The Leslie Matrix and Population Change

Anahita Hedayatifard
Melika Hassani Barsari
Nazila Zarei

Department of Mathematics,
College of Sciences,
University of Tehran

July, 2023

Abstract

Matrices are frequently used in all the biological fields as organizers of numerical information since they greatly increase the ease and efficiency of analyzing data. Lezlie matrix model is a popular and useful method for population prediction based on age-specific survival and fecundity rate. This model is easy to understand by biologists and since it has many theoretical features that can be easily studied using matrix algebra, it has become very popular. This article will describe in detail the cases, properties, application and development of the Leslie matrix model. We will be particularly interested in determining whether a population is increasing without bound, stable or dying out.

Keywords

Leslie Matrix, Survival rate, Fecundity rate, Stable population, Growing population, Extincting population.

Introduction

Births and deaths are age-dependent in most species except very simple organisms. It is usually necessary to take account of a population's age structure to get a reliable description of its dynamics. Matrix population models have become increasingly important and useful in predicting population growth. The first work to use matrix algebra for population projection is by Harro Bernardelli, and it appeared in 1941. Bernardelli was concerned with the oscillations of human population dynamics, rather than with its long-term stabilization. Matrix population models were also suggested by E. G. Lewis in a paper that appeared in 1942. Lewis essentially suggests the first discrete model and considers the projection matrix associated with it and its characteristic equation. Patrick Holt Leslie, who was apparently unaware of Lewis's article, rediscovered the model in his 1945 paper. Leslie considered the mathematical properties of the model more thoroughly as well as the resulting convergence to a stable age distribution.

Although Matrix models are in some ways the simplest of the mathematical approaches to structured population modeling, their analysis requires computational power. For this reason, and because ecologists of the day viewed matrix algebra as an esoteric branch of advanced mathematics, they were largely neglected until the late 1960's, when they were rediscovered by ecologists and human demographers. This basic model is an increasingly popular tool for describing the population dynamics of plants and animals. The Leslie matrix model is widely used to project the present state of a population into the future, either as an attempt to forecast the age distribution or as a way to evaluate life history hypotheses by considering different sets of survival and fecundity parameters.

In a Leslie model, the population is divided into groups based on age classes. At each time step, the population is represented by a vector with an element for each age class where each element indicates the number of individuals currently in that class. 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.

Leslie Matrix

As mentioned before, one of the most common models of population growth used by demographers is the so-called Leslie model developed in the 1940s. This model describes the growth of the female portion of a human or animal population. In this model, the females are divided into age classes of equal duration. To be specific, suppose that the maximum age attained by any female in the population is L years (or some other time unit) and we divide the population

into m age classes. Then each class is years in duration. We label the age classes according to Table 1.[2]

Age Class	Age Interval
1	$[0, L/m)$
2	$[L/m, 2L/m)$
3	$[2L/m, 3L/m)$
...	...
$m-1$	$[(m-2)L/m, (m-1)L/m)$
m	$[(m-1)L/m, L]$

Table 1: Age Classes and Intervals

The usual method for organizing population data is a life table which contains the basic information needed to study changes in population density as well as rates of increase or decrease. We would like to be able to use the information in the life table to predict the growth of the population of the chosen species. To do this we must develop a method for determining the number of individuals in each age class one time period later, given the current age distribution.

Let $n_{i,k-1}$ be the number of individuals in the ith age category at time $k-1$. Then the number of individuals in the $(i+1)$th age category at time k can be found by multiplying the survival rate from age category i to $(i+1)$ by the number of individuals. Thus we can write:

$$n_{i+1,k} = s_i n_{i,k-1}, \quad i = 1, \ldots, m-1. \tag{1}$$

where s_i is the survival rate from the ith to the $(i+1)$th age category and m is the number of age categories. The long-term growth of the population is also dependent on the number of daughters born during each time period. The average number of offspring for each female usually varies with the age of the female. Thus, in order to determine the number of births at a particular time, we must multiply the number of females in each age category by the appropriate birth rate for that category. Since the number of births at time $k-1$ is the same as the number of individuals in the first age category at time k, this can be summarized as:

$$n_{1,k} = b_1 n_{1,k-1} + b_2 n_{2,k-1} + \ldots + b_m n_{m,k-1} \tag{2}$$

where b_i is the average number of daughters born to each female in the ith age category. (Note that we assume that the survival and fecundity rates of each category are constant over time and therefore not dependent on population density. [5])

For a better understanding of what's happening, it's good to take a look at the life-cycle graph. A life-cycle graph describes the transitions an individual can make during a time interval. Individuals in one age class can contribute to

3

another only by surviving to the next older age class or by reproduction to the first age class.
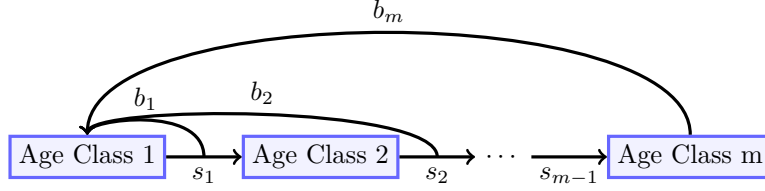
The population information contained in the n equations of expressions (1) and (2) can be written compactly in the form of a Leslie matrix (sometimes called the projection matrix). Let $\mathbf{N_k} = [n_{1k}, ..., n_{mk}]^T$ be the vector representing the number of females in each age category at time k. Then:

$$\mathbf{N_k} = \begin{bmatrix} n_{1k} \\ n_{2k} \\ \vdots \\ n_{mk} \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & b_3 & \ldots & b_{m-1} & b_m \\ s_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & s_2 & 0 & \ldots & 0 & 0 \\ & & & \ldots & & 0 \\ 0 & 0 & 0 & \ldots & s_{m-1} & 0 \end{bmatrix} \begin{bmatrix} n_{1k-1} \\ n_{2k-1} \\ n_{3k-1} \\ \vdots \\ n_{mk-1} \end{bmatrix} = \mathbb{L}\mathbf{N_{k-1}}. \quad (3)$$

The matrix $\mathbb{L}$ is the Leslie Matrix and summarizes the information necessary to describe the growth of the population. Note that $\mathbb{L}$ is a nonnegative, square matrix of order m with all the elements zero except those in the first row and in the subdiagonal immediately below the principal diagonal. ($0 < s_i < 1$, $b_i > 0$)

Survival probability s_i**:** The values of s_i can be obtained from a life table or experimental data. It is usually assumed that

$$s_i = \frac{\mathbb{L}_{i+1}}{\mathbb{L}_i}$$

where $\mathbb{L}_i = $ *the number alive in the age group i to $i + 1$ in the stationary or life table age distribution.* The value of s_i is assumed to be constant over a unit of time and not dependent on the total number in the population. Actually, seasonal differences and density dependence of the survival rate do exist, but we assume that they are not significant enough to count.

Fecundity rates b_i**:** The value of b_i is determined not only by the average number of female births per mother per time period, m_i in the life table, but also by infant survival rates. Populations can be divided into two cases of breeding systems, 'birth-flow' population with continuous reproduction and 'birth-pulse' populations

with discrete reproduction. Populations, e.g. human, of the first case produce offspring at a rate almost constant throughout the year and those of the second case produce offspring over a restricted season, e.g. blue whale. In birth-flow populations, the estimation of b_i involves integration of continuous fertility and infant mortality functions over the time interval. For birth pulse populations reproduction is during a short period of time per year. It is important to know the exact breeding period with respect to the time step in birth-pulse populations. Two different ways of counting the population, census just before or after reproduction, will influence the estimation of b_i (Cullen 1985). Some studies count the population just before reproduction and we illustrate that here. Define:

M_i = average number of female offspring per female in category i born between t and $t + 1$,

U_0 = the offspring survival rate between birth and the time when they are counted as part of the population, then

$$b_i = M_i.U_0, i = 1, 2, ..., x.$$

The age specific fecundity rate, b_i, is assumed to be constant and not dependent on density just as the age specific survival rate, s_i, is.

Solving the Leslie equation

Dependence on initial conditions

To determine the population in each category in the first time period, we let $k = 1$ in equation (3) and get:

$$\mathbf{N_1} = \mathbb{L}\mathbf{N_0} \qquad (4)$$

where $\mathbf{N_0}$ represents the population distribution at the beginning of the process. In laboratory situations, $\mathbf{N_0}$ is known by the investigator while in field situations usually it can only be estimated. For the second time period, the population distribution is:

$$\mathbf{N_2} = \mathbb{L}\mathbf{N_1}$$

from equation (3). However, N_1 is known from equation (5) so that can be written in terms of N_0 as:

$$\mathbf{N_2} = \mathbb{L}^2\mathbf{N_0}. \qquad (5)$$

Analogously, the population distribution during the kth time period is given by:

$$\mathbf{N_k} = \mathbb{L}^k\mathbf{N_0}. \qquad (6)$$

Thus, once the initial age distribution vector of the population and its associated Leslie matrix are known, the population age distribution vector at any subsequent time k can be determined by applying equation (7).

Although equation (7) supplies an expression for the solution for all times k, it may take a lot of time to compute $\mathbb{L}^k$ if k is large. Also, sometimes we are not actually interested in an exact description of the population at a given time, but rather in the general trend of the population. Since a trend or pattern is sometimes difficult to distinguish in a mass of data (which may itself take a long time to compute), we shall describe a method by which an estimate for $\mathbf{N_k}$ can be obtained without finding $\mathbb{L}^k$.

The dominant eigenvalue

The eigenvalues or characteristic values of a matrix $\mathbb{L}$ are the numbers which we will use to describe the behavior of $\mathbb{L}^k$ as k gets large. In order to find the eigenvalues, we must solve the equation,

$\det(\mathbb{L} - \lambda\mathbb{I}) = 0$. From this formulation, it can be seen that the eigenvalues of a matrix are those values which make the matrix $\mathbb{L} - \lambda\mathbb{I}$ singular or noninvertible. As we learned during the course, another way to describe the eigenvalues is that they are the values λ for which:

$$\mathbb{L}x = \lambda x \tag{7}$$

for some vector $x \neq 0$. In other words, if we premultiply some particular nonzero x by λ we only change the magnitude of the resultant vector and not its direction. The nonzero vector x is called an eigenvector of the matrix.

The eigenvalues of the Leslie matrix in equation (3) are the roots of the equation:

$$\det(\mathbb{L} - \lambda\mathbb{I}) = \lambda^m - b_1\lambda^{m-1} - b_2s_1\lambda^{m-2} - b_3s_1s_2\lambda^{m-3} \\ - \cdots - b_ms_1\ldots s_{m-1} = 0. \tag{8}$$

Clearly, if we have a lot of age categories in the population, i.e. if m is large, then equation (9) will be difficult to solve exactly. However, we can derive some qualitative information from equation (9) without actually finding all the roots.

Provided all the roots of equation (9) are nonzero, we can divide through by λ^m to get:

$$\frac{b_1}{\lambda} + \frac{b_2s_1}{\lambda^2} + \cdots + \frac{b_ms_1\ldots s_{m-1}}{\lambda^m} = 1. \tag{9}$$

Calling the expression on the left $q(\lambda)$ and Noting that all the b_i and s_i are greater than zero, we see that as λ approaches zero, $q(\lambda)$ tends toward infinity and as λ goes to infinity, $q(\lambda)$ goes monotonically to zero. Thus, the graph of $q(\lambda)$ will have the general form shown in Fig. 1. The only positive solution to equation (10) is λ, as indicated in Fig. 1. The other roots of equation (10) must either be negative or complex.
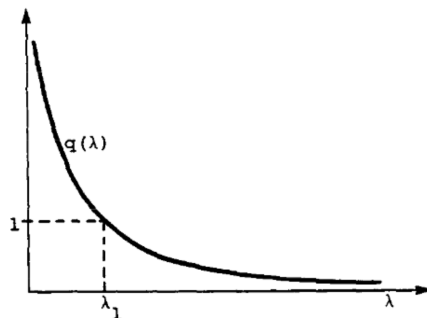
Figure 1:

There's a very famous and important theorem in matrix theory proved by Oskar Perron (1907) and Georg Frobenius (1912) that we're going to use:

Theorem 1 (Perron-Frobenius). *If all entries of an $n \times n$ matrix A are positive, then it has a unique maximal eigenvalue. Its eigenvector has positive entries. [6]*

Since $\mathbb{L}$ is what is known as an irreducible nonnegative matrix, Perron-Frobenius' theorem can be applied to guarantee that λ_1 is a simple root (not repeated) of equation (10) and that $|\lambda_i| \leq \lambda_1$ for all the other roots, λ_i, of equation (10). Because of this, λ_1 is called the dominant eigenvalue of the matrix. Further, if $|\lambda_i| < \lambda_1$ for $i = 2, \ldots, m-1$, it is called strictly dominant and the long-term behavior of the population is governed by only λ_1 and its associated eigenvector x_1.

The dominant eigenvector

The eigenvector associated with an eigenvalue is the nonzero vector which satisfies equation (8). To find the eigenvector associated with λ_1, we must find a basis for the null space of $(\mathbb{L} - \lambda_1 \mathbb{I})$. That is, we must find a nonzero vector x_1 such that:

8

$$(\mathbb{L}-\lambda_1\mathbb{I})x_1 = \begin{bmatrix} b_1-\lambda_1 & b_2 & b_3 & \dots & b_{m-1} & b_m \\ s_1 & -\lambda_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & -\lambda_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s_{m-1} & -\lambda_1 \end{bmatrix} x_1 = 0. \quad (10)$$

Since a null vector of a simple eigenvalue will have 1 degree-of-freedom, we choose the first component of x_1 to be 1 and the other components can be found in terms of λ_1 by solving $m-1$ equations. The result of this process yields the eigenvector:

$$x_1 = \begin{bmatrix} 1 \\ \frac{s_1}{\lambda_1} \\ \frac{s_1 s_2}{\lambda_1{}^2} \\ \vdots \\ \frac{s_1 s_2 \dots s_{m-1}}{\lambda_1{}^{m-1}} \end{bmatrix}. \quad (11)$$

Diagonalization of $\mathbb{L}$

If $\mathbb{L}$ has m distinct eigenvectors, it can be written as the diagonal form:

$$\mathbb{L} = \mathbb{S}A\mathbb{S}^{-1}. \quad (12)$$

where A is an $m \times m$ diagonal matrix with the eigenvalues of $\mathbb{L}$ on the diagonal and $\mathbb{S}$ is the matrix with the eigenvectors of $\mathbb{L}$ as its columns. This diagonalization provides an easy method for computing $\mathbb{L}^k$. First note that:

$$\mathbb{L}^2 = \mathbb{S}A\mathbb{S}^{-1}\mathbb{S}A\mathbb{S}^{-1} = \mathbb{S}A^2\mathbb{S}^{-1}.$$

and in fact:

$$\mathbb{L}^k = \mathbb{S}A^k\mathbb{S}^{-1}. \quad (13)$$

Expression (14) can now be used in equation (7) to get:

$$\mathbf{N_k} = \mathbb{S}A^k\mathbb{S}^{-1}\mathbf{N_0}. \quad (14)$$

We can divide equation (15) by λ_1^k, yielding:

$$\frac{1}{\lambda_1^k}(\mathbf{N_k}) = \mathbb{S}\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\lambda_2^k}{\lambda_1^k} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\lambda_m^k}{\lambda_1^k} \end{bmatrix}\mathbb{S}^{-1}\mathbf{N_k}. \tag{15}$$

Limiting Behavior

The strict dominance of λ_1 ensures that for $i = 2, \dots, m$,

$$\lim_{k\to\infty} \frac{\lambda_i^k}{\lambda_1^k} = 0.$$

so that in the limit, Equation (16) becomes:

$$\lim_{k\to\infty} \frac{N_k}{\lambda_1^k} = \lim_{k\to\infty} \mathbb{S}\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\lambda_2^k}{\lambda_1^k} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\lambda_m^k}{\lambda_1^k} \end{bmatrix}\mathbb{S}^{-1}\mathbf{N_0}. \tag{16}$$

In equation (16) $\mathbb{S}^{-1}\mathbf{N_0}$ is a vector with constant entries. In the long term, the only entry of importance is the first component which we will call c. The product of the matrix $\mathbb{S}$ with the matrix which is all zeros except for the 1 in the first row, first column is just the first column of $\mathbb{S}$ - the eigenvector x_1. Thus, we can write:

$$\lim_{k\to\infty} \mathbf{N_k} = \lim_{k\to\infty} \lambda_1^k c x_1. \tag{17}$$

For large values of k, equation (18) without the limit signs can give a good approximation to the population distribution.

From equation (18), it is clear that the population vector $\mathbf{N_k}$ will grow without bound if $\lambda_1 > 1$. If $\lambda_1 = 1$, then $\mathbf{N_k} = c x_1$ for all k, and the population is described as stable. This stable case corresponds to the popular term *zero population growth*. However, if $\lambda_1 < 1$, then $\lim_{k\to\infty} \lambda_1^k = 0$ so that the population is dying out. This latter case is that of a species facing extinction.

If the population of interest is stable, then its dominant eigenvalue is $\lambda_1 = 1$, and from equation (9) we see that:

$$b_1 + b_2 s_1 + b_3 s_1 s_2 + \cdots + b_m s_1 \ldots s_{m-1} = 1. \qquad (18)$$

The expression on the left of equation (19) is called *the net reproduction rate, R*, of a population. It corresponds to the average number of daughters born to each female during her lifetime. If $R > 1$, then the population is growing. If $R < 1$, then females are not producing sufficient offspring to replace themselves, so the population is dying out.

Conclusion

The primary purpose of this essay was to demonstrate how the Leslie matrix can be constructed, providing a thorough understanding of the mathematics behind the parameters in the matrix. One of the major findings we made was that the matrix depends only on the fertility and survival rates, referred to as the vital rates. A few necessary estimations were made in order to predict the future behaviour of the population. We normally tend to deal with two kinds of populations, birth-flow and birth-pulse populations. While it is theoretically possible to construct a more elaborate matrix to include all the possible transitions, it was to our advantage to instead consider a simpler case.

Matrix model have been widely used in biological research because of their simplicity. A recent review showed that they constitute a powerful tool for analysing population dynamics, the influence of the present population state on its capacity for growth, and the sensitivity of the dynamics to quantitative changes in the vital rates. When using the matrix model, it is possible to combine the computational case of the matrix with the flexibility of continuous time models.

In summary, we archived an understanding of the construction and contents of the matrix. We demonstrated that all age classes eventually grow at the same rate, that the asymptotic rate of growth is the leading or most positive eigenvalue, and the eigenvector corresponding to the largest eigenvalue gives the stable age structure.

References

[1] P. H. Leslie, *On the use of matrices in certain population mathematics*, Biometrika 35, 213-245, 1968.

[2] H. Anton, C. Rorres, *Elementary linear algebra with applications*, 11th edition, 673 (2014)

[3] S. Tuljapurkar, H. Caswell, *Structured-population models in Marine, Terrestrial, and Freshwater Systems*, 21(1997)

[4] M. lannelli, M. Martcheva, F.A. Milner, *Gender-Structured Population Modeling*

[5] W. c. Li, *Application Of Leslie Matrix Models To Wild Turkey Populations*,1994

[6] Perron-Frobenius theorem