and intuitive framework for modeling systems with probabilistic transitions between states
- The assumption of the Markov property, which states that the future state of the system depends only on its current state and is independent of its past history. This assumption can be limiting in certain scenarios where the system's future behavior is influenced by events or factors that occurred before the current state. In such cases, the Markov model may not accurately capture the dependencies and dynamics of the system .
- Assumption of stationary behavior, meaning that the transition probabilities between states remain constant over time. However, in many real-world scenarios, the underlying system may exhibit non-stationary behavior, where the transition probabilities change over time due to various factors or external influences. In such cases, the Markov model may fail to adapt to these changes, leading to inaccurate predictions.

6.1.2 Monte Carlo simulation

As a control setup to compare the Markov model to, Monte Carlo simulation has been chosen. It depicts patient health state progression through computer simulation of multiple individuals with random pathways based on model probabilities. It involves the following steps:

- Define the Model: Specify the mathematical model representing the system of interest, including its parameters and relationships.
- Generate Random Inputs: Randomly sample input values for the model's parameters based on their probability distributions. These inputs represent the uncertainty or variability in the system.
- Run Simulations: Execute the model using the randomly sampled input values and observe the resulting output. Repeat this process multiple times to obtain a range of outcomes.

It is an extremely flexible and realistic model and is computationally complex. To take same parameters used in Markov model, certain parameters from the model have been removed :

- Death due to old age and infection
- Birth rate : Since population has been assumed to be constant
- Lifetime immunity: Once an individual is recovered , he becomes immune and that immunity to the virus is never lost.

6.2 Comparing results

Following will be the details of the simulation:

- Probability of infection: 65%
- Probability of recovery due to immunity: 25%
- Initial fraction of individuals infected: 0.5
- Initial number of individuals immune: 0
- Immunity to the infection can be gained from the first week itself
- Once immune, although an individual is healthy, they do not contribute to the Healthy population

First, we will make the transition matrix using the above data:

$$T = \begin{pmatrix} 0.35 & 0.65 & 0 \\ 0 & 0.75 & 0.25 \\ 0 & 0 & 1 \end{pmatrix}$$

Largest eigen value, and corresponding eigen vector:

$$\lambda = 1, \quad \nu = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (3)$$

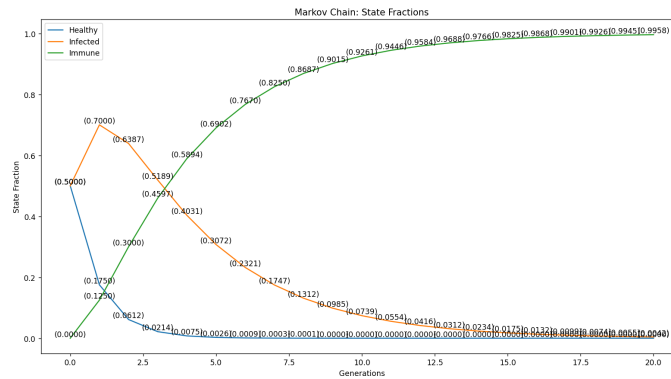
Thus, the model predicts that at after sufficiently long time , all individuals will become immune.

Population size 1000:

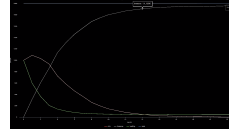
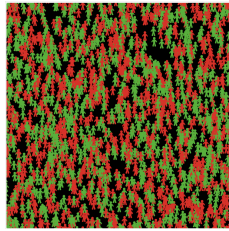


We can see there is a significant discrepancy between the realistic model's simulated results and the predicted data obtained from Markov. Reason for this Markov model assumes that the system's behavior can be adequately represented by a fixed set of transition probabilities. However, in reality, small populations often exhibit more variability and randomness due to limited sample size. Smaller sample size also leads to higher uncertainty in data. Thus we conclude that the above Markov model is not reliable for calculating virus outbreaks in localities, where number of individuals is not high.

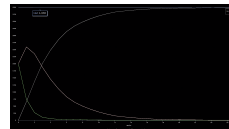
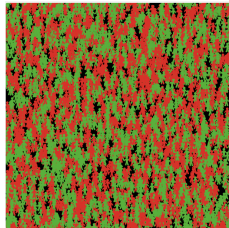
First, we see the predicted output by the Markov model.



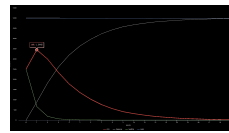
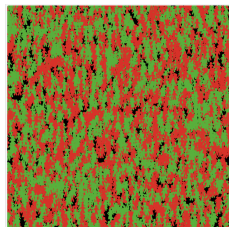
Population size 2000:



Population size 4000:



Population size 5000:



Following observations can be made:

- In the second week of the simulation, with a larger population of 5000, approximately 70.2% of the population was infected. This indicates a higher rate of transmission compared to the previous week, potentially due to factors such as increased contact or a higher initial number of infected individuals. However, by the second week, the fraction of infected individuals decreased to 60.3%. This decline suggests that some individuals are recovering from the infection or becoming immune, leading to a reduction in the overall number of infected individuals.
- The spread of the virus is initially high but gradually decreases over time. The decline in the fraction of infected individuals indicates that some individuals are recovering or becoming immune to the virus.
- Within 20-25 weeks, almost the entire population has become immune to the virus.

Conclusions:

- Initial peak in infected people occurs since immunity has not developed in the first week.

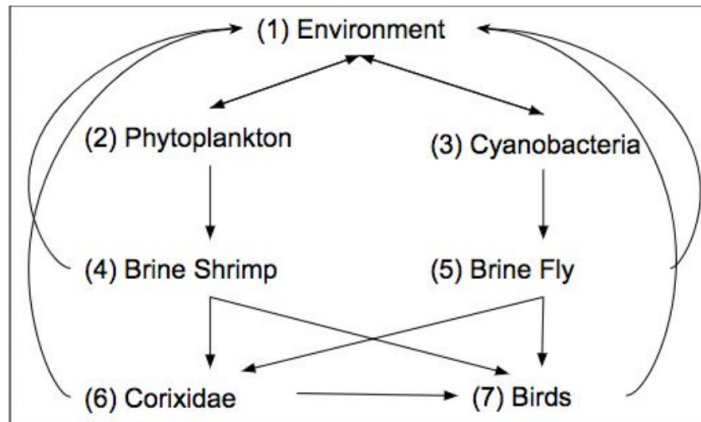
- Our initial prediction about fraction number of individuals with immunity being 1 was correct.
- Increasing population size helps in better corroboration of the results with the predicted outcome.

7 Population Ecology Network: Food Web

Ecosystem analysis is a fundamental field in understanding the intricate interdependencies and stability of natural systems. It investigates the potential of linear algebra as a powerful tool for unraveling complex ecological networks by leveraging the concept of food webs and adjacency matrices. By constructing adjacency matrices from food webs, various quantitative characteristics of ecosystems can be mathematically derived, enabling in-depth analysis of species interactions, potential food chains, and structural diversity.

7.1 Analysis of a Ecosystem using Adjacency Matrix

To illustrate the application of linear algebra in ecological analysis, let's consider a simplified food web of the Great Salt Lake ecosystem. In this ecosystem, different species interact through complex feeding relationships. For instance, phytoplankton serves as a food source for brine shrimp, while brine flies depend on cyanobacteria for their sustenance. Corixidae, in turn, prey on both brine shrimp and brine flies. Birds, occupying the highest trophic level, consume multiple species including brine shrimp, brine flies, and corixidae. It is important to note that the environment itself plays a significant role in the ecosystem, providing resources such as dead matter, nutrients in the water, and sunlight. The environment, in a sense, "feeds" phytoplankton and cyanobacteria, and is "fed" by the death of every species within the system.



To analyze the food web, we create adjacency matrices that capture the relationships between species. First, we construct matrix A_1 , which represents a matrix indicating

the connections between species. Each element in A_1 , denoted by $A_1[i, j]$, represents one-way predation relationship between species i and j . In A_1 , a value of 1 in the $A_1[i, j]$ position indicates that species i eats species j . While A_1 has 1s in every position, matrix A_2 incorporates variability by assigning different values to indicate differences in the relative quantities of consumed species. In this analysis, both A_2 and A_1 are used for various purposes, enabling a comprehensive examination of the ecosystem.

$$A_1 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 10 & 10 & 1 & 0 \end{bmatrix}$$

7.1.1 Calculation of Number of Direct Food Sources

To determine which species in the ecosystem has the most direct sources of food, we can examine the row sums of matrix A_1 . The row sums represent the total number of feeding interactions each species has with other species in the ecosystem. This information is crucial for understanding the flow of energy and resources within the food web.

Computing the row sums of matrix A_1 can be done through matrix multiplication A_1 with the column vector of ones v_1 . $A_1 v_1$ gives a column vector whose elements represent the sum of each row in A_1 . Each element in the resulting vector corresponds to a species in the ecosystem, and a higher value indicates a greater number of direct food sources for that species.

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad A_1 v_1 = \begin{bmatrix} 6 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$

This shows that Environment has 6 direct sources, Phytoplankton has 1, Cyanobacteria has 1, Brine Shrimp has 1, Brine Fly has 1, Corixidae has 2 and Birds have 2 direct sources of food.

Similarly, to determine the number of times each organism serves as a direct food source for another, we can calculate the column sums of matrix A_1 by multiplying matrix A_1 with row matrix v_2 . This analysis allows us to identify the species that

occupy pivotal roles in the food web, as they serve as primary food sources for multiple other species. Understanding the distribution of feeding interactions and identifying key species in the ecosystem contributes to our knowledge of species interdependencies and the overall stability of the ecological system.

$$v_2 = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1] \quad v_2 A_1 = [2 \quad 2 \quad 2 \quad 3 \quad 3 \quad 2 \quad 1]$$

7.1.2 Calculation of Number of Indirect Food Sources

In addition to direct food sources, organisms in our food web also have indirect sources of food. By utilizing the properties of adjacency matrices, we can determine the count of indirect food sources for each organism.

The entry a_{ij} of A_1^2 represents the number of 2-paths, indicating the number of indirect food sources of length 2, between the organism at vertex i and the organism at vertex j .

Similar to our previous approach, we can calculate the number of indirect food sources for each organism by summing the rows of matrix A_1^2 . We can do so by multiplying the matrix A_1^2 with v_1 . By combining the counts of direct food sources and indirect food sources, we obtain the total number of direct and indirect food sources for each organism.

$$A_1^2 = \begin{bmatrix} 2 & 1 & 1 & 2 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \quad A_1^2 v_1 = \begin{bmatrix} 9 \\ 6 \\ 6 \\ 1 \\ 1 \\ 2 \\ 4 \end{bmatrix} \quad A_1 v_1 + A_1^2 v_1 = \begin{bmatrix} 15 \\ 7 \\ 7 \\ 2 \\ 2 \\ 4 \\ 7 \end{bmatrix}$$

So, total number of direct and indirect food sources for Environment is 15, for Phytoplankton it is 7, for Cyanobacteria it is 7, for Brine Shrimp it is 2, for Brine Fly it is 2, for Corixidae it is 4 and for Birds it is 7.

7.1.3 Potential Food Chains

One fascinating aspect of studying food webs is understanding the potential number of food chains within an ecosystem. By employing matrix multiplication, we can calculate this quantity. To determine the number of food chains, we raise matrix A_1 to the power of the number of trophic levels, which in this case is 5. This process accounts for the propagation of energy and feeding relationships across multiple trophic levels. Next, we focus on the first column of the resulting matrix, considering the last n entries, where n represents the number of species in the last trophic level, which is 1 in this case. By

summing these entries, we arrive at the total number of potential food chains within the ecosystem. A_1^5 obtained by multiplying A_1 by itself 5 times:

$$A_1^5 = \begin{bmatrix} 10 & 14 & 14 & 14 & 14 & 10 & 8 \\ 8 & 5 & 5 & 6 & 6 & 4 & 2 \\ 8 & 5 & 5 & 6 & 6 & 4 & 2 \\ 2 & 4 & 4 & 3 & 3 & 2 & 2 \\ 2 & 4 & 4 & 3 & 3 & 2 & 2 \\ 4 & 2 & 2 & 4 & 4 & 2 & 0 \\ 4 & 4 & 4 & 6 & 6 & 4 & 2 \end{bmatrix}$$

As the last entry in first column of matrix A_1^5 is 4, therefore, 4 food chains from the food web is possible from environment to birds.

7.1.4 Structural Diversity

In addition to understanding the number of food chains, it is important to assess the structural diversity of an ecosystem. Structural diversity indicates the robustness and stability of the system, reflecting the ability of the ecosystem to sustain itself in the face of perturbations or species loss. To quantify structural diversity, we turn to the Perron-Frobenius theorem, which highlights the dominant eigenvalue of a matrix. By calculating the dominant eigenvalue, which represents the largest absolute value among all eigenvalues of the matrix, we gain insight into the ecosystem's structural stability. According to the theorem, as the values within the matrix increase, the dominant eigenvalue also increases. This holds true for ecological systems as well, as species with increased feeding opportunities contribute to a more structurally sound ecosystem.

Considering matrix A_1 from the previous analysis, we can calculate its dominant eigenvalue using python script. The dominant eigenvalue of A_1 provides a measure of the ecosystem's structural diversity. In this case, the dominant eigenvalue calculated is found to be 2.043.

$$\text{Eigenvalues of } A_1 = \begin{bmatrix} 2.043 \\ -3.913 \times 10^{-2} + 1.099i \\ -3.913 \times 10^{-2} - 1.099i \\ -1.375 \\ -5.890 \times 10^{-1} \\ 8.668 \times 10^{-9} \\ -8.668 \times 10^{-9} \end{bmatrix}$$

To further explore the concept of structural diversity, let's conduct a hypothetical scenario. Suppose we modify the food web by restricting birds to consume only brine shrimp. We can update matrix A_1 by replacing the last row of A_1 with a new row vector to form a new matrix A_3 , which represents this change. In this case, the vector

would be $\langle 0, 0, 0, 1, 0, 0, 0 \rangle$, indicating that only species 4 (brine shrimp) is consumed by birds.

$$A_3 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

By calculating the dominant eigenvalue of A_3 , we observe that it decreases to 1.952. This decrease in the dominant eigenvalue reflects the reduction in structural diversity caused by the restriction of bird diet. It demonstrates that when species have limited feeding options, the ecosystem's overall robustness decreases. The eigenvalues of matrix A_3 are presented below:

$$\text{Eigenvalues of } A_3 = \begin{bmatrix} 1.952 \\ -1.436 \\ -2.582 \times 10^{-1} + 1.002i \\ -2.582 \times 10^{-1} - 1.002i \\ 1.978 \times 10^{-6} + 3.425 \times 10^{-6}i \\ 1.978 \times 10^{-6} - 3.425 \times 10^{-6}i \\ -3.955 \times 10^{-6} \end{bmatrix}$$

The transformation of A_1 into A_2 by increasing certain values should be reflected in the dominant eigenvalue of A_1 . As previously mentioned, the dominant eigenvalue of A_1 is 2.043. The increase in these values results in the dominant eigenvalue of matrix A_2 becoming 3.641, indicating a greater level of structural diversity. The eigenvalues of matrix A_2 are presented below:

$$\text{Eigenvalues of } A_2 = \begin{bmatrix} 3.641 \\ -3.123 \\ -0.213 + 3.091j \\ -0.213 - 3.091j \\ -0.092 \\ -1.542 \times 10^{-17} + 1.539 \times 10^{-8}j \\ -1.542 \times 10^{-17} - 1.539 \times 10^{-8}j \end{bmatrix}$$

7.2 Limitations and Potential for Further Analysis

7.2.1 Limitations

It is important to acknowledge the limitations of this analysis, particularly in relation to the simplifications made. The Great Salt Lake ecosystem analysis presented here is based on simplified data constraints. The relative quantities assigned to the matrix

elements (e.g., 5 and 10) do not accurately represent the actual consumption rates between species, which can vary significantly. Additionally, the top trophic level is represented as "Birds" without considering the diverse bird species present in reality. A more detailed and accurate analysis could be performed by accounting for different bird species and their specific interactions within the ecosystem.

7.2.2 Determining Importance of Species in Coextinction

Forecasting the consequences of species' extinctions is a critical challenge in ecology. The complex network of ecological relationships among species makes it difficult to accurately predict the secondary extinctions resulting from species losses. However, by adapting the algorithm used by Google to rank web pages, it becomes possible to determine the importance of each species for coextinctions (the extinction of one species could cause the elimination of another, which in turn would cause the loss of a third species). This algorithm provides a sequence of losses that leads to the fastest collapse of the network, allowing us to identify the most critical species. It comes that the species that the greatest number of other species rely on for food are the ones that are most essential to the health of an ecosystem. Additionally, this approach bridges the gap between qualitative and quantitative descriptions of food webs by considering both the "who eats whom" aspect and the rate of interactions. The algorithm demonstrates its effectiveness in identifying the most efficient extinction sequence for secondary extinctions in all analyzed networks. It serves as a foundation for further advancements in analyzing extinction risk in ecosystems.

8 State of the Art

8.1 Matrix Representation of Population Models:

- **Advanced Matrix Structures** Researchers are exploring more sophisticated matrix structures to capture complex population dynamics. This includes the use of structured matrices that incorporate additional factors such as spatial distribution, social interactions, and environmental variables. By expanding the dimensions of the matrices or introducing additional dimensions, these models can provide more detailed and realistic representations of populations.
- **Stochastic Models** Traditional matrix models assume deterministic transitions between population states. However, recent advancements include the integration of stochasticity into matrix models. Stochastic matrix models consider random variations in transition rates, allowing for a more accurate representation of population dynamics and the incorporation of uncertainty into predictions.

8.2 Leslie Matrix Model:

- **Extensions to Age-Structured Models** The Leslie matrix model traditionally focuses on age-structured populations, but recent advancements have expanded its application to include other types of structured populations. For instance, stage-structured models capture distinct life stages (e.g., larval, juvenile, adult), allowing for a more comprehensive representation of population dynamics.
- **Sensitivity and Elasticity Analysis** Sensitivity and elasticity analyses are employed to identify key parameters or life stages that have the most substantial influence on population growth and stability. By evaluating the sensitivity of population growth rate or other population metrics to changes in matrix entries, researchers can prioritize conservation efforts or interventions targeting specific life stages.
- **Nonlinear and Complex Dynamics** Leslie matrix models traditionally assume linear relationships between population states. However, recent advancements aim to incorporate nonlinear dynamics and complex feedback mechanisms into the models. This includes the exploration of chaotic behavior, tipping points, and alternative stable states, enhancing the understanding of population dynamics under more realistic and intricate scenarios.

8.3 SIR Model

- **Incorporating Network Structure** Traditional SIR models assume a well-mixed population, where every individual has an equal chance of interacting with any other. However, real-world interactions are often structured as networks. Recent research focuses on incorporating network structures into the SIR model, considering factors such as social connections, transportation networks, or contact networks.
- **Heterogeneous Mixing** The standard SIR model assumes homogeneous mixing, implying that individuals have equal contact rates with others. However, in reality, contact rates can vary significantly due to factors like age, occupation, or location. Researchers have developed SIR models that account for heterogeneity in contact rates, allowing a more accurate representation of how diseases spread within different subgroups of the population.

8.4 Population Ecology Food Web

- **Metapopulation Models** Metapopulation models focus on the dynamics of interconnected subpopulations within a larger landscape. They consider factors like migration, colonization, and extinction of populations in fragmented habitats. SOTA approaches in metapopulation ecology involve incorporating spatial heterogeneity, landscape connectivity, and the effects of environmental changes to understand how metapopulations persist and interact.

- Ecological Network Analysis Ecological network analysis provides quantitative metrics to understand the structure, complexity, and functioning of food webs. SOTA approaches in food web ecology leverage network theory and advanced statistical techniques to analyze large-scale food web datasets.

9 End

In conclusion, matrix models have advanced our understanding of population growth, guiding conservation and management efforts. Integrating ecological concepts and refining methodologies will enable accurate predictions and promote the sustainability of ecosystems in the face of environmental changes.

The Code and Images used are uploaded on the Git Repo

10 Contribution

1. Aryaman Bahl

- Interim report in LaTeX.
- Introduction to Leslie Matrix and Markov chain.
- Python code implementation for the same.

2. Sujal Keshri

- Studying population ecology using linear algebra.
- Network analysis of ecological networks using adjacency matrix.
- Python code implementation for the same.

3. Siddharth Agarwal

- Studying and implementing SIR model for application in epidemiology field.
- Modifying existing model codes to meet ideal parameters.
- Data analysis and limitation of Markov chain in SIR model.