Demography: Age and stage dependent population growth

Erol Akçay

February 6, 2019

The fundamental equation of population dynamics from the last lecture assumed that all individuals in a population are identical when it comes to contribution to population growth. This assumption is clearly violated in many ways. One of the most fundamental dimensions in which individuals differ is time. Old individuals do different things than young individuals, die and reproduce at different rates. The study of how these individual characteristics change over time in a population and how they affect the population's dynamics is called demography.

Lifetables

The life table is simply is a table giving the fraction of individuals of a single cohort (i.e., born at the same time) that survive to a given age. For example, Grant and Grant (1992) tracked the survival of two cohorts each of two species of Darwin's finches (*Geospiza fortis* and *G. scan dens*) on Daphne Major in the Galapagos, starting in 1975 and 1978 (see figure below). They banded individuals when they fledged and in their first year, and tracked how many were still alive in each subsequent year. The result is something like the table below, for one cohort of *G. fortis*.

TABLE 3. Life table of the 1975 cohort of *Geospiza fortis*. Age classes are designated in years. The starting number in the cohort (N_0) is given in parentheses because it was estimated, as explained in Table 1.

Age class	Number alive each year (N_x)	Proportion surviving at start of age x (l_x)	Proportion dying between x and $x + 1$ (d_x)	Mortality rate (q_x)	Mean expectation of further life (e_x)
0	(314)	1.000	0.370	0.370	2.02
1	198	0.630	0.347	0.550	1.92
2	89	0.283	0.159	0.562	2.66
3	39	0.124	0.035	0.282	4.42
4	28	0.089	0.003	0.036	4.96
5	27	0.086	0.010	0.112	4.13
6	24	0.076	0.003	0.042	3.58
7	23	0.073	0.010	0.130	2.28
8	20	0.063	0.038	0.600	1.55
9	8	0.025	0.012	0.502	2.13
10	4	0.013	0.000	0.000	2.75
11	4	0.013	0.007	0.538	1.75
12	2	0.006	0.000	0.000	2.00
13	2	0.006	0.003	0.500	1.00
14	1	0.003	0.003	1.000	0.50
15	0	0			***

The basic data in the life table is the fraction of a cohort that survives to age x, called survivorship, and denoted customarily by l_x (or l(x) when x is taken to be continuous). The rest of the quantities in the above table are all calculated from this. For instance, the fraction surviving from age x to age y is simply (which would be a point estimate of survival probability from x to y) how many individuals survived to y out of how many survived to x, so: $s_{x,y} = l_y/l_x$, or

in the case of surviving to the next age $s_x = l_{x+1}/l_x$. Likewise, the fraction (or probability of) dying between y and x is $s_{x,y} = (l_x - l_y)/l_x = 1 - q_{x,y}$ or in the case of dying before next age $q_x = 1 - s_x$.

The life table is complemented by the fertility (or fecundity) table, giving the reproductive output of individuals of different ages, usually denoted by m_x . In one of the cohorts that Grant and Grant tracked, it looks like the one on the margin. Note that this table gives the fertility of males and females of each age separately, reflecting the fact that in sexually reproducing species both sexes contributions need to be tracked. Most demographic models however, abstract away from this fact and track only one sex, almost always the females (hence m_x , for "maternity").

Armed with the survivorship and fertility schedules, one can calculate the dynamics of the population in question. Below are some of the techniques and important quantities one can calculate.

Population growth with age-structured demography

One of the most basic quantities one can calculate with the demographic information above is how many offspring each individual can expect to produce over its lifetime. In a discrete age-class representation, the total expected reproduction of an individual is given by the sum over the probability of surviving to each age, multiplied by how many offspring an individual produces at that age.

$$R_0 = \sum_{x=0}^{k} l_x m_x \,, \tag{1}$$

where k is the maximum age possible¹. In the continuous-time case, the sum is replaced by an integral:

$$R_0 = \int_0^k l(x)m(x)dx. \tag{2}$$

If $R_0 > 1$ each individual can expect to produce more than enough offspring to "replace" herself in the population before she dies. Thus, such a population will grow. Conversely, with $R_0 < 1$, the population will shrink. But R_0 is not equal to the growth rate r from the exponential population growth model we saw before. That's because the timing of reproduction matters as well: a population where every individual produces two offspring once in their first year of adulthood will grow after than a population where individuals produce the same number after a 5 year delay after reaching adulthood, even though their R_0 might be the same.

Another quantity of interest is the generation time, which is defined in multiple ways, but commonly as as the expected age of the parent of randomly selected offspring. A randomly selected offspring will have probability $\frac{l_x m_x}{\sum_{k=1}^{n} l_x m_y}$

TABLE 7. Survival (l_i) and fertility (m_i) table for 116 males and 67 females of the 1975 cohort of *Geospiza fortis*, based on assumptions given in Table 6. Age classes are designated

Age class (x)	Probability of surviving to x (/x)	Fledg- lings per season (m,)	Product of survival and re- production (l,m,)	Repro- ductive value (V_x)
Males				
0	1.000	0.0	0.0	0.746
1	0.630	0.096	0.060	1.192
2	0.336	0.0	0.0	2.189
1 2 3 4 5 6 7 8	0.235	0.458	0.107	3.126
4	0.204	0.240	0.049	2.783
5	0.204	0.575	0.011	2.543
6	0.188	0.708	0.133	2.760
7	0.180	0.760	0.014	2.142
8	0.157	1.900	0.298	2.369
9	0.063	0.406	0.026	1.168
10	0.031	0.0	0.0	1.550
11	0.016	0.0	0.0	3.000
12	0.016	3.000	0.048	3.000
13	0.008	0.0	0.0	0.0
14	0.008	0.0	0.0	0.0
15	0			
		$R_0 = \Sigma l_x m$	$_{\rm v} = 0.746 \ \Sigma$	=28.568
Females	i			
0	1.000	0.0	0.0	0.484
ĭ	0.630	0.644	0.406	0.769
;	0.364	0.0	0.0	0.214
2 3 4 5 6	0.072	1.083	0.078	1.083
4	0.024	0.0	0.0	0.0
5	0.024	0.0	0.0	0.0
6	0			
7				
8				
9				
10				
11				
12				
13				
14				
15				
		D 227	$= 0.484 \Sigma$	2 5 5 6

 $^{^{1}}$ k could be in principle infinity or in practice very large, if the organism can potentially survive indefinitely.

of belonging to a parent of age x, so the expected age of a parent of randomly selected offspring will be

$$T_G = \sum_{x=0}^{k} \frac{x l_x m_x}{\sum_{y=0}^{k} l_y m_y}$$
 (3)

Euler-Lotka equation

Consider the number of births in a population at time t (take time as continuous for the moment), denoted by B(t). We can express B(x) as follows:

$$B(t) = \int_0^k B(t-x)l(x)m(x)dx, \qquad (4)$$

where l(x) is the survivorship to age x, and m(x) (for maternity) is the fertility of an individual aged x. Here, k is the maximum age in the population. This equation basically states the fact that all individuals that give birth now must have been born some time back, and have survived up to now.

To find the solution to this equation, let's make the Ansatz that the population is growing exponentially, such that $B(t) = e^{rx}B(t-x)$ and substitute B(t-x) into the integral to get:

$$B(t) = \int_0^k e^{-rx} B(t) l(x) m(x) dx$$

$$\Rightarrow 1 = \int_0^k e^{-rx} l(x) m(x) dx , \qquad (5)$$

(6)

since B(t) doesn't depend on x. This the continuous form of Euler's equation.² The discrete version replaces the integral with sums, essentially:

$$1 = \sum_{x=0}^{k} e^{-rx} l_x m_x \,, \tag{7}$$

or remembering that the discrete rate of increase was $\lambda = e^r$:

$$1 = \sum_{x=0}^{k} \lambda^{-x} l_x m_x \,, \tag{8}$$

Stable age distribution

Now, let's consider the ratio of the number of individuals of two ages, x and y. The individuals of age x were born t - x years ago and have survived with probability l(x), so their numbers are B(t-x)l(x), whereas the number of

² The discrete version was first derived by Euler himself; Lotka generalized it to the continuous version here.

individuals of age y is B(t-y)l(y). If the population grows exponentially, the ratio of these is

$$\frac{\text{\# of individuals of age x}}{\text{\# of individuals of age y}} = \frac{B(t-x)l(x)}{B(t-y)l(y)} = e^{-r(x-y)}\frac{l(x)}{l(y)}, \tag{9}$$

which is independent of t. Therefore, we can conclude that the ratio of individuals in different age classes doesn't change if the population is growing exponentially. So, the population will exhibit a steady state distribution of individuals of different ages. In fact, a stronger statement is true: most demographic models with constant vital rates converge to a steady state age distribution from any initial stage. That stable age distribution is determined purely by the vital rates, as is the intrinsic rate of increase, r.

Reproductive value

Another concept that is of central importance in demography and life-history theory is that of reproductive value (RV), due to R.A. Fisher. The RV of an individual of certain age class is defined as how many offspring that individual can expect to contribute to the population for the remainder of its life, relative to the value of the individual's reproduction now. The reproductive value of age class x, v(x) is defined by:

$$v(x) = \frac{e^{rx}}{l(x)} \int_x^k e^{-ry} l(y) m(y) dy$$
 (10)

To derive this expression, note first the fact that in an exponentially growing population, the "share" of total population size that belongs to one offspring born at time x if going to be $\frac{1}{B_0e^{rx}}$, where B_0 is a constant denoting the population size at time 0. The same fraction belonging to an individual born at a later time y is likewise $\frac{1}{B_0e^{ry}}$. So, the ratio of how much of the population an individual can "claim" by having an offspring at times x vs. y is $e^{-r(y-x)}$. Given an individual is of age x, its probability of survival to age y is given by l(y)/l(x), and the number of offspring it will produce at age y is m(y), so the total number of offspring an individual can expect to produce at age y given it is now age x is: l(y)m(y)/l(x). Multiplying it with the discount rate we found above $(e^{-r(y-x)})$ and integrating from y=x to the largest age possible, we find expression (10). In other words, the reproductive value of an individual at age xis the total share of the population that the individual can expect to contribute going forward from x.

Likewise, for discrete age classes, we have

$$v_x = \frac{\lambda^x}{l_x} \sum_{y=x}^k \lambda^{-y} l_y m_y \tag{11}$$

Reproductive value allows us to compare the effects of changing the vital rates of different classes of individuals on the overall trajectory of the population. For example, suppose a change in the environment (such as change in predation or harvesting pressure) increases the mortality rate of individuals of a certain age class (as it routinely happens with hunting, fishing, and other forms of human exploitation of natural populations). This will have a larger effect on population dynamics if the age class in question is one with high reproductive value, relative to one with low, since the former was expected to contribute more to population growth. Precisely the same idea applies to the action of natural selection on different demographic classes: natural selection "cares" about different ages in accordance to their reproductive value.

Matrix projection models: Leslie matrices

The Euler-Lotka equation above describes the growth of the overall population size. But the information contained in the survivorship and fertility schedules allow us a more detailed description of the dynamics of an age-structured population. In particular, in the discrete case, we can describe the dynamics of all age classes with the following equations, where n_i is the number of individuals in the *i*th age class:

$$n_{1,t+1} = \sum_{x} m_{x,t} n_{x,t} \tag{12}$$

$$n_{x+1,t+1} = s_x n_{x,t}$$
 for all $x \ge 2$, (13)

where again s_x is the survival probability from x to x+1. These equations can be conveniently summarized using matrix notation:

$$\mathbf{n}_{t+1} = \mathbf{A} \cdot \mathbf{n}_t \,, \tag{14}$$

where \mathbf{n}_t is a vector³ of length k with entries giving the size of each age class, \cdot represents matrix multiplication⁴, and **A** is a $k \times k$ matrix given by:

$$\mathbf{A} = \begin{pmatrix} m_1 & m_2 & m_3 & \cdots & m_k \\ s_1 & 0 & 0 & \cdots & 0 \\ 0 & s_2 & 0 & \cdots & 0 \\ 0 & 0 & s_3 & \cdots & 0 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$
 (15)

In other words, the first row of **A** is given by the maternities of all age classes, where the subdiagonals of the matrix is populated by the survival probabilities from one age to the next. The rest of the entries are zero. The matrix A is called a Leslie matrix, after ecologist Patrick Leslie who first devised it in analyzing demographic models of animal populations. It is also a special case of a projection matrix, so named because it projects the current population state into the next time step.

³ It is a standard notational convention to represent vectors and matrices with bold letters; the former usually lowercase while the latter are uppercase. We denote the entries of a matrix **X** with $x_{i,j}$, where the first subscript gives the row, the second the column of the entry. ⁴ The matrix product of two matrices **X Y** of

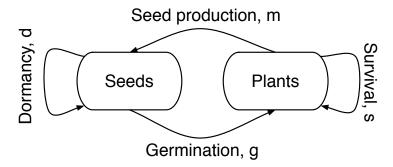
dimensions $p \times m$ and $m \times q$, respectively, is a $p \times q$ matrix with the i, jth entry given by the scalar product of the ith row of \mathbf{X} and jth column of **Y**, or $(\mathbf{X} \cdot \mathbf{Y})_{i,j} = \sum_{l} x_{i,l} y_{l,j}$

Matrix population models: stage structured populations

Age is not the only, or the best, characteristic to describe the differential contribution of individuals to population dynamics. For example, in many plants, contribution to population growth is a function of the size of the shoot system, regardless of how old it is (within certain limits). In those cases, describing the population with age structure is not desirable. But we can generalize the Leslie matrix method to deal with these types of models as well. Instead of considering classifying individuals by age, we can classify them by some other characteristic, say size. These models are called "stage structured population models" where stage can be size, height, different classes (e.g., helpers vs. breeders in cooperative breeding), etc. In this more general case, the entries of the matrix are no longer constrained to be non-zero only in the first row and the subdiagonal⁵.

One of the ways one can visualize a stage-structured model with a graph where the stages between individuals can transition in one time-step are connected with a line. The example below depicts a very simple life-history for a perennial plant that only has two stages: seeds and mature plants. Naturally enough, seeds can germinate and become plants, and plants produce new seeds, but the life history also includes a seed bank composed of seeds that lie dormant for more than a year, germinating at some later time, and plants that survive and continue to produce seeds in the next year. The symbols for each arrow gives the rates of those transition (i.e., g is the probability of germination, d the probability of seed survival in the dormant stage, s the probability of survival for adult plants, and m the number of seeds produced per adult plant). This

5 which reflected the facts that all newborns are age class 1 and no individual can stay the same age for more than one time step.



life history can be characterized by the following matrix:

$$\mathbf{A} = \begin{pmatrix} d & m \\ g & s \end{pmatrix} \tag{16}$$

Note that this model makes two very strong assumptions, namely that the ages of either adult plants or seeds do not matter for their vital rates. A 100-year old seed is just as likely to germinate as one from last year, and the same for the plants. Obviously, such assumptions might not hold, in which case, one might

add more stages to the life history accounting for the age of the seeds and/or the plants.

Matrix population models: asymptotic analysis

We introduced matrix population models as simply mathematical short-hands for the discrete-time demographic models we have seen in the previous lecture. But they are more than that, and provide powerful methods for analyzing population trajectories⁶.

A bit of linear algebra

We need a couple of fundamental concepts from linear algebra, in particular that of "eigenvalues" and "eigenvectors". An eigenvector⁷ of a square matrix (of dimensions $n \times n$) is a vector of length n that, when multiplied with that matrix, results in itself, multiplied with a constant, i.e.:

$$\mathbf{A} \cdot \mathbf{w} = \lambda \mathbf{w} \tag{17}$$

where λ is a constant and it is called the eigenvalue of **A** associated with **w**. Now, if the vector \mathbf{w} was the population state at time t and \mathbf{A} our projection matrix, that means the population stage at time t+1 is simply all the stages multiplies with a constant λ , so the proportion of individuals in each age- or stage-class does not change.

An $n \times n$ matrix will in general have n eigenvectors and values⁸. The eigenvalues are found from the characteristic equation (or polynomial):

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0, \qquad (18)$$

where **I** is the identity matrix with 1s on the diagonal and zeros everywhere else and det() is the determinant of the matrix⁹. The determinant of a 2×2 matrix is given by¹⁰:

$$\det(\mathbf{A}) = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$$
 (19)

For example, in the plant model above, the characteristic equation would read:

$$\det(\mathbf{A} - \lambda I) = 0 = \det\begin{pmatrix} d - \lambda & m \\ g & s - \lambda \end{pmatrix}$$
 (20)

$$= (d - \lambda)(s - \lambda) - mg \tag{21}$$

$$= \lambda^2 - (d+s)\lambda - mg \tag{22}$$

From which we get

$$\lambda = \frac{1}{2}(d+s \pm \sqrt{(d+s)^2 + 4gm})$$
 (23)

⁶ Caswell (2001) is the "bible" for these methods, of which we are going to cover only a tiny fraction. For anyone interested in demography and its ecological or evolutionary consequences, Caswell is mandatory (and enjoyable) reading.

^{7 &}quot;Eigen" means "self" in German.

⁸ Though in some cases, there might be degenerate ones; with some eigenvectors and values repeated two or more times. This rarely happens in demographic models, so we will not deal with that case.

⁹ This equation is a necessary condition for the equation (17) to hold.

 $^{^{10}\,\}mathrm{The}$ determinant of an $n\, imes\,n$ matrix is defined recursively. Define by A_{ij} the matrix one gets by deleting the ith row and jth column of matrix A. Then the determinant $\det(\mathbf{A}) = \sum_j (-1)^{i+j} a_{i,j} \det(A_{ij})$ for any i. The determinant of the matrix A_{ij} itself might need to be expanded with the same formula, all the way until one gets 2×2 matrix.

Since all the terms in the square root are positive, we know both eigenvalues are real and the one with the plus is greater than the one with the minus. The eigenvalue with the largest magnitude is usually called the dominant eigenvalue, and represented with λ_1 ; with the corresponding eigenvector \mathbf{w}_1 . The next highest one is λ_2 , \mathbf{w}_2 , and so on.

Finally, if the eigenvalues of an $n \times n$ matrix are all distinct (no repeated roots to the characteristic polynomial), the eigenvectors are also all distinct, and they are linearly independent, meaning that no eigenvector can be expressed as the sum of the others, each multiplied with some constant coefficient. One can think of such eigenvectors as forming a coordinate basis (like the x and y axes on a Cartesian plot, in which all vectors in n-dimensions can be expressed.

Long-term behavior of a population

We have seen that age-structured models have a characteristic stable age distribution. Matrix population models extend this stage-structured models as well. In particular, any initial population state \mathbf{n}_0 can be expressed as

$$\mathbf{n}_0 = \sum_i c_i \mathbf{w}_i , \qquad (24)$$

where c_i are constants. Applying the projection matrix to \mathbf{n}_0 t times, we get the population state at time t:

$$\mathbf{n}_t = \mathbf{A}^t \mathbf{n}_0 = \mathbf{A}^t \sum_i c_i \mathbf{w}_i = \sum_i c_i \mathbf{A}^t \mathbf{w}_i = \sum_i c_{i,t} \lambda_i^t \mathbf{w}_i , \qquad (25)$$

since we can exchange the order of matrix multiplication with multiplication with a constant, and also distribute it over a sum. Since the multiplication of A with **w** only yields λ **w**, repeating this t times will yield λ ^t**w**. Now, let's factor out the dominant eigenvalue (assume for the moment it is distinct and that all eigenvalues are real):

$$\mathbf{n}_t = \lambda_1^t \sum_i c_{i,t} \left(\frac{\lambda_i}{\lambda_1} \right)^t \mathbf{w}_i , \qquad (26)$$

Since the magnitude of λ_1 is greater than all the others, as $t \to \inf$, the tth power of the ratio of eigenvalues will vanish, i.e. $(\lambda_i/\lambda_1)^t \to 0$ for all i greater than 1. That means that in the long term (large t), the population will be approximated by $\lambda_1^t c_1 \mathbf{w}_1$:

$$\mathbf{n}_t \approx \lambda_1^t c_1 \mathbf{w}_1 \tag{27}$$

Which means that in the long term, the age (or stage) structure of the population is given by the eigenvector corresponding to the dominant eigenvalue, and the growth rate of the population is the dominant eigenvalue itself¹¹. These are exactly the kinds of results we had from the previous lecture using the Euler-Lotka

¹¹ Thus revealing that it was no coincidence that we used the symbol λ both for the population growth rate in the discrete population model, and for eigenvalues in matrix models.

equation for age-structured populations, but generalized to a much greater class of demographic models. Moreover, once we construct the projection matrix from the life table or other demographic measurements, we have the powerful tools of linear algebra, both mathematical and computational, at our disposal to calculate important quantities.

As one example, the existence and uniqueness of the positive dominant eigenvalue and a strictly positive eigenvector associated with it is guaranteed by the Perron-Frobenius theorem for the class of matrices that are called "primitive" 12. A matrix **A** is primitive if all the elements of \mathbf{A}^k are strictly positive for some large enough k. Intuitively, this means that in a population described by a primitive matrix, any individual you pick at a given time point, regardless of its age (or stage) class has a positive expected contribution to every age or stage class at some point far enough in the future. Most demographies result in primitive projection matrices. One notable exception is a age-structured demography with a single reproductive age class, in which case, one can have cycles where an individual of age x will only contribute to age class x + t in any t time-step (and to x + t - h + 1 if t > h), where h is the sole age-class that reproduces, and nothing to others.

Reproductive value in matrix models

The stable age or stage distribution and asymptotic growth rate are not the only quantities the matrix population model just spits out for us. We can also get the reproductive value for each class easily. In particular, the reproductive value is given by the **left** eigenvector of A^{13} associated with the dominant eigenvalue, \mathbf{v}_1 that satisfies ¹⁴:

$$\mathbf{v}_1^{\mathsf{T}} \mathbf{A} = \lambda_1 \mathbf{v}^{\mathsf{T}} \,, \tag{28}$$

So, the dominant eigenvalue of A gives us the asymptotic growth rate of a population, the right eigenvector associated with it the stable age distribution, and the left eigenvector the reproductive values of every state. Pretty neat.

Sensitivity of λ_1 *to vital rates*

One of the major uses of matrix models is to study the evolution of life-history, as they allow us to quantify how population growth λ_1 responds to changes in the vital rates, i.e. changes in elements of the Leslie matrix. This is called "sensitivity" of λ_1 to vital rates. In general, vital rates will evolve in the direction of increasing λ_1 . The sensitivity of λ_1 is also useful for trying to figure out what demographic interventions are most effective for either conserving a population (increasing λ_1) or controlling it (decreasing λ_1).

Technically, the sensitivity of λ_1 to the change of a matrix element a_{ij} is defined simply as the partial derivative of λ_1 w.r.t. to a_{ij} , which can be shown

¹² See Caswell (2001), ch 4, section 5.

¹³ what we called the eigenvector for now, w, is more properly known as right eigenvector, since it is multiplied to the right of the matrix 14 The symbol $^{\top}$ denotes the transpose of a matrix or vector. i.e., the flipping of rows and columns, such that a row vector becomes a column one, and vice versa.

to be:

$$\frac{\partial \lambda_1}{\partial a_{ij}} = \frac{v_i w_j}{\mathbf{v} \cdot \mathbf{w}} \,, \tag{29}$$

where v_i and w_j are the ith and jth elements of the reproductive value and stable age distribution vectors, respectively, and \cdot is the scalar product. An intuitive interpretation of (29) is that the sensitivity of the long-term growth rate of the population to increasing the transition rate a_{ij} is proportional to the fraction of population in stage j (the source stage) times the reproductive value of stage i(the target stage), the latter giving the long-term contribution of stage i to the population. If $\frac{\partial \lambda_1}{\partial a_{ij}} > 0$, we'd expect a_{ij} to increase.

Now, since both reproductive values and stable stage distribution is nonnegative, that means the right-hand side of (29) will also be non-negative, and therefore you might think that all elements of a projection matrix will be selected to increase. That would be true if there were no trade-offs, but in life, everything has trade-offs. For instance, investing more in reproduction at a given age frequently decreases survival to the next age. In the Leslie matrix, that can be expressed by making s_i (the probability from surviving from age i to i + 1) a decreasing function of m_i (the number of offspring produced at age i). Then, the change in s_i will depend on the sensitivities of λ_1 to both s_i and m_i .

References

Caswell, H. 2001. Matrix Population Models. Sinauer Associates, Sunderland MA.

Grant, P. R., and B. R. Grant. 1992. Demography and the genetically effective sizes of two populations of darwin's finches. Ecology pages 766–784.