Introduction to mathematical biology and application in R

Koissi Savi

January 1, 2021

zef
Center for
Development Research
University of Bonn

Ministerium für Innovation,
Wissenschaft und Forschung
des Landes Nordrhein-Westfalen

UNIVERSITÄT BONN

Outline

Goals of the seminar-lecture

Although the material was developed for a fairly wide audience, researchers with prerequisite in basic calculus, linear algebra and ordinary differential equations will find the material much more accessible.

- Hope to familiarize researchers with mathematical methods they can customize and use for their own research.
- Get inspired by previous models to develop a model that addresses your own research.

Outline

History of mathematical biology

- **Leonardo Bonacci** [Fibonacci] created a series denoted **Fibonacci series** (each number is the sum of the two preceding ones) to explain the dynamics of rabbits around the 13th century

- **Daniel Bernouilli** in the 18th century applied a system in two ordinary differential equations (ODEs) to describe the dynamics of smallpox on the human population and defined how vaccine could be used to prevent the spread of this disease.

- **Thomas Malthus** in 1798 theorized on the **exponential growth** of human population.

- **Pierre François Verhulst** in 1836 formulated the **logistic growth** of the population

- **Fritz Müller** in 1878 theorized on the mimicry of species that share a common predators. The theory was first described based on species coloration eg: coral snake vs scarlet kingsnake.

Motivation of the seminar-lecture

Some reasons motivate this seminar lecture,

- As a researcher (biologist/ ecologist) mathematics help to design your experiment or clinical trials so that your results will be meaningful (statistical significant) statically
- Sometimes it is impossible or too difficult to answer a research question through traditional lab experiments, so biologists rely on math to develop models that represent the system they're studying. These approaches allow scientists to indicate the likelihood of certain outcomes as well as refine the research questions in an ethical manner.
- Mathematical modeling helps biologist to identify pattern in their studies.

Society for Mathematical Biology

CrossMark

What Has Mathematics Done for Biology?

Michael C. Mackey · Philip K. Maini

Outline

Chap I: Dynamics of biological populations

Goal

In this first chapter we will review the classical models of dynamics of the population of a single species.

Exponential growth

Definition

This growth occurs when the instantaneous rate of change is proportional to the quantity itself. It is also called **geometric growth** or **Malthusian growth model**.

- This growth is observed for a colony of bacterial, of mitosis-borne cells.
- A virus (for example COVID-19, or smallpox) typically will spread exponentially at first, if no artificial immunization is available.
- In finance the Pyramid schemes or Ponzi schemes also show this type of growth resulting in high profits for a few initial investors and losses for large numbers of investors.

Notation and computation

Let N be the number of individuals in a population at time t,
b the average per capita birth rate and
d the average per capita death rate
After a short period of time dt, the fraction of change in the population N is given by:

$$\frac{dN}{dt} = bN - dN \tag{1}$$

$$(1) \leftrightarrow \frac{dN}{dt} = (b-d)N$$

$$\frac{dN}{N} = (b-d)dt$$

$$\int_0^\infty \frac{dN}{N} = \int_0^\infty (b-d)dt$$

$$N = N_0 e^{(b-d)t} \tag{2}$$

Assuming that $b - d = r$ with r representing the **per capita growth**;
$(2) \leftrightarrow N = N_0 e^{rt}$

Example of the dynamics of rabbit populations

Let's assume that we have a population of 1000 rabbit living in an unlimited space and with unlimited resources. Besides, let's consider that the monthly growth rate of the rabbit is 10%. How will this population grow?

Months / time	Population	Growth factor
t	N_0	r
0	1000	
1	1000	1.1
2	1000	1.1×1.1
3	1000	$(1.1)^3$
$\vdots$	$\vdots$	$\vdots$
t	1000	$(1.1)^t$

Thus, the population of rabbit is given by $N(t) = 1000 \times (1.1)^t$.
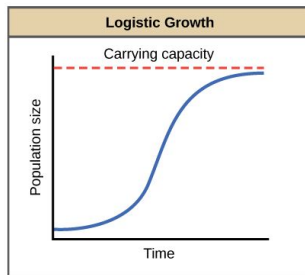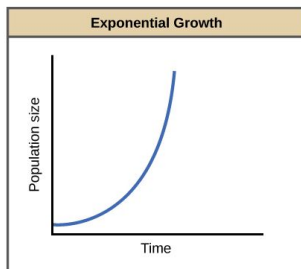The population after 10 years (120 months) is given by
$N(120) = 1000 \times (1.1)^{120} == 93,000,000$ rabbits

Logistic growth

Definition

Contrary to the exponential growth that has a constant **per capita growth**, in **logistic growth** the per capita growth rate gets smaller as the population size approaches its maximum imposed by limited resource also called **carrying capacity** denoted K.

- An exponential growth has a **J-shaped curve** while a logistic growth has an **s-shaped curve**.

Image credit: "Environmental limits to population growth: Figure 1." by   Image credit: "Environmental limits to population growth: Figure 1." by

Notation and computation

Let us modify the premises of the equation (1).

Using an r (per capita growth rate) that depends on population size N and how close it is to carrying capacity K.

Assuming that the population has a base growth rate of r when it is very small, we can write the following equation:

$$\frac{dN}{dt} = r\frac{(K - N)}{K}N \qquad (3)$$

$$\frac{dN}{dt} = rN(1 - \frac{N}{K})$$

We can observe that this equation has two fixed points.

1. when $\lim_{N \to 0} \frac{dN}{dt} = rN ==>$ Exponential growth
2. when $\lim_{N \to K} \frac{dN}{dt} = 0 ==>$ The slope of the curve is zero indicating an asymptote in K.

The general form of the solution to this ode is given by:

$$N = \frac{K}{1 + be^{-rt}} \qquad (4)$$

With b be a constant

Exercise

Demonstrate that the general form of a logistic growth model is given by the Eq (4)

$$\frac{dN}{dt} = r\frac{(K-N)}{K}N \rightarrow N = \frac{K}{1 + be^{-rt}}$$

Solution

$$\frac{dN}{dt} = r\frac{(K-N)}{K}N \rightarrow N = \frac{K}{1 + be^{-rt}}$$

Let's now integrate both sides of the above-mentioned equation

$$\int \frac{dN}{N(1 - \frac{N}{K})} = \int rdt$$

Let's apply the partial fraction decomposition property on the left hand side of the previous equation. We obtain

$$\int \frac{1}{N} + \frac{1}{K-N}dN = \int rdt$$

Solution

We can apply the additive property of integral and it is easy to see that:

$$\int \frac{1}{N} + \frac{1}{K-N} dN = \int r dt \rightarrow ln(N) - ln(K-N) = rt + C$$

With C being a constant. Using the natural log property we have

$$ln(N) - ln(K-N) = rt + C \leftrightarrow ln(\frac{N}{K-N}) = rt + C$$

We can then exponentiate the previous expression and as C is a constant let denoted the e^{-c} b. Thus, it is easy to see that

$$\frac{dN}{dt} = r\frac{(K-N)}{K}N \leftrightarrow N = \frac{K}{1 + be^{-rt}}$$

At $t = 0$ it is easy to see that $b = \frac{K-N_0}{N_0}$ with N_0 being the initial population. Thus,

$$N = \frac{N_0 K}{N_0 + (K-N_0)e^{-rt}}$$

Example of cows - Problem

100 cows were released on an island in 2005. By 2012, there were 324 cows. The island has a big pasture land that can carry a maximum of 5000 cows.

1. Assuming logistic growth, write the general equation that describes the population $N(t)$ at any time t?

2. How many cows will be there by 2020?

3. In what year will the population reach 1000?

Example of cows - Solution

1- Equation describing the population

$N(t) = \frac{K}{1+be^{-rt}}$ In 2005 $t = 0 \rightarrow N(0) = 100$

In 2012 $t = 7 \rightarrow N(7) = 324$

The carrying capacity is $K = 5000$ $100 = \frac{5000}{1+be^{-r(0)}} \rightarrow b = 49$

$324 = \frac{5000}{1+49e^{-r(7)}} \rightarrow r = 0.1746$

$N(t) = \frac{5000}{1+49e^{-0.1746(t)}}$

2- Number of cows in 2020 (t=15)

1094 cows

3- Determination of t

$t = 14.35 \rightarrow 2019$

r and K-selected species

- Like the elephant, **K-selected** species have a logistic growth.

- They are defined by fewer offspring and they provide a long-term care after birth

- **r-selected** species like jellyfish have an exponential growth

- They are defined by large numbers of offspring and short to no care after birth

- More debate in this category: The oak for instance has a large number of offspring but because a lot also survive they are defined as a K-selected species.

Summary

1. The dynamics of a given population could be modeled mathematically using differential equations

2. In absence of competition or self-limiting factors (or intra-specific competition), species dynamics could be modeled using an exponential growth model

3. However, when the resource is limited or when there is intra-specific competition, the dynamics is modeled using a logistic growth model

4. Thus, what will happen if there is a competition between species?

Outline

Chap II: Modeling interactions between species: Lotka-Volterra model

In this section we will explore the dynamics of biological population when we have two or more species with respect to the interactions between them.

Historical background

- In 1910, Alfred J. Lotka proposed a set system of ODEs in the theory of autocatalytic. This theory postulated that from a hypothetical chemical reaction, this reaction could exhibit periodic behaviour.
- Then, in 1920 and 1925 he extended his equation to explain the dynamics of plant species and herbivores and then to the predator-prey dynamics.
- In 1926, an independent work of Vito Volterra, a mathematician who became interested in biology, published the same set of equations to explain the dynamics of fish catched in Adriatic Sea where he noticed an asynchronism in the density of the population of the catches fish (prey) and their predators.

Prey-predator

Assumptions

- The prey have unlimited resources and in absence of predator grow exponentially
- The predator depends on the presence of its prey to survive.
- Births and deaths are the only way the populations of both the prey and the predator can increase and decrease
- There is no migration to or from the populations
- The only way of prey dies is by consumption by predators
- No other preys and predators are incorporated into this model

Notation and the model

- Let N be the population of prey; P be the population of predator ; a the capture rate of the predator; b the expected increase in the number of predator and m the mortality rate of the predator; and r the per capita growth the of the prey
- Initially we assumed that N grows exponentially. Thus the population of prey is given by .

$$\frac{dN}{dt} = rN - aNP \qquad (5)$$

- aNP represents the number of prey consumed by the predator. We need to incorporate N cause the likelihood of a predator to encounter a prey depends on the density of the prey.
- The population of the predator is given by

$$\frac{dP}{dt} = baNP - mP \qquad (6)$$

- $baNP$ represents the transformation of the predator consumption into offspring and mP represents the natural death that occurred in the population of predator

Solution to the system of ODEs

The equations (5) and (6) define a solvable system of ODEs.

- It is easy to see that there are two states of the system of ODEs that does not depend on time also called **steady state**.
- There are $(0,0)$ and $(\frac{m}{ab}, \frac{r}{a})$. This last one can be proved while posing $(5) = 0$ and $(6) = 0$ $[\frac{dN}{dt} = rN - aNP = 0$ & $\frac{dP}{dt} = baNP - mP = 0]$
- The Jacobian of the system is given by

$$J = \begin{pmatrix} r - aP & -aN \\ baP & baN - m \end{pmatrix}$$

- For the steady state 1

$$J = \begin{pmatrix} r & 0 \\ 0 & -m \end{pmatrix}$$

- For the steady state 2

$$J = \begin{pmatrix} 0 & \frac{-m}{b} \\ br & 0 \end{pmatrix}$$

Solution to the system of ODEs

We can derive from the steady state 1 two eigenvalue

$$\lambda_1 = r \quad \& \quad \lambda_2 = -m$$

And from the steady steady state 2 the eigenvalue is given by

$$\lambda_{1,2} = \pm\sqrt{rmi}$$

Meaning of the solution

- First, we can notice that the steady-state 2 the population of prey is independent of its own growth rate or mortality ($N = \frac{m}{ab}$); rather, it depends on parameters associated with the predator.

- A similar result holds for steady-state levels of the predator ($P = \frac{r}{a}$)

- The presence of predator ($P \neq 0$) means that the available prey has to just suffice to make growth rate due to predation, dN, equal predator mortality m for a steady predator population to persist.

Meaning of the solution

- Similarly, when prey are present ($N \neq 0$), predators can only keep them under control when prey growth rate r and mortality a due to predation, by, are equal. This helps us to understand the steady-state equations.

- Second, the steady state is neutrally stable. The eigenvalues are pure imaginary with a non real part and predict oscillations around the steady state.
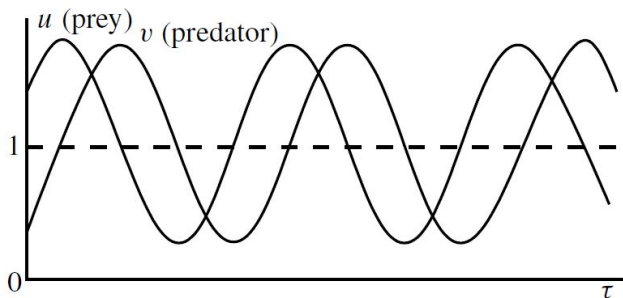
Meaning of the solution

- If $N(0) = 0$ but $P(0) > 0$, we see that an explicit solution of the system is $N(t) = 0, P(t) = P_0 e^{-mt} ==>$ The predator population will decrease at an exponential rate

- Similarly if $N(0) > 0$ but $P(0) = 0$ the solution $N(t) = N_0 e^{rt}$ goes to infinity along the line $P = 0$ exponentially as $t \to \infty$. $==>$ Exponential growth of the prey.

- If $N(0) \geq 0, P(0) \geq 0$ then $N(t) \geq 0, P(t) \geq 0$ for all $t \geq 0$. $==>$ The solution here could be observed from the phase plane of the model

Example

Canadian lynx–snowshoe hare interaction in the fur catch records of the Hudson Bay Company from about 1845 until the 1930's. We assume that the numbers reflect a fixed proportion of the total population of these animals. Although this assumption is of questionable accuracy, as indicated by what follows, the data nevertheless represent one of the very few long term records available.
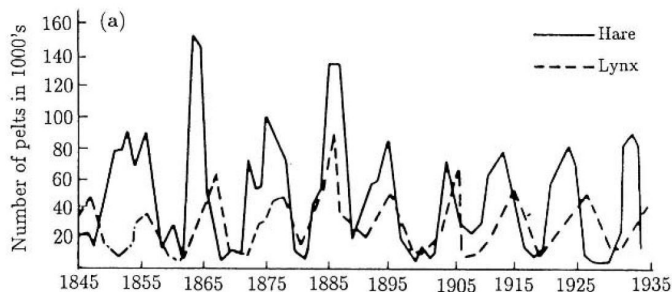
Competition

Recall that for a single population of density N, the *Logistic equation* is read

$$\frac{dN}{dt} = rN(1 - \frac{N}{K})$$

and the explicit solution

$$N = \frac{K}{1 + be^{-rt}} \rightleftharpoons N = \frac{N_0 K}{N_0 + (K - N_0)e^{-rt}}$$

- With this explicit expression for the density, it is easy to see that $N(t) \rightarrow K$ as $t \rightarrow \infty$ for all $N_0 > 0$.

- Note that as $t \rightarrow \infty$ we have $N(t) \rightarrow K$. The density K is the maximum population density that the environment can carry and is called the *environmental carrying capacity*.

- The quadratic term in the first equation (i.e. $1 - \frac{N}{K}$) represents competition between members of the same population for resources, i.e. **intraspecific competition.**

Formulation of the model

To build a simple model for such competition we start with two independent logistic models for the population densities N_1, N_2 and add an extra term in each to model the **interspecific competition**:

$$\frac{dN_1}{dt} = r_1 N_1 (1 - \frac{N_1}{K_1} - c_1 N_2)$$

$$\frac{dN_2}{dt} = r_2 N_2 (1 - \frac{N_2}{K_2} - c_2 N_1)$$

(7)

- Note that in the absence of interspecific competition, each species grows to its respective carrying capacity.
- The relative sizes of $c1, c2 > 0$ determine the competitiveness of each species.

Solution

We note that, employing the solution to the single-species logistic equation,

$$N_1(t) = 0, N_2(t) = \frac{N_{20}}{\frac{N_{20}}{K_2} + (1 - \frac{N_{20}}{K_2})e^{-r_2 t}}$$

is the solution to (7), with $N_1(0) = N_{10} = 0, N_2(0) = N_{20} > 0$.

- Thus, all initial states $(0, N_{20})$ with $N_{20} > 0$ go exponentially to $(0, K_2)$. Similarly, all states $(N_{10}, 0)$ with $N_{10} > 0$ end up at $(K_1, 0)$
- To ease calculations, we first set $u_i = \frac{N_i}{K_i}$ for $i = 1, 2$, $a_{12} = c_1 K_2, a_{21} = c_2 K_1$.
- We also introduce a dimensionless time $\tau = r_1 t$ and set $\rho = r_1/r_2$

Generalization of the model

(7) becomes

$$\frac{du_1}{d\tau} = u_1(1 - u_1 - a_{12}u_2)$$

$$\frac{du_2}{d\tau} = \rho u_2(1 - u_2 - a_{21}u_1) \qquad (8)$$

In absence of growth of u_1 and u_2 (8) $\rightarrow$:

$$u_1 = 0 \text{ and } 1 - u_1 - a_{12}u_2 = 0$$
$$u_2 = 0 \text{ and } 1 - u_2 - a_{21}u_1 = 0$$

Hence steady states occur at points

$$(u_1^*, u_2^*) = (0,0), (1,0), (0,1), P = (\frac{1 - a_{12}}{1 - a_{12}a_{21}}, \frac{1 - a_{21}}{1 - a_{12}a_{21}})$$

Ecological considerations

In terms of the ecology, we understand the 4 cases as follows:

- Case I $a_{12} < 1$ and $a_{21} < 1$;
  If the interspecific competition is not too strong the two populations can co-exist stably, but at lower populations than their respective carrying capacities;

- Case II $a_{12} > 1$ and $a_{21} > 1$;
  Interspecific competition is aggressive and ultimately one population wins, while the other is driven to extinction. The winner depends upon which has the starting advantage;

- Case III, IV $a_{12} < 1$ and $a_{21} > 1$ or $a_{12} > 1$ and $a_{21} < 1$;
  Interspecific competition of one species dominates the other and, since the stable node in each case globally attracts, the species with the strongest competition always drives the other to extinction

Summary

1. We learn how to model the dynamics of prey and predator population
2. We have also demonstrated how other type of interactions could be modeled using Lotka-Volterra model.

Outline

Chap III: Modeling infectious diseases

Goal

We will model the dynamics of infectious diseases in populations, but in principle, mathematical modeling can be applied to any system, biological or otherwise.

Historical background

- Ronald Ross, a trained physician and poet published *Report on the Prevention of Malaria in Mauritius* in 1908 and *The Prevention of Malaria* in 1910.

- Despite his proof of the role of certain mosquitoes in the transmission of malaria, Ross met skepticism when he claimed that malaria could be eradicated simply by reducing the number of mosquitoes.

- In the second edition of his book *The Prevention of Malaria* published in 1911, he tried to build mathematical models of the transmission of malaria in order to support his claim. One of his models consisted of a system of differential equations initiating the mathematical modeling of diseases.

- As such, mathematical models are a simplified representation of how an infection spreads across a population over time.

Notation and computation

- Let's assume that human total population N is divided into 3 compartments.
- With S being the compartment of **susceptible** i.e. the class of individuals who are susceptible to infection;
- I being the compartment of **Infected** i.e. in this class, the number of parasite is sufficiently large within the host and there is potential in transmitting the infection to other susceptible individuals;
- R being the compartment of **Recovered** i.e. this class includes all individuals who have been infected and have recovered and got removed from the initial population by acquired immunity.
- the total population size is considered constant and is given by

$$N = S + I + R$$

Notation and computation

- Let β be the transmission rate of the disease and γ the recovery rate.
- At anytime the dynamics of the disease **in absence of birth and death of the population** is read

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I \qquad (9)$$
$$\frac{dR}{dt} = \gamma I$$

- To be well defined mathematically the initial condition $S(0) \geq 0, I(0) \geq 0, R(0) \geq 0$
- This model is called the SIR model or SIR system.

Mathematical properties of the SIR model

- Because $\frac{dS}{dt} < 0$ for all t, the number of susceptible individuals is always declining, independently of the initial condition $S(0)$. Since $S(t)$ is monotone and positive, we have

$$\lim_{t \to \infty} S(t) = S_\infty$$

- The number of recovered individuals also has monotone behavior, independently of the initial conditions since $\frac{dR}{dt} > 0$

$$\lim_{t \to \infty} R(t) = R_\infty$$

- the number of infected individuals may be monotonically decreasing to zero, or may have nonmonotone behavior by first increasing to some maximum level, and then decreasing to zero. The prevalence first starts increasing if

$$I^{'}(0) = (\beta S(0) - \gamma)I(0) > 0$$

Basic reproduction number

- Hence, a necessary and sufficient condition for an initial increase in the number of infected is $\beta S(0) - \gamma > 0$ or $\frac{\beta S(0)}{\gamma} > 1$

- At the outset of an epidemic, nearly everyone (except the index case) is susceptible. So we can say that $S = 1$, we arrive at the following inequality

$$\frac{\beta}{\gamma} = R_0 > 1$$

Basic reproduction number

- R_0 also denoted **reproduction number** represents the number of secondary infections produced by a single infected individual and can be expressed as the product of the expected duration of the infectious period and the rate at which secondary infections occur.

- R_0 is mathematically and epidemiologically interesting due to the behaviour of epidemics in relation to the threshold $R_0 = 1$

- Specifically, under a deterministic ODE model:
  - If $R_0 > 1$ the epidemic will take off
  - If $R_0 < 1$ the epidemic will die out

Exercise: Basic Covid-19 modeling in the NRW

Hardly anyone was immune at the beginning of the epidemic, so almost everyone was susceptible. We will assume that there was a trace level of infection in the population, say, 10 people. Thus, our initial values for the population variables are $S(0) = 7,900,000; I(0) = 10; R(0) = 0$ We don't know the values for the parameters beta and gamma yet, but we can estimate them, and then adjust them as necessary to fit the excess death data. We have already estimated the average period of infectiousness at three days, so that would suggest gamma $= 1/3$. If we guess that each infected would make an infecting contact every two days, then beta would be $1/2$.
The solution is given in the application section

Summary

1. We have leaned to model the dynamics of a disease using an SIR model
2. We computed the reproduction number which is an epidemiological parameter expressing the speed of the disease

Outline

Application

Please visit this link and download the R script
https://github.com/SaviKoissi/Mathematical-Biology

Outline

Evaluation sheet

Please visit this link and give your feedback
https://form.jotform.com/203655184201347