AN AGE STRUCTURED POPULATION MODEL

BY

KIPCHIRCHIR ISAAC CHUMBA

A dissertation submitted in partial fulfilment for the degree of Master of Science in Mathematical Statistics in the Department of Mathematics

University of Nairobi
Nov. 1993
DECLARATION

This dissertation is my original work and has not been presented for a degree in any other University.

KIPCHIRCHIR ISAAC CHUMBA

Nov. 1993

This dissertation has been submitted for examination with my approval as University Supervisor.

Dr. JOHN OWINO

Nov. 1993

Department of Mathematics
University of Nairobi
P.O. Box 30197
NAIROBI
KENYA.
(ii)

ACKNOWLEDGEMENT

I wish to thank Dr. John O.Owino, member of staff, Department of Mathematics who stimulated my interest in the application of matrix algebra to biological systems and also as my Supervisor. He devoted his valuable time to read my project drafts and made useful comments.

I also wish to thank many members of staff in the Department of Mathematics for their lectures in courses that are directly or indirectly related to the present work and in particular Mr. H. Mwambi who on a number of occasions discussed the problems with me.

I'm indebted to the University of Nairobi for the award of scholarship which facilitated my pursuance of postgraduate studies.

Finally, thanks goes to my family for having been kind, caring, and encouraging throughout my study period.
# LIST OF CONTENTS

## CHAPTER 1. INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Objective and Significance of the Study</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Brief Literature Review</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>Eigenvalues and Eigenvectors</td>
<td>11</td>
</tr>
</tbody>
</table>

## CHAPTER 2. AGE STRUCTURED MATRIX POPULATION MODEL

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Basic concepts of matrix population model</td>
<td>19</td>
</tr>
<tr>
<td>2.2</td>
<td>Age structured model</td>
<td>35</td>
</tr>
<tr>
<td>2.3</td>
<td>Stable population Theory</td>
<td>47</td>
</tr>
<tr>
<td>2.4</td>
<td>Harvesting in matrix population model</td>
<td>58</td>
</tr>
</tbody>
</table>

## CHAPTER 3. APPLICATION OF THE MODEL

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>67</td>
</tr>
<tr>
<td>3.2</td>
<td>Application to generate and maintain the stable population structure</td>
<td>68</td>
</tr>
<tr>
<td>3.3</td>
<td>Concluding Remarks</td>
<td>71</td>
</tr>
</tbody>
</table>

## APPENDIX

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>74</td>
</tr>
</tbody>
</table>

## REFERENCES

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>80</td>
</tr>
</tbody>
</table>
1.1 Introduction.

Studies in population dynamics have been carried out solely to predict future characteristics of a population when the past or the present is known.

One such characteristic is the population size which will be discussed in the sequel. Various models have been derived to predict future population size given the present/past.

These models can further be classified as:
(i) Calculus models
(ii) Matrix models

In calculus models the population parameters are assumed to be continuous on time scale whereas in matrix models they are assumed to be discrete.

The notable calculus model describing the growth of a population is

\[ n_t = n_0 e^{rt} \]  \hspace{1cm} (1.1.1)

where \( n_t \) is the population size at time \( t \), \( n_0 \) is the initial population size (i.e. at time \( t=0 \)) and \( r \) is the intrinsic rate of natural increase and often referred to as "Lotka's \( r \". (1.1.1) is also called the Malthusian model (Malthus (1798)) and studied by Lotka (1925).

Remark.

\( r > 0 \) will result in an indefinite population growth and \( r < 0 \)
will result in extinction of the population. These two cases are not applicable hence this type of model is only suitable for short term projection.

An improved model yields the logistic equation (so-called Verhulst-Pearl logistic equation, (Verhulst (1838)); Pearl (1927)).

\[
n_t = \frac{n_0 K}{n_0 + (K-n_0) e^{-r_0 t}}
\]

where \( n_t \) and \( n_0 \) are as defined for the exponential growth equation given by (1.1.1.). \( r_0 \) is some initial growth rate and can be estimated using the available data and \( K \) is the carrying capacity of the population.

Remark
The logistic equation has an advantage of long term projection.

In arriving at the above equations the assumptions taken were;

(i) Birth and death rates were independent of age or that population growth occurred in such a way that the age distribution remained unaltered.

(ii) Density dependence i.e. a decrease in birth rate and an increase in death rate as the population become larger.

The study of population dynamics concepts has demonstrated the complexity of calculus and the simplicity of matrices. In reversing the above assumptions we have the
following:

(i) An individual's chances of reproducing and dying are a function of its age.

(ii) These chances are unaffected by the size of the population in which it finds itself, perhaps because the population numbers never rise so high that density dependence begins to exert any influence.

This latter set of assumptions characterizes the matrix models. The need to employ the matrix models rather than the calculus models is because the calculus models merely predict the size of the population after a lapse of time, given the initial size, but matrix models have an added advantage that it also gives the age structure(distribution).

Other inherent advantages of matrix models over the calculus models are:

(i) Matrix processes are assumed to take place in discrete time units whilst the very nature of calculus implies continuous process taking place in infinitesimally small time intervals. It would appear then that the matrix approach is more reminiscent of biological process of birth and death with which we are concerned with when modelling.

(ii) Matrices are relatively dealt with by numerical rather than algebraic techniques. This is a feature of importance when simulation of populations is required.

The Leslie Matrix model or the Age-Structured Model is
deterministic. It predicts the age structure of a population of female animals, given the age-structure at some past/present time and given the age-specific survival and fecundity rates.

Instead of the differential and integral calculus, matrix algebra is used, a step which leads to a great economy in the use of symbols and consequently to equations which are more easily handled. Moreover, many quite complicated arithmetical problems can be solved with great ease by manipulating the matrix which represents the given system of age specific rates (fecundity and survival rates).

But the question then arises whether these advantages mentioned may not be offset by a greater degree of inaccuracy in the results as compared with those obtained from the previous methods of computation. It is not easy, however, to settle this point satisfactorily at the moment and requires investigation.

The study shall be based on data of a Gorilla species, Gorilla Berengei (Mountain Gorilla).

The data compiled from several sources including:-
1.2 Objective and Significance of the study.

Concerned to predict the size and age structure of a population after a lapse of time (or at successive intervals of time) given the age-specific rates of fecundity and survival.

To employ matrix algebra and numerical methods to determine the latent vector of the stable age distribution and the intrinsic rate of natural increase.

More precisely, if $A$ is the matrix describing the model then

$$\lim_{t \rightarrow \infty} A^t n_0 = n_s$$

(1.2.1)

where $n_0$ is the initial age distribution and $n_s$ is the age distribution at time $s$. The elements of $n_s$ are proportional to the sizes of the age groups when the age distribution is stable. We shall establish that when the population has stabilized

$$n_{s+1} = A n_s = \lambda n_s$$

(1.2.2)

where $\lambda$ is a dominant eigenvalue of $A$.

If the population size increases from $n$ to $\lambda n$ over one period of time, we shall show that the harvest that can be taken is given by

$$H = 100(\lambda - 1)/\lambda$$

(1.2.3)

where $H$ is expressed as a percentage of total population. A constant population size is maintained if the harvest is spread proportionately over all of the age groups. However, if the exploitation is aimed at only one age group then an
enhanced exploitation can be taken.

If there is no harvesting the model predicts that the population size continues to increase at a constant rate, namely by a multiplicative factor of $\lambda$ every period of time. The model thus closely approximates the calculus model describing exponential growth of population given by (1.1.1). Assuming stability the matrix model is described by the equation

$$n_t = \lambda n_{t-1}$$  \hspace{1cm} (1.2.4)

and by induction

$$n_t = \lambda^t n_0$$  \hspace{1cm} (1.2.5)

where $\lambda$ is the dominant eigenvalue of $A$. These two models in population dynamics can be related by the equation

$$r = \log_e \lambda$$  \hspace{1cm} (1.2.6)

It is remarkable to note that the matrix model takes into consideration the age structure of the population.

The essence of predicting the size and age structure of a population at some given time in future is to be utilized in economic and social planning with respect to the population under study.

In human populations, in particular, the classes are usually the quinquennial (or other) age groups 0-4, 5-9, ..., 85 and over, and once the number of individuals in an age group has been predicted one can plan for the requirements appropriate for that age group. The requirements being referred to above are schools, hospitals,
infrastructure, recreation facilities, employment opportunities, housing, food e.t.c.

The system of age specific rates can also be used to model the dynamics of a wide variety of ecological populations e.g. Forests, Wildlife, Insects e.t.c. In populations of pupating insects the classes might be the three stages larvae, pupae and adults.

The results of the study also provide information on harvesting. Harvesting here refers to removal of some individuals from a given age group for consumption or otherwise. Often the harvest is not spread proportionately over all age groups but is taken chiefly from some subset of the age groups.

Finally the results of the study will provide reading and reference material for further researches/investigations by scientists both in Physical/Biological sciences and Social Sciences. The study is also useful to Biostatisticians in their research activities.

1.3 Brief literature review.

Since Leslie (1945, 1948) first put their use on a firm footing, projection matrices have proved exceedingly valuable in ecological and human demography. Further research still continues to expand their use in many directions. A very similar model had, independently been described earlier by Bernardelli (1941) and Lewis (1942).

Herebelow is a brief review on the use of the Leslie's
Matrix Model and its developments.

(i) Polard (1966) showed that the predictions based on the model are not markedly affected by changes in the time interval. Thus he compared two estimates of the rate of increase per annum of Australian human females and showed that the estimates were very close, regardless of whether the age classes were one year or five years long. The error introduced by altering the time interval appears to be small.

(ii) When a projection matrix is used it is normally assumed that the age intervals are of equal length. However, Lefkofitch (1965) has generalized Leslie's model by considering unequal age groupings as characterized by the insect populations. The structure of these populations can most simply be in terms of the four developmental stages - eggs, larvae, pupae and adults. The lengths of time occupied by each of these four stages are in general not equal and hence one of the premises of the basic model, namely that an animal moves up by exactly one class is no longer tenable. Lefkofitch demonstrated the usefulness of his model in studies of the growth of experimental populations of the cigarette beetle, Lasioderma Serricone (Fabricus).

(iii) A modified projection matrix for the two sex situation has been explored by Williamson (1959). A considerable extension of the basic model for the two sex situation has
been made by Goodman (1969). He applies his model to the structure of human populations of the United States. His model records not the actual population age structure at time $t$ but rather how many female descendants aged under five that a female alive in the $i^{th}$ age group will have. From such studies Goodman is able to predict the eventual age structure of the United States, exactly analogous to the Stable age structure predicted by the basic model.

(iv) Forest trees are generally classified according to their size rather than their age. Usher (1966) developed the basic Leslie model for selection forests, these forests which contain an uneven age and size structure of the trees and in which there is no regeneration phase. In a similar manner to the development for the insect populations, trees can either remain in the same size class during the period of time over which the matrix operates or they move to a larger class. Usher (1966, 1967/68, 1969) has assumed that the period of time over which the matrix operates is sufficiently small for a tree not to move up by more than one size class. The major difficulty for the forest model concerns the analogue to the fecundity terms of the model. Usher (1972) has applied the model to forest management on pine trees.

(v) Stochastic versions have been described by Sykes (1969b) and Polard (1966/73). In particular, Polard (1966) has developed a stochastic form of Leslie's basic model, giving
for each integral point of time the mean and variance of the female population of Australia, showing that after 250 periods of time (in this case one year) the ratio of the numbers in the age groups gives a very close approximation to the latent root.

(vi) Keifitz (1964) has applied the model to human demography. In particular, he discussed empirically the approach to stability and illustrated several different ways of calculating the intrinsic rate of natural increase, and examined how these calculations are related to one another.

(vii) Polard (1966) has described some deterministic extensions which allow for immigration, namely:

(a) It is assumed that each year a constant immigration vector \( b \) is added to the population. Then

\[
 n_{t+1} = A n_t + b \tag{1.3.1}
\]

(b) It is assumed that, instead, a population vector proportional to the actual population vector is added to the population each year. That is

\[
 n_{t+1} = ( A + \alpha I ) n_t \tag{1.3.2}
\]

(c) It is assumed that immigration may be represented by a combination of the following three factors:

* An immigration vector whose elements are proportional to those in the population.
* A constant immigration vector; and
* An immigration vector whose elements grow exponentially at the same rate with time.
Mathematically the model may be represented by the recurrence equation

\[ n_{t+1} = A n_t + a n_t + b + 6^{t+1} c \]  \hspace{1cm} (1.3.3)

1.4 Eigenvalues and Eigenvectors.

Generally, we shall be concerned with the exposition and proof of results of the eigenvalue/eigenvector problem.

Defn.1. characteristic root/vector

If A is an nxn matrix and \( \lambda \) is a scalar the characteristic polynomial of A is defined as

\[ f(\lambda) = \det (A - \lambda I) \] \hspace{1cm} (1.4.1)

and when it is expanded it produces an \( n^{th} \) degree polynomial in \( \lambda \). \( f(\lambda) = 0 \) is called the characteristic equation of A i.e.

\[ \lambda^n + k_1 \lambda^{n-1} + k_2 \lambda^{n-2} + \ldots + k_{n-1} \lambda + k_n = 0 \] \hspace{1cm} (1.4.2)

and it factors out linearly as \( f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n) \) and the n roots \( \lambda_1, \lambda_2, \ldots, \lambda_n \) of \( f(\lambda) \) are called the characteristic roots of A (also called latent roots, proper roots, secular roots, eigenvalues). The set \( \{ \lambda_1, \lambda_2, \ldots, \lambda_n \} \) is called the spectrum of A. Although an n x n matrix always has n characteristic roots, they need not be distinct.

If \( u \) is a non-zero column vector such that

\[ A u = \lambda u \] \hspace{1cm} (1.4.3)

for some scalar \( \lambda \), \( u \) is called a characteristic vector of A corresponding to \( \lambda \) (other names latent vector, proper vector, eigenvector). Clearly \( u \) is any non-trivial solution
of the homogeneous system of linear equations
\[(A - \lambda I)u = 0.\] Such a system has a non-trivial solution
iff \[|A - \lambda I| = 0,\] thus iff \(\lambda\) is a characteristic root of \(A\).

Defn. 2. Diagonal matrix
A square matrix \(A=(a_{ij})\) of order \(n\) is called diagonal if
its elements off the diagonal are all zero i.e. \(a_{ij} = 0\ \forall\ i \neq j\) and is denoted by
\[A = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn})\] (1.4.4)

Theorem 1.
\(A\) and \(A'\) (\(A\) transpose) have the same characteristic equation.

Proof.
We note that \(|A| = |A'|\) and \(I' = I\)

Let \(f(\lambda)=0\), \(f^*(\lambda)=0\) be the characteristic equations of \(A\)
and \(A'\) respectively. We show that \(f(\lambda) = f^*(\lambda)\).

\[f(\lambda) = |A - \lambda I| = |A' - \lambda I'| = |A' - \lambda I| = f^*(\lambda)\]

Theorem 2.
If \(\lambda\) is an eigenvalue of \(A\) corresponding to \(u\) then \(1+\lambda\) is an
eigenvalue of \((I+A)\) corresponding to the same eigenvector \(u\).

Proof.
\(A u = \lambda u\) by definition.

\[(I+A)u = u+Au = u+\lambda u = (1+\lambda)u\] and by definition the proof is complete.

Corr.

\((1+\lambda)^{-1}\) is an eigenvalue of \((I+A)^{-1}\) corresponding to \(u\).
Defn. 3. Similarity.

Two $n \times n$ matrices $A$ and $B$ are called similar if

$$B = P^{-1} A P \quad (1.4.5)$$

where $P$ is an arbitrary non-singular $n \times n$ matrix.

Theorem 3.

Similar matrices have the same characteristic equation (and hence same eigenvalues).

Proof.

Let $A$ and $B$ be similar matrices and $f(\lambda) = 0$, $f^*(\lambda) = 0$ be their characteristic equations respectively. We show that

$$f(\lambda) = f^*(\lambda)$$

$$f(\lambda) = \det(B - \lambda I) = \det(P^{-1} (A - \lambda I) P)$$

$$= \det(P^{-1}) \det(A - \lambda I) \det(P) = \det(A - \lambda I) = f^*(\lambda)$$

since $\det(P^{-1}) = 1 / \det(P)$

Remark.

If two matrices do not have the same characteristic equation, then we can say categorically that they are not similar, however, if two matrices have the same characteristic equations, they may or may not be similar.

Defn. 4.

Let $u_i$ denote the eigenvector corresponding to the eigenvalue $\lambda_i$ then for $i = 1, 2, \ldots, n$

$$Au_i = \lambda_i u_i \quad (1.4.8)$$

Arrange the column vectors $u_i$ side by side to form the $n \times n$ matrix $U$. Let $D$ be a diagonal matrix whose elements are the characteristic roots of $A$. We note that the
characteristic roots may be made to appear on the diagonal of $D$ in any desired order by suitable choice of $U$. Then

$$U^{-1} A U = D \quad (1.4.7)$$

Proof.

Given that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $A$ and $u_1, u_2, \ldots, u_n$ are the corresponding eigenvectors, provided that $\lambda_1$'s are distinct then the eigenvectors $u_1$'s are linearly independent, then $|U| \neq 0$ and $U^{-1}$ exits where $U=\{u_1, u_2, \ldots, u_n\}$. Using (1.4.6) we have

$$A U = U D \quad (1.4.8)$$

which reduces to (1.4.7) on premultiplying both sides by $U^{-1}$. $U$ is called a modal matrix for $A$, and $D$ is called a spectral matrix for $A$. The similarity transformation (1.4.7) of $A$ into $D$ is called diagonalization and has many important applications.

Remark.1.

A matrix is diagonalizable if it is similar to a diagonal matrix.

Remark.2.

Any suitable vectors, $u_1$ will do. There is no need to use normalized eigenvectors. It should be noted that if some of the $\lambda$'s are repeated then $A$ is similar to the diagonal matrix $D$ only if $A$ possesses $n$ linearly independent eigenvectors.

Remark.3.
A matrix will have \( n \) linearly independent eigenvectors if all the eigenvalues are distinct or, depending upon the matrix, even if some or all of the eigenvalues are equal. A priori, therefore, we place no restrictions on the multiplicities of the eigenvalues.

Similar results can be obtained for matrices which are not diagonalizable. We begin by generalizing the concept of the eigenvector. It will follow that every matrix \( A \) has \( n \) linearly independent generalized eigenvectors and hence similar to an "almost diagonal" matrix.

Defn.5.
A vector \( u_m \) is a generalized eigenvector of rank \( m \) corresponding to the matrix \( A \) and the eigenvalue \( \lambda \) if
\[
(A-\lambda I)^m u_m = 0 \quad \text{and} \quad (A-\lambda I)^{m-1} u_m \neq 0 \quad (1.4.9)
\]

Defn.6.
Let \( u_m \) be a generalized eigenvector of rank \( m \) corresponding to the matrix \( A \) and the eigenvalue \( \lambda \). The chain generated by \( u_m \) is a set of vectors \( \{u_m, u_{m-1}, \ldots, u_1\} \) given by
\[
u_j=(A-\lambda I)^{m-j} u_m = (A-\lambda I)u_{j+1} \quad (1.4.10)
\]
for \( j=1,2,\ldots, m-1 \).

Remark.1.
\( u_j \) given by \((1.4.10)\) is a generalized eigenvector of rank \( j \) corresponding to the eigenvalue \( \lambda \).

Remark.2.
A Chain is a linearly independent set of vectors.
Remark 3.
A generalized eigenvector of rank 1 is in fact an eigenvector.

Theorem 4. (without proof).
Every $n \times n$ matrix $A$ possesses $n$ linearly independent generalized eigenvectors, abbreviated liges. Generalized eigenvectors corresponding to distinct eigenvalues are linearly independent. If $\lambda$ is an eigenvalue of $A$ of multiplicity $\mu$, then $A$ will have $\mu$ liges corresponding to $\lambda$.

Defn 7.
A set of $n$ liges is a canonical basis if it is composed entirely of chains.

We now generalize the concept of a modal matrix yielding the following definition.

Defn 8.
Let $A$ be an $n \times n$ matrix. A generalized modal matrix $M$ for $A$ is an $n \times n$ matrix whose columns, considered as vectors, form a canonical basis for $A$ and appear in $M$ according to the following rules.

(i) All chains consisting of one vector appear in the first columns of $M$.

(ii) Each chain appears in $M$ in order of increasing rank (i.e. the generalized eigenvector of rank 1 appears before the generalized eigenvector of rank 2 of the same chain, which appears before the generalized eigenvector of rank 3 of the same chain, e.t.c.)
Defn. 9. Jordan canonical matrix

A square matrix $J$ is in Jordan canonical form if it is a diagonal matrix or can be expressed in either one of the following two partitioned forms.

\[
\begin{bmatrix}
D & 0 & \ldots & 0 \\
0 & J_1 & 0 & \ldots & 0 \\
0 & 0 & J_2 & 0 & \ldots & 0 \\
& & & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & 0 & J_d
\end{bmatrix}
\]  
\hspace{1cm} (1.4.11)

or

\[
\begin{bmatrix}
J_1 & 0 & \ldots & 0 \\
0 & J_2 & 0 & \ldots & 0 \\
0 & 0 & J_3 & 0 & \ldots & 0 \\
& & & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & 0 & J_d
\end{bmatrix}
\]  
\hspace{1cm} (1.4.12)

where $D$ is a diagonal matrix, $0$'s are null matrices and $J_k$ $(k=1,2,\ldots,q)$ is given by

\[
J_k = \begin{bmatrix}
\lambda_k & 1 & 0 & \ldots & 0 & 0 \\
0 & \lambda_k & 1 & 0 & \ldots & 0 \\
& & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & 0 & \lambda_k \\
0 & 0 & 0 & \ldots & 0 & \lambda_k
\end{bmatrix}
\]  
\hspace{1cm} (1.4.13)

Remark.

Here $k$ represents some +ve integer and has no direct bearing on the order of $J_k$. A matrix in Jordan canonical form has
non-zero elements on the superdiagonal and the main diagonal, and that the elements on the superdiagonal are restricted to be either zero or one. In particular, a diagonal matrix is a matrix in Jordan canonical form that has all its superdiagonal elements equal to zero.

Theorem 5.

Every $n \times n$ matrix $A$ is similar to a matrix in Jordan canonical form.

Proof.

If $\{u_m, u_{m-1}, \ldots, u_1\}$ is a chain generated by $u_m$, then rewriting (1.4.10) we have

$$Au_j = \lambda u_j + u_j$$ (1.4.14)

for $j = 1, 2, \ldots, m-1$. Next, each complete chain of more than one vector in length that goes into composing $M$ will give rise to a $J_k$ submatrix in $J$. The order of $J_k$ is identical to the length of the chain. The chains consisting of only one vector give rise collectively to the $D$ submatrix in $J$. The elements on the diagonal of $D$ would be the eigenvalues corresponding to the one element chains. Analogously to (1.4.8) we have

$$AM = MJ$$ (1.4.15)

Noting that $|M| \neq 0$ (see thm. 4.), we postmultiply (1.4.15) by $M^{-1}$ yielding

$$A = MJM^{-1}$$ (1.4.16)
CHAPTER 2: AGE STRUCTURED MATRIX POPULATION MODEL

2.1. Basic Concepts of Matrix Population Model

Introduction.

We shall now explore the basic concepts associated with the matrix population models. These concepts shall be discussed under a special class of matrices viz the Non-negative matrices.

We shall be devoted to the exposition and proof of results concerning the eigenvalues and eigenvectors of Non-negative matrices. We will, in particular, be interested in the existence of a non-negative or strictly positive eigenvector and of a positive largest eigenvalue.

Non-negative matrices

Defn.1. Non-negative matrix

A matrix $A$ with real elements $A=(a_{ij})$ ($i=1,2,\ldots,m;j=1,2,\ldots,n$) is said to be non-negative denoted $(A \geq 0)$ or positive denoted $(A > 0)$, if all the elements of the matrix $A$ are non-negative (respectively positive): $a_{ij} \geq 0$ (respectively $a_{ij} > 0$).

Corr.

If $A \geq 0$, $B > 0$, $AB = 0$ then $A = 0$
If $A \geq 0$ then $A^m \geq 0$ for a finite integer $m$.
If $A > 0$ then $A^m > 0$ for a finite integer $m$.

Defn.2. Permutation matrix.

An $n \times n$ Permutation matrix $P$ is one in which there is a
single element equal to 1 in every row and column, all other elements being zero. In other words, the rows of \( P \) consists of a permutation of the rows of \( I_n \). Premultiplying any \( n \times m \) matrix \( A \) by \( P \) has the effect of applying the same permutation to the rows of \( A \). Similarly regarding the columns of \( P \) as a (different) permutation of the columns of \( I_n \), then postmultiplying an \( m \times n \) matrix by \( P \) performs this permutation on its columns. We remark that \( P' = P^{-1} \) for a permutation matrix.

Defn. 3. Reducible matrix.

An \( n \times n \) matrix \( A \) is reducible (or decomposable) if there exists a permutation matrix \( P \) such that

\[
P'AP = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}
\]

(2.1.1)

where \( A_1 \) and \( A_3 \) are square submatrices, \( \emptyset \) is a rectangular null submatrix and \( A_2 \) is a rectangular submatrix. If \( A \) is not reducible it is called irreducible (or indecomposable).

Lemma.

Let \( Z(x) \) denote the number of zero components of the vector \( x \). Then, if \( A \) is non-negative and irreducible, and \( x \geq \emptyset, x \neq \emptyset \) then

\[
Z[(I+A)x] < Z(x)
\]

(2.1.2)

Proof.
Put $y = (I+A)x$. It is clear that $Z(y) \leq Z(x)$ (Since $Ax \geq 0$ and $y = (I+A)x$, any coordinate of $y$ will be positive if the corresponding coordinate of the vector $x$ is positive).

Suppose that $Z(y) = Z(x)$, i.e. the zero components of $y$ are the same as those of $x$. Without lose of generality, it can be supposed that the column vectors $y$ and $x$ have the form

$$y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} ; \quad x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

where $x_1$, $y_1$, have the same dimension and are strictly positive. Corresponding to the above subdivision, we set

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{ij}$, $(i,j=1,2)$ are submatrices and in particular $A_{11}$ are square.

$$\begin{bmatrix} y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} I + A_{11} & A_{12} \\ A_{21} & I + A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

$y_1 = (I+A_{11})x_1$ implies $0 = A_{21}x_1$. These requires $A_{21}=0$, since $x_1 \geq 0$, which contradicts the assumption of irreducibility. Therefore the number of zero components of $y$ is strictly less than those of $x$.

Corr.

If $A$ is non-negative, and irreducible of order $n$, then

$$(I+A)^{n-1}>0 \quad (2.1.3)$$

Proof.

If we take any non-negative vector $x$ and apply the lemma
repeatedly, then \((I+A)^{n-1} x\) can have no zero components. As \(x\) is arbitrary we must have \((I+A)^{n-1} > 0\).

Theorem 1. Due to Perron.

An \(n \times n\) positive matrix \(A = (a_{ij})\) always has an eigenvalue \(r\) which is real and positive, which is a simple root of the characteristic equation, and which exceeds in modulus all other eigenvalues. To this dominant (or maximal) eigenvalue there corresponds an eigenvector of the matrix \(A\) with positive coordinates.

Remark.

A positive matrix is a special case of a non-negative matrix. Frobenius generalized the theorem of Perron in an investigation of the spectral properties of irreducible non-negative matrices.

Theorem 2. Due to Frobenius.

An irreducible non-negative matrix \(A\) of order \(n \times n\) always has a positive eigenvalue \(r\), which is a simple root of the characteristic equation. The moduli of all the other eigenvalues are at most \(r\). In the special case when \(A\) is positive then \(r\) is strictly greater than the modulus of any other eigenvalue.

Remark.

Since the theorem of Perron is a special case of the theorem of Frobenius, it suffices to give a proof of the latter.

Proof.

Define a number \(r(\mathbf{v^*})\) associated with the matrix \(A\) and
demonstrating its properties. The number is defined by

$$r(v^*) = \max_{v \in S} \left\{ \min_{i=1}^n \frac{(Av)_i}{v_i} \right\}$$

for $i=1,2,\ldots,n$; where

$$S = \{v / v \geq 0, \sum_{i=1}^n v_i = 1\}$$

$(Av)_i$ is the $i^{th}$ component of the vector $Av$ and $v^*$ is the vector for which the maximum occurs. In words $r(v^*)$ is found by taking weighted sums of the rows of $A$ and choosing the row $i$ with the lowest ratio of this weight to the weight for the $i^{th}$ element of each row (i.e. the $i^{th}$ column); then we maximize this value over all possible weightings. Now since $A \geq 0$, $v \geq 0$, we can easily show that $r(v^*) \geq 0$ (strictly positive if at least one element of each row and column is strictly positive).

Let $v = u = (1/n, 1/n, \ldots, 1/n) \in S$. Then

$$r(u) = \min_i \frac{(Au)_i}{u_i} = \min_i \frac{(1/n)(A1)_i}{(1/n)} = \min_i \sum_{j=1}^n a_{ij} \geq 0$$

Since $A \geq 0$. By definition

$$r(v^*) \geq r(u) \geq 0 \quad (2.1.5)$$

Thus, we see that $r(v^*)$ is greater or equal to the minimum row sum of the matrix. We now show that $r(v^*)$ gives us the maximum value which might be taken by an eigenvalue of $A$; in
other words \( r(v^*) \) is the largest number satisfying the inequality:

\[ A v \geq r v \]

The \( k^{th} \) component of the inequality gives

\[ \frac{(Av)_k}{v_k} \geq r \quad \text{i.e.} \quad r \leq \min_k \frac{(Av)_k}{v_k} \]

Clearly for a maximum of \( r \) we will have the equality with \( r=r(v^*) \). If this is not an equality, then

\[ [Av^* - r(v^*)v^*] \geq 0 \]

(i.e. is a non-negative vector). We multiply it by the strictly positive matrix \((I+A)^{n-1}\) (see (2.1.3)) to obtain

\[ (I+A)^{n-1}[Av^* - r(v^*)v^*] > 0. \]

Noting that \((I+A)^{n-1}A = A(I+A)^{n-1}\) and putting \( y = (I+A)^{n-1}v^* \) we get

\[ A y - r(v^*)y > 0. \]

However, this contradicts the fact that \( r(v^*) \) is the maximum over all \( r \), so that we must have equality in \( A v^* \geq r(v^*)v^* \) i.e. in other words \( r(v^*) \) is an eigenvalue of \( A \) and \( v^* \) the associated eigenvector, more precisely

\[ A v^* = r(v^*)v^* \quad (2.1.6) \]

If \( r(v^*) \) is an eigenvalue of \( A \), \( 1+r(v^*) \) must be an eigenvalue of \( I+A \) (see thm.2, sec.1.4), and thus \( [1+r(v^*)]^{n-1} \) will be an eigenvalue of \((I+A)^{n-1}\) with \( v^* \) as the eigenvalue in each case.

Since \((I+A)^{n-1}\) is strictly positive we have
\[ (1 + r(v^*)) n^{-1} v^* = (I + A) n^{-1} v^* > 0 \]
and using (2.1.5) we have
\[ v^* > 0 \]
(2.1.7)

Now consider any other eigenvalue \( \lambda \) with eigenvector \( x \) so that \( A x = \lambda x \). Take the absolute values and put
\[ x^+ = (|x_1|, |x_2|, \ldots, |x_n|) \]
so that we get
\[ |\lambda| x^+ \leq |\lambda x| = |Ax| = Ax^+ \]
(Since \( |a| |b| \leq |a| \) for any real numbers \( a \) and \( b \) and \( |Ax| = Ax^+ \) by the non-negativity of \( A \)). But \( r(v^*) \) is the largest number satisfying this inequality.

Therefore \( |\lambda| x^+ \leq r(v^*) x^+ \leq Ax^+ \). Hence
\[ r(v^*) \geq |\lambda| \]  
(2.1.8)
for every eigenvalue of \( A \), we will call \( r(v^*) \) the dominant root of \( A \) and denote it by \( \lambda^* \).

Further it is clear that \( \lambda^* \) cannot be a repeated eigenvalue. Suppose not, so that there was another vector \( x^* \) which would have to be strictly positive by the above argument and not equal to \( v^* \) such that \( A x^* = \lambda^* x^* \). Now we can choose \( x^* \) such that \( \Sigma_i x^*_i = 1 \), so that since \( x^* \neq v^* \) for some \( i \), \( x^*_i > v^*_i \). Subtracting the two eigenvector equations we get
\[ A(v^* - x^*) = \lambda^*(v^* - x^*) \]  
(2.1.9)
In other words the vector \( (v^* - x^*) \), which must have a negative element, is another eigenvector associated with \( \lambda^* \).

This contradicts the demonstration that the vector associated with \( \lambda^* \) must be strictly positive (see (2.1.8) & (2.1.7)). Hence \( \lambda^* \) cannot be a repeated eigenvalue.
If $A$ is a strictly positive matrix we can show that $\lambda^* > |A|$ for all $\lambda \neq \lambda^*$. Suppose not, so that $\lambda^* = -\lambda_1$. Consider $B = A - \delta I$ with two of its eigenvalues $\lambda^* - \delta$, $\lambda_1 - \delta$. Since $A > 0$, for small enough $\delta$, $B > 0$ so that its largest eigenvalue is $\lambda^* - \delta$ (see thm. 2, sec. 1.4). But $|\lambda_1 - \delta| = |-\lambda^* - \delta| > |\lambda^* - \delta|$, which contradicts the fact that $\lambda^* - \delta$ is the dominant root of $B$. This shows that, if $\alpha$ is the smallest element in $A$, then $\lambda^* - \alpha \geq \lambda$ for any $\lambda \neq \lambda^*$ i.e. $\lambda^* > |A| + \alpha$ and hence

$$\lambda^* > |A|$$

(2.1.10)

for all $\lambda \neq \lambda^*$.

Defn. 4. Primitive Matrix

Let $A \geq 0$ be an irreducible matrix, and let the dominant eigenvalue be $\lambda^*$. Suppose there are exactly $h$ eigenvalues of modulus $\lambda^*$, say $\lambda^* = \lambda_1$ then $\lambda_1 = |\lambda_2| = \ldots = |\lambda_h|$. If $h = 1$, the matrix is called primitive. If $h > 1$, the matrix is called imprimitive and the number $h$ is called the index of imprimitivity. The index of imprimitivity can be quickly calculated if all the non-zero coefficients of the characteristic equation

$$\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \ldots + a_{n-2} \lambda^2 + a_{n-1} \lambda + a_n = 0$$

of the matrix are known. More precisely, we suppose that $a_1 > 0$ for $(i = a_1, a_2, \ldots, a_{q-1}; q \leq n)$ and $n$ where $a_1 < a_2 < \ldots < a_{q-1} < n$. The exponents of $\lambda$ appearing in the characteristic equation are then $n$, $n-a_1$, $n-a_2$, $\ldots$, $n-a_{q-1}$, $0$ and considering the differences $a_1, a_2 - a_1, \ldots, a_{q-1} - a_q - 2, n - a_{q-1}$:
A is primitive iff the g.c.d. of these differences is one. i.e. \( h = 1 \) where

\[
h = \gcd\{a_1, a_2-a_1, \ldots, a_{q-1}-a_{q-2}, n-a_{q-1}\}
\]

(2.1.11)

This is also a necessary and sufficient condition for primitivity.

The following theorem gives an important property of primitive matrices.

Theorem 3.

A non-negative matrix \( A \geq 0 \) is primitive iff there is a power of \( A \) which is positive i.e. \( A^p > 0 \) for some integer \( p \).

Defn. 5. Change in sign

Given a polynomial \( f(\lambda) \) of degree \( n \), a change in sign is said to occur if two successive terms have opposite signs, missing terms being ignored. On the other hand a continuation is said to occur whenever the signs of two consecutive terms are the same.

Defn. 6. Descarte's rule of signs

(i) The number of positive roots cannot exceed the number of changes in sign of \( f(\lambda) \).

(ii) The number of negative roots cannot exceed the number of changes in sign of \( f(-\lambda) \).

Remark.

Whilst Descarte's rule does not give the exact number of real roots in an equation, it has the merit of simplicity. It still leaves an uncertainty as to the exact number of real roots in an equation, it only gives an upper limit to
Defn. 6. Companion matrix

An nxn Companion matrix C is defined as

\[ C = \begin{bmatrix} k_1 & k_2 & k_3 & \cdots & k_{n-2} & k_{n-1} & k_n \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \]  

(2.1.12)

The nth degree polynomial

\[ f(\lambda) = \lambda^n - k_1 \lambda^{n-1} - k_2 \lambda^{n-2} - \cdots - k_{n-1} \lambda - k_n \]  

(2.1.13)

is associated with the Companion matrix or more precisely the characteristic equation of C is

\[ | \lambda I_n - C | \]  

(2.1.14)

and when expanded yields (2.1.13). The first row of C consists of coefficients of \( f(\lambda) \) (except leading coefficient which is unity) with opposite signs. The polynomial \( f(\lambda) \) is called monic, since its leading coefficient is one.

Remark.

Companion matrix is a basic tool which enables polynomial problems to be handled using matrices.

Defn. 7. Leslie matrix

An nxn Leslie matrix L is defined as
where $f_i \geq 0$ and $0 < p_i \leq 1$ ($i=1,2,\ldots,n$).

Remark.

If $p_i = 1$ for all $i=1,2,\ldots,n-1$; then $L$ reduces to a Companion matrix $C$ where $k_i = f_i$.

Defn. 8. Vandermonde matrix

Given an $nxn$ matrix $A$, having distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$; the corresponding Vandermonde matrix has the form

$$V = \begin{bmatrix}
   \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \\
   \lambda_1^{n-2} & \lambda_2^{n-2} & \cdots & \lambda_n^{n-2} \\
   \vdots & \vdots & \ddots & \vdots \\
   \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\
   \lambda_1 & \lambda_2 & \cdots & \lambda_n \\
   1 & 1 & \cdots & 1
\end{bmatrix}$$

(2.1.16)

Remark.

If $\lambda_i$ is an eigenvalue of $C$ then an associated eigenvector is

$$v_i = [\lambda_1^{n-1}, \lambda_1^{n-2}, \ldots, \lambda_1^2, \lambda_1, 1]^T$$

(2.1.17)

which is the $i$th column of the Vandermonde matrix.
Similarity between C and L.

Define a transformation matrix

\[ A = \text{diag}(p_1p_2...p_{n-1}, p_2p_3...p_{n-1}, p_3p_4...p_{n-1}, ..., p_{n-1}) \]

and let \( p_{(i)} = p_ip_{2p...p_i} \); \( i=1,2,...,n-1 \) then we have

\[
\begin{bmatrix}
 f_1 & f_2 & p_{(1)} & f_3 & p_{(2)} & \cdots & f_{n-1} & p_{(n-2)} & f_n & p_{(n-1)} \\
 1 & 0 & 0 & \cdots & 0 & 0 \\
 0 & 1 & 0 & \cdots & 0 & 0 \\
 0 & 0 & 1 & \cdots & 0 & 0 \\
 \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 0 & 0 & 0 & \cdots & 1 & 0 \\
\end{bmatrix}
\]

we note that \( A_1A = C \) where \( k_1 = f_1 \) and \( k_1 = f_ip_{(i-1)} \), \( i=2,3,...,n \). We conclude that L and C are similar (see defn.4, sec.1.4) and therefore they must have the same characteristic equation given by

\[
\lambda^n - f_1\lambda^{n-1} - f_2p_{(1)}\lambda^{n-2} - f_3p_{(2)}\lambda^{n-3} - \cdots - f_{n-1}p_{(n-2)}\lambda - f_np_{(n-1)} = 0
\]

(2.1.19)

Theorem.4.

Given that the \( i \)th column of \( V \) is an eigenvector of C corresponding to \( \lambda_i \), then the \( i \)th column of \( A^{-1}V \) is an eigenvector of L corresponding to \( \lambda_i \).

Proof:

Let \( \lambda_i, i=1,2,...,n \) be distinct eigenvalues of a Companion matrix C and \( v_i \) their corresponding eigenvectors where \( v_1 \) is given by (2.1.17).

\[ Cv_1 = \lambda_1v_1 \quad i = 1,2,...,n \quad \text{and in matrix form} \]

\[ C [v_1,v_2,...,v_n] = [v_1,v_2,...,v_n] D \]
where $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.  

From (2.1.16), (2.1.17) & (2.1.18) we have $C V = V D$ and $A A^{-1} V = V D V^{-1}$ which yields

$$L(A^{-1}V) = (A^{-1}V)D \quad (2.1.20)$$

Therefore the columns of $A^{-1}V$ constitute the eigenvectors of $L$. In particular if $v_1$ is the eigenvector of $C$ then $u_1 = A^{-1}v_1$ is the eigenvector of $L$ corresponding to $\lambda_1$ in each case and

$$u_1 = [1, p_1/\lambda_1, p_1 p_2/\lambda_1^2, p_1 p_2 p_3/\lambda_1^3, \ldots, p_1 p_2 p_3 \ldots p_{n-1}/\lambda_1^{n-1}]' \quad (2.1.21)$$

Determining dominant eigenvalue.

Basically the application of the model which will be discussed in the next section will require the determination of the positive real root with maximum modulus and its corresponding eigenvector. One of the most widely used procedure is the power method.

The Power method.

Our basic assumptions are that $A$ is a real $n \times n$ matrix, its eigenvalues satisfy

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geq \ldots \geq |\lambda_n| \quad (2.1.22)$$

and that $A$ has $n$ linearly independent eigenvectors $u_i$, $i=1,2,\ldots,n$ where

$$A u_i = \lambda_i u_i \quad (2.1.23)$$

$\lambda_1$ is the dominant eigenvalue.

Since corresponding eigenvectors $u_1, u_2, \ldots, u_n$ are linearly independent they form a basis, so any arbitrary
column n-vector \( x_0 \) can be uniquely expressed as a linear combination of the basis vectors.

\[
x_0 = a_1 u_1 + a_2 u_2 + \ldots + a_n u_n
\]

(2.1.24)

where \( a_i \)'s are constants not all zero. If \( x_1 = A x_0 \), then from (2.1.23) we have

\[
x_1 = a_1 A u_1 + a_2 A u_2 + \ldots + a_n A u_n = a_1 \lambda_1 u_1 + a_2 \lambda_2 u_2 + \ldots + a_n \lambda_n u_n
\]

Similarly if \( x_2 = A x_1 \), \( x_3 = A x_2 \) e.t.c. we have

\[
x_k = a_1 \lambda_1^k u_1 + a_2 \lambda_2^k u_2 + \ldots + a_n \lambda_n^k u_n, \quad k = 1, 2, \ldots
\]

or equivalently

\[
x_k / \lambda_1^k = a_1 u_1 + a_2 (\lambda_2 / \lambda_1)^k u_2 + \ldots + a_n (\lambda_n / \lambda_1)^k u_n
\]

(2.1.25)

Since \( |\lambda_1 / \lambda_i| < 1 \) for \( i > 1 \), it follows that as \( k \to \infty \) the terms \( (\lambda_1 / \lambda_i) \to 0 \), so that as \( k \to \infty \)

\[
x_k / \lambda_1^k \to a_1 u_1
\]

(2.1.26)

i.e. \( x_k / \lambda_1^k \) tends to a multiple of the eigenvector \( u_1 \).

However, as \( k \) increases the elements of \( x_k \) are growing quite large, so to avoid this it is preferable to scale \( x_k \) at each step so that its largest element is unity in order to keep within computational limits. We therefore define a modified sequence \( x_1, x_2, \ldots \) as follows (i.e. 2.1.27/28/29)

\[
y_{k+1} = A x_k \quad k = 0, 1, 2, \ldots
\]

(2.1.27)

Let \( \beta_{k+1} = \text{element of } y_{k+1} \text{ having largest modulus} \)

\[
x_{k+1} = y_{k+1} / \beta_{k+1}
\]

(2.1.28)

(2.1.29)

as \( k \) tend to \( \infty \) we will have \( x_k \) tending to some multiple of \( u_1 \), so we can write

\[
x_k \to m u_1
\]

(2.1.30)

and then using (2.1.23), (2.1.27) & (2.1.30) we have
$y_k$ tend to $A(mu_1)=m(Au_1)=m(\lambda_1u_1)=\lambda_1(mu_1)$ and using (2.1.30) again we have

$$y_k \rightarrow \lambda_1 x_k$$  \hspace{1cm} (2.1.31)

as $k \rightarrow \infty$.

Because the largest element of $x_k$ is unity and using (2.1.28) & (2.1.29), it follows that the element of $y_k$ having largest modulus approaches $\lambda_1$, as $k$ tend to $\infty$. Thus $\beta_k$ and $x_k$ defined by (2.1.27/28/29) respectively provide successively better approximations to $\lambda_1$ and $u_1$ and the process is terminated when $x_{k+1}$ and $x_k$ are sufficiently close or equivalently when $\beta_{k+1}$ and $\beta_k$ are sufficiently close.

We wish to be fairly careful about the description of the power method, for this description has two parts, namely the "practical" part which shows how the method is actually carried out, and the "theoretical" part (described above) which shows why the method works.

**Power method "practical".**

Recalling (2.1.25), it is natural to write $x_k$ as

$$x_k = \lambda_1^k\{a_1u_1 + a_2(\lambda_2/\lambda_1)^ku_2 + \ldots + a_n(\lambda_n/\lambda_1)^ku_n\}$$  \hspace{1cm} (2.1.32)

so that for large $k$, the approximation

$$x_k \approx \lambda_1^k(a_1u_1)$$  \hspace{1cm} (2.1.33)

should be good because of (2.1.22).

In order to estimate $\lambda_1$ we replace $k$ by $k+1$ in (2.1.31) to obtain

$$x_{k+1} \approx \lambda_1 x_k$$  \hspace{1cm} (2.1.34)
This suggests a number of ways to approximate $\lambda_1$. For example $x_k'x_{k+1} \approx \lambda_1 x_k'x_k$ and thus

$$\lambda_1 \approx \frac{x_k'x_{k+1}}{x_k'x_k} \quad (2.1.35)$$

Another common procedure is to divide the largest component of $x_k$ into the corresponding element of $x_{k+1}$. If the largest component is the $j$th, then this amounts to

$$\lambda_1 \approx \frac{e_j'x_{k+1}}{e_j'x_k} \quad (2.1.36)$$

where $e_j$ is the $j$th unit-vector. The estimates of $\lambda_1$ can be found by generating the sequence of scalars $\{\beta_k\}$ from the equation

$$\beta_k = v'x_{k+1}/v'x_k \quad (2.1.37)$$

where $v$ is usually $x_k$ or $e_j$.

When programming the power method, it is a little easier to generate the estimate $\beta_k$ from (2.1.35) than (2.1.36) which includes a test at each step to determine the maximum component of $x_k$ (moreover the program will execute faster). The advantage in finding the maximum component of $x_k$ is that round-off errors will tend to be reduced.

Finally, as $k$ increases, the vectors $x_k$ are usually growing quite large if $|\lambda_1| > 1$ (or quite small if $|\lambda_1| < 1$); it is desirable to "scale" or "normalize" $x_k$ to keep within computational limits. This can be done by dividing $x_k$ by $\|x_k\|$ at each step. The process is terminated when $\beta_k$ and $\beta_{k+1}$ are sufficiently close.

**Power method program algorithm.**

First an initial (arbitrary) vector $x_0$ is chosen where
The power method then proceeds for $i=1,2,\ldots$ as follows

(a) Let $y_1 = Ax_{i-1}$  
(b) Set $\beta_1 = y_1'x_{i-1}$  
(c) Let $\eta_1 = \sqrt{y_1'y_1}$  
(d) Set $x_1 = y_1/\eta_1$ and return to step a. We note that $\beta_1$ is the approximation to $\lambda_1$ and that each $x_1$ is an approximate eigenvector.

Remark.

If $\lambda_1 = \lambda_2$, and $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$, then the $\beta_1$ still converge to $\lambda_1$ (that is, multiple eigenvalues do not affect the power method). If $\lambda_1 = -\lambda_2$ and $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$, then the $\beta_1$ would exhibit a periodic behaviour while $v'x_{i+2}/v'x_i$ would provide an estimate for $(\lambda_1)^2$. Dominant complex eigenvalues cause the most problems.

2.2 Age-Structured Model

Introduction.

In the sequel we shall introduce and discuss a discrete time model with a discrete age scale viz, the age-structured model put forward by Leslie (1945). The model is deterministic and predicts the age structure of a population of female animals, given the age structure at some past time (or present) and given the age-specific survival and fecundity rates.

Assumptions.

(i) The age-specific rates remain constant over a period of time.

(ii) Age groups are of equal length.
(iii) The same unit of age is adopted as that of time i.e. time intervals have the same duration as the age intervals.

(iv) No migration.

Formulation of the Matrix Projection Model.

Given any arbitrary age distribution at time \( t \) the female population in one unit time, can be expressed in the form of \( m+1 \) linear equations, where \( m \) to \( m+1 \) is the last age group considered. Let

\[
U_t = (U_{0t}, U_{1t}, \ldots, U_{mt})'
\]  
(2.2.1)

be a column vector representing the population's age structure at time \( t \), where \( u_{it} \) is the number of females alive in the age group \( i \) to \( i+1 \) at time \( t \). A female in the age group \( i \) to \( i+1 \) is described as of age \( i \) and since \( m \) to \( m+1 \) is the last age group considered, none can live to be older than \( m \). The \( m+1 \) elements in the vector (2.2.1) represent \( m+1 \) different age groups. \( u_{t+1} \) is a column vector similar to \( u_t \), but representing the age structure at time \( t+1 \).

Age-specific rates.

Let

\[
p_i = \text{the probability that a female aged } i \text{ to } i+1 \text{ at time } t \text{ will be alive in the age group } i+1 \text{ to } i+2 \text{ at time } t+1. \quad 0 < p_i \leq 1 \text{ for } i=0,1,2,\ldots,m-1. \]  
These values are termed as the survival rates.

\[
f_i = \text{the average number of daughters who will be alive at} \]
time \( t+1 \) in the age group 0-1 born in the interval \( t \) to \( t+1 \) to each female who was in the age group \( i \) to \( i+1 \) at time \( t \). \( f_i \geq 0 \) for \( i = 0,1,2, \ldots, m \) and until the first unit of time has passed their ages will be 0. These values are termed as the *fecundity* rates.

Now, given the age distribution at time \( t=0 \)

\[
\mathbf{u}_0 = (u_{00}, u_{10}, u_{20}, \ldots, u_{m0})'
\]

(2.2.2)

the age distribution at the end of one unit's interval

\[
\mathbf{u}_1 = (u_{01}, u_{11}, u_{21}, \ldots, u_{m1})'
\]

(2.2.3)

is obtained as follows:-

The total number of daughters who will be alive in the age group 0-1 at time \( t=1 \) born by females who were alive in the age group \( i \) to \( i+1 \) at time \( t=0 \) (\( i=0,1, \ldots, m \)) is given by

\[
u_{i0} f_i
\]

(2.2.4)

By definition of \( u_{01} \) we have

\[
u_{01} = \sum_{i=0}^{m} u_{i0} f_i
\]

(2.2.5)

The number of individuals who were in the age group \( i \) to \( i+1 \) at \( t=0 \) who will be alive in the age group \( i+1 \) to \( i+2 \) at \( t=1 \) for \( i=0,1,2, \ldots, m \) is given by

\[
u_{i+1,i} = p_{i} u_{i0}
\]

(2.2.6)

Thus
where in general the \((m+1)\times(m+1)\) matrix \(M\) describes the transition of the population from one age structure to another over one period of time. The matrix \(M\) is called the *Projection matrix*, and often referred to as *Leslie matrix* (cf.\((2.1.15)\)). \((2.2.8)\) defines the *Projection Matrix Model* or the *Age-Structured Model*.

\[
\begin{align*}
\mathbf{u}_1 &= M \mathbf{u}_0 \\
\mathbf{u}_2 &= M \mathbf{u}_1 = M(M \mathbf{u}_0) = M^2 \mathbf{u}_0 \\
\mathbf{u}_t &= M^t \mathbf{u}_0
\end{align*}
\]

(2.2.9)

and in general for \(t=0,1,2,...\)

Properties of the Basic matrix \(M\).

The matrix \(M\) is square of order \(m+1\); it is not necessary, however, in what follows to consider this matrix as a whole. For if \(i=k\) is the largest age within which reproduction occurs, or rather females are sterile in the last \(m-k\) age groups, \(f_k\) is the last \(f_i\) figure which is not equal to zero.
The matrix $M$ can now be partitioned as follows

$$M = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$$

(2.2.11)

The submatrix $A$ is square of order $k+1$, $B$ is of order $(m-k) \times (k+1)$, $C$ is square of order $m-k$ with only non-zero elements in the subdiagonal immediately below the principle diagonal. The remaining submatrix is null and of order $(k+1) \times (m-k)$. On investigating the powers of $M$ it is found that

$$M^t = \begin{bmatrix} A^t & 0 \\ f(ABC) & C^t \end{bmatrix}$$

(2.2.12)

where

$$f(ABC) = \sum_{j=0}^{t-1} C^j B A^{t-j-1}$$

(2.2.13)

$C^0 = A^0 = I$ and since $C^{m-k} = 0 \Rightarrow C^t = 0 \ \forall \ t \geq m-k$  \hspace{1cm} (2.2.14)

In biological terms this is merely an expression of the obvious fact that individuals alive in the post-reproductive ages contribute nothing to the population. The terms of $f(ABC)$ are zero for $j > m-k-1$, and so $f(ABC)$ has only $m-k$ terms, each of which is readily computed if we know the powers of $A$.

$$\lim_{t\to\infty} f(ABC) = \lim_{t\to\infty} (B A^{t-1} + C B A^{t-2} + \ldots + C^{m-k-1} B A^{t-(m-k)})$$

$$= (I + C + \ldots + C^{m-k-1}) B \lim_{t\to\infty} A^t$$
\[(I-C)^{-1}(I-C^{m-k}) B \lim_{t \to \infty} A^t \]
\[(I-C)^{-1} B \lim_{t \to \infty} A^t \quad (2.2.15)\]

so that if \( \lim_{t \to \infty} A^t \) is known, \( \lim_{t \to \infty} M^t \) is easy to find.

It is a submatrix \( A \) of \( M \) which is principally of interest, and in mathematical discussion which follows, attention is focused entirely on it and on age distributions confined to the pre-productive and reproductive age groups.

The matrix \( A \) is of order \((k+1)\times(k+1)\) where \( i=k \) is the last age group in which reproduction occurs, and written in full,

\[
\begin{bmatrix}
  f_0 & f_1 & f_2 & \ldots & f_{k-1} & f_k \\
  p_0 & 0 & 0 & \ldots & 0 & 0 \\
  0 & p_1 & 0 & \ldots & 0 & 0 \\
  \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
  0 & 0 & \ldots & 0 & p_{k-1} & 0 \\
\end{bmatrix}
\]

\[(2.2.16)\]

Therefore we shall only consider the reduced model

\[
\mathbf{u}_{t+1} = A \mathbf{u}_t 
\]

for \( t=0,1,2,\ldots \) where

\[
\mathbf{u}_t = (u_0t, u_1t, \ldots, u_{kt})
\]

is of course now taken to be \((k+1)\)-dimensional. We shall from now henceforth treat the population under study only that part of the total population that is of reproductive age, i.e. we consider \( A \) only.

Repeated premultiplication by \( A \) of the column vector \( u_0 \) permits prediction of the future growth and age distribution of the reproducing part of the population made up of females
aged k or less, i.e.

\[ u_t = A^t u_0 \]  \hspace{1cm} (2.2.19)

**Contribution to future population.**

Moreover, it will be seen that with the help of the \((j+1)\)th column of \(A^t\) we can predict the contribution of some one particular age group, i.e. those of age \(j\) at time \(t=0\) say, to the population at some future time \(t\). This can be obtained as follows.

Let the population projection matrix be \(A\) and \(A^t = (a_{ij})\) \(i,j = 0, 1, 2, \ldots, k\). Then using (2.2.19) we have

\[
\begin{bmatrix}
  a_{00} & a_{01} & \ldots & a_{0k} \\
  a_{10} & a_{11} & \ldots & a_{1k} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{k0} & a_{k1} & \ldots & a_{kk}
\end{bmatrix}
\begin{bmatrix}
  u_{00} \\
  u_{10} \\
  \vdots \\
  u_{k0}
\end{bmatrix}
= 
\begin{bmatrix}
  u_{0t} \\
  u_{1t} \\
  \vdots \\
  u_{kt}
\end{bmatrix}
\]

Showing only those contributions to the elements in \(u_t\) that come from the elements \(u_{j0}\) of \(u_0\) we have

\[
\begin{bmatrix}
  \ldots + a_{0j} u_{j0} + \ldots \\
  \ldots + a_{1j} u_{j0} + \ldots \\
  \ldots + a_{kj} u_{j0} + \ldots 
\end{bmatrix}
= 
\begin{bmatrix}
  u_{0t} \\
  u_{1t} \\
  \vdots \\
  u_{kt}
\end{bmatrix}
\]  \hspace{1cm} (2.2.20)

At time \(t\), provided \(t>k\), every element in \(u_t\) has some contribution from the \(u_{j0}\) individuals that were of age \(j\) at the start. At time \(t\) the number of living individuals contributed to the total population at time \(t\) by this particular age group for \(j=0, 1, 2, \ldots, k\) is
which is the product of the number in the age group \( j \) to \( j+1 \) (referred to as of age \( j \)) at time \( t=0 \) and the sum of the elements in the \( (j+1) \text{th} \) column of \( A^t \). The age group \( j \) to \( j+1 \) is the ancestral age group we are concerned with in knowing the number of individuals it will contribute to the total population at time \( t \). Moreover, it is noticeable that for \( i=0,1,2,\ldots,k \) the number contributed to \( u_{it} \) in particular, is given by

\[
a_{ij}u_{j0}
\]

i.e. the product of the \((i,j)\text{th}\) element of \( A^t \) and the number in the age group \( j \) to \( j+1 \) at time \( t=0 \).

**Backward Projection.**

The matrix \( A \) is non-singular, since the determinant

\[
|A| = (-1)^{k+2} (p_0 p_1 p_2 \ldots p_{k-1} f_k) \neq 0
\]

and therefore there exists an inverse of \( A \).

In order to discover the age composition of a population prior to the time say \( t \) at which it came under observation provided we concern ourselves only with reproducing part of the population, having a projection matrix of the form \( A \) there is no difficulty since \( A^{-1} \) exits. The backward projection is described by the backward series \( A^{-1}u_0, A^{-2}u_0, A^{-3}u_0, \ldots \) e.t.c.
Remark.

Whereas the forward series can be continued indefinitely into the future given any initial age distribution, the backward series can only be performed so long as $u_{1t} \geq 0$, since biologically a negative number of individuals in an age group is meaningless.

Algebraic Properties of the Matrix $A$.

$A$ is square matrix of order $k+1$. $A$ is a non-negative matrix, a characteristic of projection matrices. $A$ is non-singular (see (2.2.23)).

We shall now show that $A$ is irreducible so that the Perron-Frobenius theorem (see P-F thm. sec. 2.1) applies to this matrix. $A$ can be expressed as the sum of two matrices $A_a$ and $A_r$ as follows:

$$A = A_a + A_r \quad \text{i.e.}$$

$$A = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & P_{k-1} & 0 \end{bmatrix} + \begin{bmatrix} f_0 & f_1 & \ldots & f_{k-1} & f_k \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}$$

(2.2.24)

Without loss of generality, we may restrict ourselves to matrix $A_a$ with $f_0 = f_1 = \ldots = f_{k-1} = 0$ and $f_k > 0$, since we can easily show that if $A_a$ is irreducible, then so is a more general matrix $A$, where $A_r$ can have non-zero elements only in the first row.

Suppose that $A_a$ is irreducible and $A$ is reducible, then
there exist a permutation matrix $P$ (see defn.3, sec.2.1) such that

$$P'AP = P'(A_s + A_r)P = P'A_sP + P'A_rP = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}$$

(2.2.25)

where $A_1$ and $A_3$ are square. Since $A_r \geq 0$ then $P'A_rP \geq 0$ and $P'A_sP$ is not of the required form by hypothesis. Hence $A$ is irreducible, since the addition of the non-negative matrix to the matrix not already of the required form cannot result in a sum of that form.

Using (2.1.19), the characteristic equation of $A$ is

$$\lambda^{k+1} - f_0\lambda^k - f_1p_0\lambda^{k-1} - f_2p_0p_1\lambda^{k-2} - \ldots - f_{k-1}p_0p_1\ldots p_k - 2\lambda^{-1} - \ldots = 0$$

(2.2.26)

We can now assert that the matrix $A$ has exactly one positive real root say $\lambda_1$ since on applying Descarte's rule of signs (see defn.6, sec. 2.1) to the characteristic equation (2.2.26) we find that the number of positive real roots is at most one, and since $A$ is non-negative and irreducible, the Perron-Frobenius theorem tells us there is at least one. Hence the remaining roots of $A$ are negative or complex.

Also by Perron-Frobenius theorem the moduli of these remaining roots are at most $\lambda_1$. $\lambda_1$ is referred to as the dominant root and a characteristic vector $u_1$ say, corresponding to $\lambda_1$ is positive (see pf of P.F.thm.sec.2.1)
A is primitive iff the index of primitivity is one. In particular, if \( f_i > 0 \) for all \( i = 1, 2, \ldots, k+1 \), \( h = 1 \) and hence \( A \) is primitive. If \( h > 1 \) then \( A \) is imprimitive (see defn.4, sec.2.1).

We remark that if a projection matrix \( A \) is primitive the powers \( A^n \) approach a limiting form. In fact \( A^n \) (and hence \( u_t \)) will behave periodically if \( A \) is imprimitive with period equal to the index of imprimitivity. Pure periodicity is displayed when the initial age distribution \( u_0 \) is repeated periodically. We suppose that \( h = k+1 \) such that

\[
A^{k+1} = nI
\]  \((2.2.27)\)

i.e. \( nu_0 \) will be repeated periodically with period \( k+1 \).

Cayley Hamilton theorem says that every square matrix satisfies its own characteristic equation therefore from the characteristic equation of \( A \).

\[
A^{k+1} - f_0 A^k - f_1 p A^{k-1} - \ldots - f_{k-1} p \varphi_1 \ldots p_{k-2} A - f_k p \varphi_1 \ldots p_{k-1} I = 0
\]  \((2.2.28)\)

and for \((2.2.27)\) to be satisfied we must have

\[
f_0 = f_1 = \ldots = f_{k-1} = 0 \quad (2.2.29)\)

and

\[
f_k p \varphi_1 \ldots p_{k-1} = n \quad (2.2.30)\)

Expression \((2.2.29)\) results in the special case of the matrix \( A \) having only a single non-zero element in the first row. Whenever \( n = 1 \) in \((2.2.30)\) pure periodicity will be displayed.

The question of approximation to case of pure
periodicity may be of practical importance, for some lower animals with few age groups whose individuals are fertile for only a short period right at the end of their lives e.g. some locust species.

An interesting example of a case in which the requirement of primitivity is not met is provided by a hypothetical insect population (such as beetle), which lives for only three years and which propagates in the third year of life, considered by Bernardelli (1941) and referred to by Leslie (1945). The system of survival and fecundity rates is given by

\[
A = \begin{bmatrix}
0 & 0 & 6 \\
1/2 & 0 & 0 \\
0 & 1/3 & 0 \\
\end{bmatrix}
\]

(2.3.31)

Biologically this can be interpreted as each female in the age group 2-3 produces, on the average, 6 new living females.

The matrix A above is clearly not primitive and the dominant root is not larger than the moduli of all other roots. In fact roots of this matrix are of the same modulus. Any initial age distribution repeats itself regularly every three years.

In the next section we shall particularly be concerned with primitive projection matrices and investigate on their limiting behaviour.
2.3. Stable Population Theory.

Introduction.

Formally it was remarked that if a projection matrix $A$ is primitive then the powers $A^t$ approach a limiting form. In this section we are concerned with the behaviour of the sequence $\{u_t\}$ over long periods of time or the stable population theories in discrete version.

Theorem 1.

Let $A$ be a primitive population projection matrix with dominant eigenvalue $\lambda_1$ and associated (right) positive eigenvector $u_1$. Then

$$\lim_{t \to \infty} (A/\lambda_1)^t = H_1$$

exists, where $H_1$ is a matrix whose columns are positive multiples of $u_1$. ($H_1$ is of rank 1).

Remark.

The requirement that $A$ be primitive ensures that $\lambda_1$ is the only eigenvalue of maximum modulus. To establish the above result (i.e. (2.3.1)) we shall assume that the remaining eigenvalues $\lambda_i$, $(i=2,3,\ldots,k+1)$ of $A$ are distinct and then show that they need not be distinct in order for the above limit to exist.

Proof.

Suppose that $\lambda_i$, $(i=2,3,\ldots,k+1)$ are distinct then $A$ is diagonalizable and is similar to a diagonal matrix $D=\text{diag}(\lambda_1,\lambda_2,\ldots,\lambda_{k+1})$. Since complex roots occur in conjugate pairs, the moduli of these eigenvalues need not be
distinct.

Let $U$ be a matrix whose columns constitute the eigenvectors of $A$. The eigenvalues may be made to appear on the diagonal of $D$ in any desired order by suitable choice of $U$ (see defn. 4, sec. 1.4). Suppose they are arranged in order of decreasing moduli, i.e. $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_{k+1}|$.

$$A = U D U^{-1} \quad (2.3.2)$$

and for the $t^{th}$ power of $A$ we have

$$A^t = U D^t U^{-1}$$

and dividing both sides by $\lambda_1^t$ we have

$$\frac{A^t}{\lambda_1^t} = U \begin{bmatrix} \lambda_1^t & 0 & \ldots & 0 \\ 0 & \lambda_2^t & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \lambda_{k+1}^t \end{bmatrix} U^{-1} \quad (2.3.3)$$

But $|\lambda_i/\lambda_1| < 1$ for $i = 2, 3, \ldots, k+1$ so that

$$\lim_{t \to \infty} \frac{A^t}{\lambda_1^t} = U \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & (\frac{\lambda_2}{\lambda_1})^t & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & (\frac{\lambda_{k+1}}{\lambda_1})^t \end{bmatrix} U^{-1} = H_1$$

$$\lim_{t \to \infty} \frac{A^t}{\lambda_1^t} = U \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} U^{-1} = H_1$$
Hi is a square matrix of rank 1. Moreover, columns of Hi are positive multiples of u1. The fact that columns of Hi are positive multiples will be made clear later.

Suppose now that the remaining eigenvalues \( \lambda_i, i=2,3,\ldots, \) k+1 are not distinct. Let the distinct roots be \( \lambda_1 \) and \( \lambda_i, i=2,3,\ldots,q < k+1 \) and if \( \lambda_1 \) has multiplicity n1 then A is similar to a matrix J in Jordan canonical form i.e. there exists a generalized modal matrix M for A (see defn.8 and thm.5, sec. 1.4) such that

\[
A = M J M^{-1}
\]

where \( J_i \), i=1,2,\ldots, q; is given by (1.4.12) is an \( n \times n \) square matrix with diagonal elements \( \lambda_i, 1's \) on the superdiagonal and zeros elsewhere. The \( t^{th} \) power of A is

\[
A^t = M \begin{bmatrix}
J_1^t & 0 & \cdots & 0 \\
0 & J_2^t & 0 & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & J_q^t
\end{bmatrix} M^{-1}
\]

(2.3.7)

\( J_1 \) can be expressed as
\[ J_1 = \lambda_1 I_n + B_n \] (2.3.8)

where \( I_n \) is the identity matrix of order \( n_1 \) and \( B_n \) is a \( n_1 \times n_1 \) square matrix whose elements are zero everywhere except for ones on the superdiagonal. Noting that \( IB = BI \) and for \( x \geq n_1 \)

\[ B^x = 0 \] (2.3.9)

Using binomial theorem and (2.3.9) (dropping subscript \( n_1 \) for convenience) the expansion of \( J_1^t \) is given by

\[ J_1^t = (\lambda_1 I + B)^t \]

\[ = \sum_{x=0}^{t} \binom{t}{x} (\lambda_1 I)^x B^{t-x} \]

\[ = \sum_{y=0}^{n_1-1} \binom{t}{t-y} (\lambda_1 I)^{t-y} B^y \]

(2.3.10)

Since \( A \) is primitive, \( J_1 = \lambda_1 \) and first column of \( M \) is \( u_1 \), (see defn.8, sec. 1.4 and defn.4, sec.2.1) so that

\[ \lim_{t \to \infty} (J_1/\lambda_1)^t = 1 \] (2.3.11)

For \( i \neq 1 \)

\[ \lim_{t \to \infty} \left( \frac{J_1}{\lambda_1} \right)^t = \sum_{y=0}^{n_1-1} \left\{ \lim_{t \to \infty} \left( \frac{t}{t-y} \right) \left( \frac{\lambda_1 I}{\lambda_1} \right)^{t-y} B^y \right\} \]

(2.3.12)

We note that

\[ \frac{1}{y!} \{ t(t-1) \ldots (t-(y-1)) \} \leq t^y \] (2.3.13)

and for \( \lambda_1 > \lambda_1 \)

\[ \theta_1 = \log_\lambda \lambda_1 - \log_\lambda \lambda_1 = \log_\lambda (\lambda_1/\lambda_1) > 0 \] (2.3.14)
\[ (\lambda_i/\lambda_1)^t = \exp(-\theta_1 t) \] (2.3.15)

Using (2.3.13/14/15) we have

\[ \lim_{t \to \infty} (J_i/\lambda_1)^t = 0 \] (2.3.16)

and finally using (2.3.11/16)

\[
\lim_{t \to \infty} \frac{A^t}{\lambda_1^t} = M \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} M^{-1} = H_1
\] (2.3.17)

Therefore, regardless of \( \lambda_1, i=2,3,..., k+1 \) being not distinct, \( A^t/\lambda_1^t \) still converges to \( H_1 \).

On investigation it is found that

\[ H_1 = u_1v_1' \] (2.3.18)

where \( u_1 \) is the right positive eigenvector and \( v_1' \) is the left positive eigenvector of \( A \) both corresponding to \( \lambda_1 \). Consequently, since \( v_1' \) is a positive vector then the columns of \( H_1 \) are positive multiples of \( u_1 \).

This result will be established through a lemma but first we shall examine

\[ \lim_{t \to \infty} A^t z \] (2.3.19)

where \( z \) is an arbitrary \( k+1 \) vector. Suppose that the eigenvalues of \( A \) are \( \lambda_i \), the corresponding eigenvectors \( u_i, i=1,2,..., k+1 \) and \( u_1 \) form a basis for \( \mathbb{R}^{k+1} \). Then we can write

\[ z = \sum_{i=1}^{k+1} x_i u_i \]
where $\alpha_1$ are constants not all zero.

Clearly as $t$ gets large this sum will be dominated by the term involving the eigenvalue with the largest absolute value (modulus), provided that the associated coefficient is not zero. Let us assume that this eigenvalue has the index $i=1$, so that we can write for large $t$

$$A^t z \approx \alpha_1 \lambda_1 t u_1$$

(2.3.21)

or equivalently

$$\lim_{t\to\infty} A^t z = \alpha_1 \lambda_1 t u_1$$

(2.3.22)

Lemma.

Let $A$ be an $n \times n$ matrix, $x$ an $n$-vector, and $\alpha, \lambda$ real numbers. Then for all $z \in \mathbb{R}^n$

$$\lim_{t\to\infty} A^t z = \alpha \lambda^t x$$

(2.3.23)

iff

$$\lim_{t\to\infty} A t = \lambda^t x y'$$

(2.3.24)

$y \in \mathbb{R}^n$ is arbitrary and $\alpha = y'z$

Proof.

Since $z$ is independent of $t$

$$\lim_{t\to\infty} A^t z = (\lim_{t\to\infty} A t) z$$

$$= (\lambda^t x y') z$$

using (2.3.24)

$$= \lambda^t x y' z$$

$$= \alpha \lambda^t x$$

conversely

$$\lim_{t\to\infty} A^t z = \alpha \lambda^t x \quad \text{where} \quad \alpha = y' z$$

$$= \lambda^t x \alpha$$
so that

$$\lim_{t \to \infty} A^t = \lambda^t x y'$$

which proves sufficiency.

Now suppose that

$$\lim_{t \to \infty} A^t = \lambda^t x y' + E$$  \hspace{1cm} (2.3.25)

where $E \neq 0$ is an $n \times n$ matrix then

$$\lim_{t \to \infty} A^t z = (\lambda^t x y' + E) z$$

$$= a \lambda^t x + E z$$  \hspace{1cm} (2.3.26)

for arbitrary $z$ we must therefore have $E = 0$, which proves necessity.

We can now use the lemma to establish (2.3.8) where $A$ is a primitive population projection matrix with dominant eigenvalue $\lambda_1$ and associated positive right eigenvector $u_1$. Noting that $A$ and $A'$ have the same eigenvalues (see thm.1, sec.1.4). Let

$$z = \sum_{i=1}^{k-1} \beta_i v_i$$  \hspace{1cm} (2.3.27)

where $v_1'$ are the right eigenvectors of $A'$ corresponding to $\lambda_1$ or equivalently $v_1'$ are the left eigenvectors of $A$. $\beta_1$'s are constants not all zero.

Using (2.3.22) we have

$$\lim_{t \to \infty} (A')^t z = \beta_1 \lambda_1 t v_1$$  \hspace{1cm} (2.3.28)

and using lemma
\[ \lim_{t \to \infty} (A')^t = \lambda_1^tv_1w' \] (2.3.29)

where \( \beta = w'z \) and since \((A')^t = (A^t)'\) we have

\[ \lim_{t \to \infty} (A')^t = (\lim_{t \to \infty} A^t)' \] (2.3.30)

From (2.3.22) and lemma

\[ (\lim_{t \to \infty} A^t)' = (\lambda_1^txu_1)' \] (2.3.31)

where \( x \in \mathbb{R}^{k+1} \) is arbitrary and \( a_1 = x'z \).

Using (2.3.29/30/31) we can take \( x = v_1 \) and \( w = u_1 \) so that

\[ \lim_{t \to \infty} = \lambda_1^txu_1' = \lambda_1^tv_1v_1' \] (2.3.32)

or equivalently

\[ \lim_{t \to \infty} A^t/\lambda_1^t = u_1v_1' \] (2.3.33)

where \( u_1, v_1' \) are respectively the right (left) positive eigenvectors of \( A \) both corresponding to the dominant eigenvalue \( \lambda_1 \). Using (2.3.5/33) then (2.3.18) follows.

Considering the spectral representation of the diagonalizable matrix \( A \) demonstrated by (2.3.18) we have

\[ A = \sum_{i=1}^{k+1} \lambda_i u_i'v_i' \] (2.3.34)

where \( u_1, v_1' \) are respectively the right(left) eigenvectors of \( A \) both corresponding to the eigenvalue \( \lambda_1 \), \( (i = 1, 2, \ldots, k+1) \) so that

\[ H_1 = u_1v_1' \] (2.3.35)

Using (2.3.35) we can write (2.3.34) as
and model (2.2.19) can be rewritten as

\[ A = \sum_{i=1}^{k+1} \lambda_i H_i \]  

(2.3.36)

\[ u_c = \lambda_1 \sum_{i=1}^{k+1} \left( \frac{\lambda_i}{\lambda_1} \right)^t H_i u_0 \]

(2.3.37)

i.e. the population is growing *approximately geometrically* so long as \((\lambda_1/\lambda_1)^t\) is close to zero.

**Theorem 2.**

For any non-negative vector \(x\) the relative values of the components of

\[ y = H_1 x \]

(2.3.38)

are independent of \(x\).

**Proof.**

\[ y = H_1 x \]

\[ = (m_1 u_1, m_2 u_1, \ldots, m_{k+1} u_1) x \]

\[ = u_1 m' x \]  

(2.3.39)

where \(m=(m_1, m_2, \ldots, m_{k+1})'\).

The relative values of the components of \(y\) in vector form is given by

\[ y(1'y)^{-1} = u_1 m' x (1' u_1 m' x)^{-1} \]

\[ = u_1 m' x (m' x)^{-1} (1' u_1)^{-1} \]

\[ = u_1 (1' u_1)^{-1} \]  

(2.3.40)
which is independent of $x$ and hence the result.

Applying theorem 1 to (2.2.19) we note that

$$\lim_{t \to \infty} ut/\lambda_1^t = \left( \lim_{t \to \infty} A^t/\lambda_1^t \right) u\varnothing$$

$$= H_1 u\varnothing$$

$$= c \zeta \quad (2.3.41)$$

The age distribution vector $\xi_t = ut (1'ut)^{-1}$ approaches a limiting value $\zeta$ which is independent of the initial population $u\varnothing$, depending only on the matrix $A$ and satisfies

$$1'\zeta = 1 \quad (2.3.43)$$

$A$ determines $\lim_{t \to \infty} ut$ to within a scalar $c$ which depends on both $u\varnothing$ and $A$.

$$\lim_{t \to \infty} ut/\lambda_1^t (1'ut/\lambda_1^t)^{-1} = \lim_{t \to \infty} ut (1'ut)^{-1} = \zeta \quad (2.3.44)$$

Applying theorem 2 to (2.3.41)

$$\lim_{t \to \infty} ut/\lambda_1^t (1'ut/\lambda_1^t)^{-1} = u_1 (1'u_1)^{-1} \quad (2.3.45)$$

From (2.3.44) & (2.3.45)

$$\zeta = u_1 (1'u_1)^{-1} \quad (2.3.46)$$

and clearly satisfies (2.3.43).

The limiting vector $\zeta$ is called the stable age distribution or stable population vector corresponding to $A$, and the fact that it is independent of $u\varnothing$ is known as the ergodic property for population growth (i.e. it forgets its past).

It is observed that as $t$ gets large the matrix $A^t$ displays a very interesting property i.e. the ratio of an element in $A^{n+1}$ to the corresponding element in $A^n$ will for large $n$ tend to $\lambda_1$ and when the population has stabilized
this ratio is $\lambda_1$.

From (2.2.19), this is equivalent to saying that when the population has stabilized

$$u_{t+1} = \lambda_1 u_t$$

(2.3.47)

and by the definition of the model (2.2.17) we conclude that

$$Au_t = \lambda_1 u_t$$

(2.3.48)

which has non-trivial solution iff

$$|A - \lambda_1 I| = 0$$

(2.3.49)

or equivalently iff the characteristic equation of $A$ has dominant root $\lambda_1$.

We can now establish (2.3.47) by supposing that it holds and show that the characteristic equation of $A$ has dominant root $\lambda_1$.

Proof.

Let $Y_0, Y_1, Y_2, \ldots, Y_k$ be a set of frequencies such that

$$u_t = (Y_0, Y_1, Y_2, \ldots, Y_k)'$$

(2.3.50)

and by hypothesis

$$u_{t+1} = (\lambda_1 Y_0, \lambda_1 Y_1, \lambda_1 Y_2, \ldots, \lambda_1 Y_k)'$$

(2.3.51)

Using model (2.2.17)

$$u_{t+1} = [\xi_1 f_1 Y_1, \rho_0 Y_0, p_1 Y_1, \ldots, p_{k-1} Y_{k-1}]'$$

(2.3.52)

It is clear from (2.3.51) & (2.3.52) that

$$\lambda_1 Y_0 = f_0 Y_0 + f_1 Y_1 + f_2 Y_2 + \ldots + f_k Y_k$$

(2.3.53)

and for $i=1, 2, \ldots, k$

$$Y_i = \rho_0 \rho_1 \rho_2 \ldots \rho_{i-1} Y_0 / \lambda_1^i$$

On substituting $Y_i$, $i=1, 2, \ldots, k$ into (2.3.53) we obtain

$$\lambda_1^{k-1} - f_0 \lambda_1^k - f_1 \rho_0 \lambda_1^{k-1} - f_2 \rho_0 \rho_1 \lambda_1^{k-2} - \ldots - f_{k-1} \rho_0 \rho_1 \ldots \rho_{k-1} = 0$$
which is the characteristic equation of $A$ given by (2.2.26) in dominant root $\lambda_1$ and hence the result.

It is clear from the established relation (2.3.47) that the dominant root $\lambda_1$ gives the rate of increase of the population in a single step; so that the population grows geometrically. It is related to Lotka's $r$(intrinsic rate of natural increase) by expression (1.2.6). $\lambda_1$ is also called population finite rate of natural increase.

If the dominant root $\lambda_1$ is greater than unity implies that the population is capable of increasing. If $\lambda_1$ is less than unity then the population is capable of decreasing. If $\lambda_1$ is unity then the population remains constant and is often referred to as stationary population.


Introduction.

Before we focus specifically on harvesting we wish to briefly study or rather examine stochastic matrices and establish a relationship with a non-negative matrix. The relationship will be utilized in a proof of a theorem on positive harvesting. We shall consider harvesting problem in terms of matrix algebra with the intention of maximizing yields subject to maintaining a constant population size and age structure after each time period.

Stochastic matrices.

Defn.1.
A square matrix, $P=(p_{ij})$ of order $n$ is called stochastic if the matrix $P$ is non-negative and if every row sum is equal to 1.

Thus, the matrix of transitional probabilities for a homogeneous Markov chain is stochastic; and conversely, an arbitrary stochastic matrix may be taken to be the matrix of transition probabilities of some homogeneous Markov chain.

A stochastic matrix is a special case of a non-negative matrix. Special addition properties of a stochastic matrix are the following. It follows from the definition, that such a matrix has characteristic root 1, to which the positive characteristic vector $z=(1,1,...,1)'$ corresponds. Conversely, it is clear that if the corresponding characteristic root is 1, then the matrix $P$ is stochastic. Further, the root 1 is the dominant characteristic root of every stochastic matrix, since the dominant characteristic root always lies between the largest and smallest row sums and each row sum is 1. These facts are restated yielding definition 2 of a stochastic matrix.

Defn. 2.

A non-negative matrix $P (\geq 0)$ is stochastic iff the vector $(1,1,...,1)'$ is a characteristic vector of $P$, with corresponding characteristic root 1. For a stochastic matrix, $1$ is the dominant characteristic root.

Now we establish a representation for certain non-negative matrices. Let $A=(a_{ij})$ be a non-negative square
matrix of order n; let the positive root r (>0) be a
characteristic root of A, and let the corresponding
characteristic vector z=(z_1, z_2, ..., z_n) (>0) be positive.

\[ A z = r z \]  \hspace{1cm} (2.4.1)

so that for i=1,2,...,n

\[ \sum_{j=1}^{n} a_{ij} z_j = r z_i \]  \hspace{1cm} (2.4.2)

We define diagonal matrix Z by

\[ Z = \text{diag}(z_1, z_2, ..., z_n) \]  \hspace{1cm} (2.4.3)

and the matrix P=(p_{ij}) by

\[ P = r^{-1}Z^{-1}AZ \]  \hspace{1cm} (2.4.4)

The elements p_{ij} of P are clearly non-negative:

\[ p_{ij} = r^{-1}z_i^{-1} a_{ij} z_j \geq 0 \]  \hspace{1cm} (2.4.5)

for i,j=1,2,...,n and using (2.4.2.) we have the relation

\[ \sum_{j=1}^{n} p_{ij} = 1 \]  \hspace{1cm} (2.4.6)

for i=1,2,...,n.

We have established the following definition,

Defn.3.

If A (>0) is a non-negative matrix which has a positive
dominant characteristic root, to which there corresponds a
positive characteristic vector z=(z_1, z_2, ..., z_n) (>0), it is
similar to the product rP of r by some stochastic matrix P.
A = Z(rP)Z^{-1} \quad (2.4.7)

where Z is given by (2.4.3).

Remark 1.
If \( r = 1 \) then A is similar to a stochastic matrix P.

Remark 2.
If A is primitive then P is also primitive.

Defn. 4.
We call a stochastic matrix P and the corresponding homogeneous markov chain proper if the matrix P has no characteristic roots \(*1\) of modulus 1, and call a matrix of chain regular if it is proper and if 1 is a simple characteristic root (simple root of the characteristic equation) of P.

Remark.
A homogeneous markov chain is called aperiodic if the corresponding stochastic matrix P is primitive.

If P is the matrix of transition probabilities of a homogeneous markov chain, then the matrix \( P^m \) of limiting probabilities exists iff the chain is proper.

\[ P^m = \lim_{m \to \infty} P^m \quad (2.4.8) \]

Remark.
The rows of the matrix \( P^m \) add to unity, and as they contain a fixed number of elements it is impossible that \( p_{ij}^m \to 0 \) \( \forall \) pairs \( i,j \) (see Feller Vol.1, Pg. 392).

Harvesting.

Under favourable conditions, natural animal populations
have a tendency to increase in numbers. It is, however, possible to remove some animals and maintain a constant population size. If the population is divided into groups by age, different proportions of the various groups can be harvested subject to the condition that a fixed population size and age structure is maintained after every time interval.

The problem of harvesting was introduced to matrix population theory by Lefkovitch and Williamson in 1967. If the population is distributed by ages according to the stable age distribution \( \zeta \), the latent vector corresponding to \( \lambda_1 \) then

\[ A \zeta = \lambda_1 \zeta \quad (2.4.9) \]

and it is possible to remove the vector

\[ (\lambda_1 - 1) \zeta \quad (2.4.10) \]

from the population after it has reproduced, restoring the initial age structure and population size. This corresponds to harvesting equal fractions \( (\lambda_1 - 1)/\lambda_1 \) from each age group or equivalently \( 100(\lambda_1 - 1)/\lambda_1 \% \) of the total population. This is called uniform harvesting.

It is required that the population size and age structure be restored after each time period (this is referred to as sustainable harvesting). Subject to this constraint, an age structure and harvesting policy may be chosen to maximize some value attached to the harvest (the number of animals harvested, the biomass harvested, etc.)
Both authors enquired whether some harvesting policy other than harvesting equal proportions from each group might produce a higher yield. Neither author attempted to solve the problem generally, but Williamson demonstrated that the above mentioned harvesting procedure is not always best (optimal).

He considers the matrices

\[
A_1 = \begin{bmatrix} 0 & 9 & 12 \\ 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 3 & 36 \\ 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \end{bmatrix}
\]

(2.4.11)

\[A_1\] and \(A_2\) both have \(\lambda_1 = 2\). The stable age structure for both is \(\xi = 1/29(24,4,1)'\). The yield

\[1'(Au - u) \quad (2.4.12)\]

is 1 per member of the population when \(u = \xi\). However if \(u = 1/9(6,2,1)'\) is used instead of \(\xi\) a yield of 8/3 per member of the population is possible for \(A_1\) and a yield of 4 per member for \(A_2\).

Harvesting problem in terms of matrix algebra.

For harvesting after reproduction, it is required to find a non-negative column vector \(u\) and a non-negative diagonal matrix

\[R = \text{diag}(r_1,r_2,\ldots,r_n) \quad (2.4.13)\]

such that

\[RAu = u \quad (2.4.14)\]

and
\[ c'(I - R) A u \quad (2.4.15) \]
is maximized over \( u \) and \( R \). \( u \) is the chosen distribution of ages, the diagonal elements of \( R \) are the proportions of various groups remaining in the population after harvesting, \( c'=(c_1,c_2,\ldots,c_n) \) is the set of weights given to the groups, (2.4.14) is the requirement that the initial age structure and population size be restored at the end of each cycle of reproduction and harvesting.

For harvesting before reproduction the value of the harvest is
\[ c'(I - R) u \quad (2.4.16) \]
and the initial age distribution and population structure is restored by the constraint
\[ A R u = u \quad (2.4.17) \]

Intuitively, it seems obviously better to harvest after rather than before reproduction. Without loss of generality the total population size can be constrained by the relation
\[ 1'u = 1 \quad (2.4.18) \]
so that the problem can be restated as that of maximizing (2.4.15) subject to (2.4.14),(2.4.18) and \( Au-u \geq 0 \), where \( u \geq 0 \) or equivalently we are maximizing
\[ c'(Au - u) \quad (2.4.19) \]
for
\[ Au - u \geq 0, \quad u \geq 0 \text{ and } 1'u = 1 \quad (2.4.20) \]
This is a linear programming problem which can be solved by the simplex method.
Since it is possible in some cases to increase the yield by not choosing \( u = \zeta \) (demonstrated by (2.4.11)), we may ask whether it is necessary that \( \lambda_1 > 1 \) for a positive harvest to be possible. If \( A \) is primitive (some power of \( A \) has positive entries)(see thm 3, sec. 2.1) then it is necessary that \( \lambda_1 > 1 \). However, if \( \lambda_1 = 1 \) and \( A \) is not primitive a positive harvest may be possible. An example is any Leslie matrix \( A \) with a zero fertility in the last age group. The last age group can be harvested completely in this case with \( u = \zeta \) in the remaining components.

Theorem.

If the dominant root, \( \lambda_1 \), of \( A \) is less than 1, no positive harvest is possible. If \( \lambda_1 = 1 \) and \( A \) is primitive, no positive harvest is possible.

Proof. (By contradiction).

For a positive harvest to be possible there must exist a \( u \geq 0 \) such that \( A u \geq u \) and strict inequality holds for at least one component. It follows that

\[ A^m u \geq u \quad (2.4.21) \]

for any positive integer \( m \).

(i) If \( \lambda_1 < 1 \) then \( A^m \to 0 \) as \( m \to \infty \) (all entries tend to zero). Therefore, there is no non-zero \( u \) satisfying \( A u \geq u \) and no positive harvest is possible.

(ii) If \( \lambda_1 = 1 \) and \( A \) is primitive, then \( u_1 \) is strictly positive. Also \( A \) is similar to a primitive stochastic matrix \( P \) (see rem.1 of defn.3,sec.2.4).
where \( Z \) is a diagonal matrix whose diagonal elements are the components of \( u_1 \).

Since \( P \) is primitive (regular in Markov chains, or aperiodic in homogeneous Markov chains) (see rem. 2 of defn. 3, sec. 2.4)

\[
\lim_{m \to \infty} P^m = P^\infty
\]  

exists and has all entries positive (see rem. of (2.4.8)). Therefore

\[
\lim_{m \to \infty} A^m = Z \left( \lim_{m \to \infty} P^m \right) Z^{-1} = Z P^\infty Z^{-1}
\]

also exists and has all entries positive.

For any \( u \),

\[
\lim_{m \to \infty} A^m(A u - u) = \lim_{m \to \infty} A^{m+1} u - \lim_{m \to \infty} A^m u
\]

\[= 0\]

But if \( A u - u \) has at least one positive component, then

\[
\lim_{m \to \infty} A^m(A u - u) > 0
\]

i.e. is strictly positive. Therefore no positive harvest is possible.

Remark.

For the special case when \( u = \zeta \), i.e. harvesting equal proportions of \((\lambda_1 - 1)/\lambda_1\) from each age group (uniform harvesting), it is quite clear that a positive harvest is not possible whenever \( \lambda_1 \leq 1 \).

In the next chapter we shall consider practical applications of the model we have described.
3.1 Introduction

We shall apply the model to a population of a gorilla species (Mountain gorilla) (see last paragraph of sec. 1.1 for source of data). The central problem is to determine the stable population vector, the intrinsic rate of natural increase and the harvest that can be taken subject to the condition that a fixed population size and age-structure is maintained after every time interval. With knowledge of $\lambda_1$, the dominant root of the matrix $A$, all the above three are readily obtainable from the appropriate formulae; namely

(i) The stable vector is given by

$$\zeta = u_1 (1' u_1)^{-1}$$

where

$$u_1 = (1, \rho_0/\lambda_1, \rho_0 \rho_1/\lambda_1^2, \ldots, \rho_0 \rho_1 \ldots \rho_{k-1}/\lambda_1^k)'$$

is the eigenvector corresponding to $\lambda_1$.

(ii) The intrinsic rate $r$ of natural increase (Lotka's $r$) is given by

$$r = \log_e \lambda_1$$

(iii) The harvest that can be taken so that a constant population vector is maintained after every time interval is given by

$$H = 100(\lambda_1 - 1)/\lambda_1$$

where $H$ is expressed as a percentage of the total population i.e. harvesting a proportion of $(\lambda_1 - 1)/\lambda_1$ from each age
group. It is remarkable to note that this harvesting policy may or may not be optimal. If it is not optimal we can alternatively maximize (2.4.19) subject to (2.4.20).

3.2. Application to generate and maintain $\zeta$

The general information includes the assumption that

(a) There is

(i) No migration, (ii) No density dependence

(b) Age-specific rates remain constant over a period of time.

(c) Same unit of age is adopted as that of time.

The sex ratio is 0.500 and the population is partitioned into 5-yr age groups and a maximum age of 6 will be considered i.e. we shall have 7 age groups. The initial age distribution and the age-specific rates are as tabulated below.

<table>
<thead>
<tr>
<th>Age</th>
<th>Initial</th>
<th>Fecundity</th>
<th>Surviva</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>80</td>
<td>0.00</td>
<td>0.682</td>
</tr>
<tr>
<td>1</td>
<td>57</td>
<td>0.00</td>
<td>0.896</td>
</tr>
<tr>
<td>2</td>
<td>46</td>
<td>1.15</td>
<td>0.730</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
<td>1.15</td>
<td>0.730</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>1.15</td>
<td>0.730</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>1.15</td>
<td>0.730</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>1.15</td>
<td>0.730</td>
</tr>
</tbody>
</table>
so that the initial age distribution is given by
\[ u_0 = (80, 57, 46, 34, 25, 19, 14)' \]  

(3.2.1)

and the projection matrix \( A \) is given by
\[
A = \begin{bmatrix}
0 & 0 & 1.15 & 1.15 & 1.15 & 1.15 & 1.15 \\
0.682 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.896 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.730 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.730 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.730 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.730 & 0 \\
\end{bmatrix}
\]  

(3.2.2)

Thus, the model is described by
\[ u_t = A^t u_0 \]  

(3.2.3)

for \( t=1,2,3, \ldots \).

The non-negativity of \( A \) is obvious. The irreducibility of \( A \) follows as already proved for a general matrix. It only remains to find the index of imprimitivity of \( A \). The characteristic equation of \( A \) is given by (2.2.26) for \( k+1=7 \) i.e.
\[
\lambda^7-0.7027328\lambda^6-0.512894944\lambda^5-0.374486309\lambda^4-0.273375005\lambda^3-
0.199563754 = 0
\]  

(3.2.4)

The index of imprimitivity \( h \) is the g.c.d. of the differences \({7-4, 4-3, 3-2, 2-1, 1-0}\). Here \( h=1 \), hence \( A \) is primitive which implies that the dominant root \( \lambda_1 \) of \( A \) is greater than the moduli of all other roots and
\[
limit_{t\to\infty} A^t/\lambda_1^t
\]

exists. In other words there exist a stable population vector \( \zeta \).
The power method procedure is used to determine the dominant root $\lambda_1$. The algorithm is on page 34 and a fortran program utilizing this method is in the appendix. The dominant root is found to be

$$\lambda_1 = 1.1856588 \quad (3.2.6)$$

Letting $p_i = p_0, p_1, p_2, \ldots, p_i, \quad i=0,1,2,3,4,5.$ we have

$$u_1 = (1, 0.575207639, 0.43468327, 0.26630778, 0.164777985, 0.10152398, 0.062463375)' \quad (3.2.7)$$

$$\lambda_1 u_1 = (1.1856588, 0.681999999, 0.515386044, 0.317318787, 0.195370468, 0.120287928, 0.07406025)' \quad (3.2.8)$$

$$A u_1 = (1.185658977, 0.682, 0.515386044, 0.317318787, 0.195370467, 0.120287929, 0.07406025)' \quad (3.2.9)$$

(3.2.7) & (3.2.8) confirms that $u_1$ is an eigenvector of $A$ corresponding to $\lambda_1$. The stable population vector is

$$\zeta = u_1/(1' u_1)$$

$$= u_1/2.606215445$$

$$= (0.383.313, 0.220706097, 0.166787158, 0.102689429, 0.063225005, 0.038927095, 0.02396708') \quad (3.2.10)$$

and satisfies

$$1' \zeta = 1 \quad (3.2.11)$$

The intrinsic rate of natural increase $r$ is

$$r = \log_e(1.1856588)$$

$$= 0.170298569$$

$$\approx 0.1703 \quad (3.2.12)$$

Noting that $\lambda_1 > 1$ a positive harvest is possible.
Subject to the condition of maintaining a constant population size and age-structure after each time period we can harvest a proportion of

\[
\frac{\lambda_1 - 1}{\lambda_1} = 0.156587038
\]

\[
\approx 0.1566
\]

(3.2.13)

from each age group i.e 15.66% of the total population. This harvesting policy is equivalent to using (2.4.19) by taking \( c' = (1,1,\ldots,1) \), \( u = \zeta \) so that the yield is

\[
c'(Au-u) = 1'(A\zeta-\zeta)
\]

\[
= (\lambda_1 - 1)1'\zeta
\]

\[
= \lambda_1 - 1 \text{ (using 3.2.11)}
\]

\[
= 1.1856588 - 1
\]

\[
= 0.1856588
\]

(3.2.14)

per member of the population.

Carrying out the above harvesting policy after reproduction will restore the stable population vector \( \zeta \) after each time period.

3.3 Concluding Remarks.

We have established that when the population has stabilized the model \( u_t = A u_{t-1} \) reduces to \( u_t = \lambda_1 u_{t-1} \) where \( \lambda_1 \) is the dominant eigenvalue of the projection matrix A, so that the population grows geometrically. Moreover, the eigenvector \( u_1 \) of A corresponding to the eigenvalue \( \lambda_1 \) is a function of the survival rates \( p_i \ (i=0,1,\ldots, k-1) \) and \( \lambda_1 \).

It is also found that when the population has not
stabilized the population grows approximately geometrically so long as \((\lambda_1/\lambda_1)^t\) is close to zero\((i=2,3,\ldots,k+1)\)

For the Gorilla data \(\lambda_1 = 1.1856588 > 1\) i.e. the population is capable of increasing at a constant rate of 1.1856588 in a single step and a positive harvest is possible. To maintain the stable population vector \(\xi\) after every time period we should harvest a proportion of \((\lambda_1 - 1)/\lambda_1 = 0.1566\) from each age group after reproduction.

The model we have considered is often referred to as the classical (traditional) model. The projections are intended only to indicate the future course of population growth if the present trends in vital rates continue. The predictions are therefore not exact if time intervals are long, however, for short intervals of time the approximations are good.

The matrix is defined as operating over a period of time, and thus the model can take the structure of the ecosystem at time \(t\), and predict the new structures at times \(t+1, t+2, t+3, \ldots\). There is no way of using the model to predict what the structure would be at time \(t+(1/2)\), or any non-integer value of time. However, many biological processes occur in discrete periods of time and are not continuous in their operations.

It has been observed in practise that population growth is a highly variable process and that the projections using classical models are quickly invalidated. Any results
obtained using this model are only as valid as the underlying assumptions made.

Models should incorporate insight into the biological mechanisms involved as well as being sufficiently general for the model to be applicable in a variety of situations merely by changes in the parameters or modifications of the underlying assumptions. Such models are called robust models.

If there is reason to suspect a violation of any of the postulates inherent in the construction of the model, or if sufficient information to judge their validity is not available then a more general model which includes variable factors is desirable (refer literature review, (sec. 1.3)).

It is possible to develop models which represent specific growth events with a certain degree of precision but it is much more difficult to develop models with a general range of applicability, however, the extensions of the classical model (refer literature review, (sec. 1.3)) has demonstrated its robustness.
APPENDIX

PROGRAM DEVALU

DIMENSION A(20,20), X(20)

REAL LAMBDA

OPEN( UNIT = 5, FILE = 'INPUT.', STATUS = 'OLD')
OPEN( UNIT = 2, FILE = 'OUTPUT', STATUS = 'NEW')

READ(5,10) ((A(I,J),I = 1,7),J = 1,7)
READ(5,20) (X(I), I = 1,7)

10 FORMAT(7F5.3)
15 FORMAT(/,7F8.3)
20 FORMAT(7F9.7)
25 FORMAT(7(F13.10//))
30 FORMAT(5X,'INITIAL ARBITRARY VECTOR OF UNIT NORM/')
35 FORMAT(5X,'PROJECTION MATRIX A')
40 FORMAT(5X,'LIMITING VECTOR OF UNIT NORM/')
45 FORMAT(4X,'******************************************
TOL = 1E-10
N = 7
M = 10000
WRITE(*,30)
WRITE(*,25)(X(I),I=1,7)
WRITE(2,30)
WRITE(2,45)
WRITE(2,25)(X(I),I=1,7)
WRITE(*,35)
WRITE(*,15)((A(I,J), I= 1,7),J= 1,7)
CALL POWERM(A,X,LAMBDA,TOL,N,M,ITERM)
50 FORMAT(/,' DOMINANT EIGENVALUE IS ',F15.12///)
WRITE(2,45)
WRITE(2,35)
WRITE(2,15) ((A(I,J),I=1,7),J=1,7)
WRITE(*,50) LAMBDA
WRITE(2,45)
WRITE(2,50) LAMBDA
WRITE(*,40)
WRITE(2,40)
WRITE(*,25)(X(I),I=1,7)
WRITE(2,25)(X(I),I=1,7)
IF(ITERM.EQ.2) GO TO 60
WRITE(2,55)
55 FORMAT(5X,'CONVERGENCE DUE TO TOLERANCE')
GO TO 70
60 WRITE(2,85)M
65 FORMAT(5X,' CONVERGENCE AFTER' ,I4,' ITERATIONS')
70 STOP
END
SUBROUTINE POWERM(A,X,LAMBDA,TOL,N,M,ITERM)
REAL LAMBDA
DIMENSION A(20,20),X(20),Y(20)

C SUBROUTINE POWERM USES THE POWER METHOD WITH SCALING TO ESTIMATE THE DOMINANT EIGENVALUE OF A MATRIX A. THE CALLING PROGRAM MUST SUPPLY THE MATRIX A, AN INITIAL VECTOR X WITH EUCLIDEAN LENGTH 1, A TOLERANCE TOL, AN INTEGER N WHERE A IS (NXN) AND AN INTEGER M = MAXIMUM NUMBER OF POWER ITERATIONS DESIRED. THE SUBROUTINE RETURNS WHEN THE DIFFERENCE OF TWO SUCCESSIVE ESTIMATES IS LESS THAN TOL IN ABSOLUTE VALUE OR WHEN M ITERATIONS HAVE BEEN EXECUTED. IN THE FIRST CASE, A FLAG, ITERM IS SET TO 1 AND IN THE SECOND CASE ITERM IS SET TO 2. THE APPROXIMATE EIGENVALUE IS RETURNED AS LAMDA AND AN APPROXIMATE EIGENVECTOR AS X. LAMBDA MUST BE DECLARED REAL IN THE CALLING PROGRAM.

ITR=1

CALCULATE THE INITIAL EIGEN VALUE APPROXIMATION

DO 1 I=1,N
   Y(I)=0.

DO 1 J=1,N
1 \( Y(I) = Y(I) + A(I,J) \times X(J) \)

\( \text{TEMP} = 0. \)

\( \text{YScale} = 0. \)

DO 2 \( I = 1, N \)

\( \text{TEMP} = \text{TEMP} + Y(I) \times X(I) \)

2 \( \text{YScale} = \text{YScale} + Y(I) \times Y(I) \)

\( \text{YScale} = \sqrt{\text{YScale}} \)

\( \text{ESTold} = \text{TEMP} \)

C

POWER METHOD ITERATION WITH SCALING

3 \( \text{ITR} = \text{ITR} + 1 \)

DO 4 \( I = 1, N \)

4 \( X(I) = Y(I) / \text{YScale} \)

DO 5 \( I = 1, N \)

\( Y(I) = 0. \)

DO 5 \( J = 1, N \)

5 \( Y(I) = Y(I) + A(I,J) \times X(J) \)

\( \text{TEMP} = 0. \)

\( \text{YScale} = 0. \)

DO 6 \( I = 1, N \)

\( \text{TEMP} = \text{TEMP} + Y(I) \times X(I) \)

6 \( \text{YScale} = \text{YScale} + Y(I) \times Y(I) \)

\( \text{YScale} = \sqrt{\text{YScale}} \)

\( \text{ESTnew} = \text{TEMP} \)
TEST FOR TERMINATION OF THE POWER METHOD ITERATION

IF(ABS(ESTNEW-ESTOLD).LE.TOL) GO TO 7
IF(ITR.GE.M) GO TO 8
ESTOLD=ESTNEW
GO TO 3

7 ITERM=1
   LAMBDA=ESTNEW
   RETURN

8 ITERM=2
   LAMBDA=ESTNEW
   RETURN
END
OUTPUT FILE

INITIAL ARBITRARY VECTOR OF UNIT NORM

0.3779644966
0.3779644966
0.3779644966
0.3779644966
0.3779644966
0.3779644966
0.3779644966

PROJECTION MATRIX A

0.000 0.000 1.150 1.150 1.150 1.150 1.150
0.682 0.000 0.000 0.000 0.000 0.000 0.000
0.000 0.896 0.000 0.000 0.000 0.000 0.000
0.000 0.000 0.730 0.000 0.000 0.000 0.000
0.000 0.000 0.730 0.000 0.000 0.000 0.000
0.000 0.000 0.000 0.000 0.730 0.000 0.000
0.000 0.000 0.000 0.000 0.000 0.730 0.000

DOMINANT EIGENVALUE IS 1.185658612523

LIMITING VECTOR OF UNIT NORM

0.2175698429
0.3782458007
0.5005250573
0.4702009261
0.4209486246
0.3409535885
0.2110264301

CONVERGENCE DUE TO TOLERANCE
REFERENCES


(24) Ogden J. and Neal E. (1979) Applications of transition matrix models in forest dynamics, *Australian*


