Chapter 2

Matrix Models for Structured Populations

If a population or species is in decline and at risk of extinction, it's clear what we need to do: increase the birth rate or decrease the death rate – or both. But exactly *whose* birth and death rates are we talking about? Given limited manpower and financial resources, would it be more effective (for example) to create additional good nesting sites so more adult birds can breed each year, to shield eggs from predators so that each nest produces more offspring, or to augment food supply so that juveniles have a better shot at surviving to breeding age?

Trying to answer these questions in advance – rather than by trial and error – is an increasing part of the conservation biologist's job. This chapter focuses on a type of model that is widely used for this task: matrix models for the dynamics of structured populations. "Structured" means that the model incorporates differences among individuals. Structured population models were developed centuries ago for the study of human populations – one basic result is credited to Euler (1707-1783) – and we will start by considering that case, where individuals are distinguished based on their age. We then consider models with other kinds of population structure. Their study requires some aspects of matrix theory, which we review from the ground up, beginning by defining matrices and reviewing basic matrix algebra. We then consider some applications of matrix models, including conservation planning, and some generalizations of the model to include factors other than individual variability.

2.1 The population balance law

The starting point for modeling population dynamics is the fundamental balance law

$$N(t+1) = N(t) + \text{Births} + \text{Immigration} - \text{Deaths} - \text{Emigration} \tag{2.1}$$

where $N(t)$ is the number of individuals or density (number per unit area) in the population at time t. The balance law becomes a complete model when we specify formulas for the quantities on the right-hand side. The simplest model is to assume a closed population without

Immigration or Emigration, and that the per-capita (i.e. per-individual) birth and death rates are constant:

$$\text{Births} = b \times N(t), \quad \text{Deaths} = d \times N(t)$$

The balance law then becomes

$$N(t+1) = N(t) + bN(t) - dN(t) = \lambda N(t) \tag{2.2}$$

where $\lambda = 1 + b - d$. This model is simple enough that we can solve it. Starting from any initial population size $N(0)$ we get

$$\begin{aligned} N(1) &= \lambda N(0) \\ N(2) &= \lambda N(1) = \lambda(\lambda N(0)) = \lambda^2 N(0) \\ N(3) &= \lambda N(2) = \lambda(\lambda^2 N(0)) = \lambda^3 N(0) \end{aligned} \tag{2.3}$$

and so on, leading to the general solution

$$N(t) = \lambda^t N(0). \tag{2.4}$$

This is *exponential population growth*: defining $r = \log \lambda$, we then have $\lambda^t = (e^r)^t = e^{rt}$ and so

$$N(t) = e^{rt} N(0).$$

(Note that "log" means the natural (base-e) logarithm; we will use $\log_{10}$ to indicate base-10 logarithms).

Growth can't go on forever, so (2.4) can't be valid forever if $\lambda > 1$. This kind of limitation bothers biologists much more than it bothers physical scientists, who are used to the idea that different models for a given system may be valid in different circumstances. Anderson and May (1992, p. 9) compare simple biological models to Newton's first law of motion:

> A body remains in its state of rest or uniform motion in a straight line, unless acted on by external forces.

Exponential population growth has the same character – it tells us what happens if current conditions persist without change:

> A closed population of self-reproducing entities – such as viruses, cells, animals, or plants – will grow or decay exponentially at a constant rate, unless a change in conditions alters the per-entity birth or death rates.

Therefore, more general models are derived by considering the factors that can alter the average per-entity birth and death rates.

2.2 Age-structured models

The theme of this chapter is that per-entity birth and death rates are affected systematically by differences among individuals. Some such differences are happenstance – who happens to

be at the spot where the tornado touches down – but others are predictable consequences of attributes that vary among individuals, such as their age. To take the effect of age into account we need to describe the population by a state *vector* listing the numbers of individuals of each age. It is natural (but not necessary) to use years as the time unit; the state variables are then $n_a(t)$, the number of a-year-old individuals in year t, with a running from 0 to the maximum possible age A.

For now, we continue to assume a closed population without immigration or emigration. The model's dynamic equations are bookkeeping expressed in mathematical symbols, as in the salmon model (Chapter 1). Consider one of the authors: 50 years old in January 2004. In order to reach that state, he must have been 49 years old in January 2003 and survived the next year. Consequently

$$n_{50}(2004) = p_{49} n_{49}(2003) \tag{2.5}$$

where p_{49} is the probability that an 49-year-old survives to age 50. In general, this line of reasoning tells us that

$$n_a(t+1) = p_{a-1} n_{a-1}(t) \text{ for } a > 0. \tag{2.6}$$

where p_a is the probability that an a-year-old individual survives to be age $a+1$.

To complete the model we need to specify the number of births each year. If we assume that per-individual birth rates are only a function of age, then we have

$$\begin{aligned} n_0(t+1) &= f_0 n_0(t) + f_1 n_1(t) + \cdots f_A n_A(t) \\ &= \sum_{a=0}^{A} f_a n_a(t) \end{aligned} \tag{2.7}$$

where f_a is the number of newborns next year, per a year old individual this year. Two conventions can be used: count everybody, or only count females – in which case f_a **only** includes female offspring. The females-only convention is by far the more common, and we will always use it here.

Exercise 2.1 Find at least two important assumptions that are necessary for (2.6) to be true.

Exercise 2.2 Define $l_0 = 1$, $l_a = p_0 p_1 p_2 \cdots p_{a-1}$, the probability of an individual surviving from birth to age a. Explain in words why

$$n_0(t+1) = \sum_{a=0}^{A} f_a l_a n_0(t-a) \tag{2.8}$$

Exercise 2.3 The general theory developed later in this chapter tells us that in the long run a population governed by the age-structured model typically grows exponentially, as in (2.4). In particular, $n_0(t) = c\lambda^t$ will hold (with greater and greater accuracy as time goes on), for some λ and c. By substituting this approximation for $x_0(t)$ into (2.8), show that the long-term population growth rate λ satisfies the equation

$$\sum_{a=0}^{A} \lambda^{-(a+1)} l_a f_a = 1. \tag{2.9}$$

This is called the Euler, Lotka, or Euler-Lotka equation.

Exercise 2.4 (a) Show that the left-hand side of (2.9) is a decreasing function of λ by computing its derivative with respect to λ. (b) Compute the values of the left-hand side in the limits $\lambda \to 0$ and $\lambda \to \infty$. (c) Explain why (a) and (b) imply that (2.9) has one and only one positive real solution.

2.2.1 The Leslie matrix

It is convenient and informative to express the age-structured model in matrix notation. In this form it is called the Leslie matrix model, after British ecologist P.H. Leslie who popularized age-structured models for animal populations in the mid 20th Century.

First we need to review a bit about matrices. A matrix is a rectangular array of numbers. A matrix $\mathbf{A}$ with entries a_{ij} is said to have size $m \times n$ if it has m (horizontal) rows and n (vertical) columns. Thus the row index i takes the values $1, 2, \ldots, m$ (1 indicating the top row) and the column index j takes the values $1, 2, \ldots, n$ (1 indicating the leftmost column). A matrix with one column is called a column vector, and a matrix with a single row is called a row vector.

Matrix algebra was invented for studying systems of linear equations in several unknowns, such as

$$\begin{aligned} 3x_1 + 5x_2 &= 1 \\ x_1 - 2x_2 &= 0 \end{aligned} \tag{2.10}$$

Solving one such equation in one unknown is a snap

$$\begin{aligned} 3x &= 1 \\ (1/3) \times (3x) &= (1/3) \times (1) \\ x &= 1/3 \end{aligned} \tag{2.11}$$

Matrix algebra let us make (2.10) look like (2.11), so that we can solve it in the same way. We put the coefficients in (2.10) into a matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 5 \\ 1 & -2 \end{bmatrix} \tag{2.12}$$

and put the variables and right-hand side into vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Matrix-vector multiplication is then defined so that (2.10) is equivalent to the single matrix equation:

$$\begin{array}{ccccc} \mathbf{A} & \mathbf{x} & = & \mathbf{b} \\ \begin{bmatrix} 3 & 5 \\ 1 & -2 \end{bmatrix} & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & = & \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{array} \tag{2.13}$$

The definition that makes (2.10) and (2.13) mean the same thing is the following: if $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is a column vector and $\mathbf{A}$ is a matrix with n columns, then $\mathbf{Ax}$ is the column vector

$$\begin{aligned} &(x_1 \times \texttt{first column of } \mathbf{A}) \\ &+(x_2 \times \texttt{second column of } \mathbf{A}) \\ &\qquad \vdots \\ &+(x_n \times \texttt{last column of } \mathbf{A}). \end{aligned} \tag{2.14}$$

Algebraically, that works out to the following formula: for a matrix $\mathbf{A}$ with n columns and a column vector $\mathbf{x}$ of length n, $\mathbf{Ax}$ is the vector whose i^{th} element is

$$(\mathbf{Ax})_i = \sum_{j=1}^{n} A_{ij} x_j \tag{2.15}$$

where A_{ij} is the number in the ith row and j^{th} column of the matrix $\mathbf{A}$. The jth term in this sum corresponds to x_j multiplying the ith element down in the jth column of $\mathbf{A}$; summing over j then gives you the total of all contributions to $(\mathbf{Ax})_i$ from the n columns of $\mathbf{A}$.

What good does that do us? Suppose that we could find a multiplicative inverse to $\mathbf{A}$ – a matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}(\mathbf{Ax}) = \mathbf{x}$ for any vector $\mathbf{x}$. Then we could solve (2.13) the same way we solved (2.11): just multiply both sides of equation (2.13) by the inverse of $\mathbf{A}$ to get the solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Figuring out when such inverses exist and how to compute them was one of the major accomplishments of 19th century mathematics.

As is often the case in mathematics, a tool invented for one purpose turns out to be useful for many others. Notice that the equation (2.7) for births in the age-structured model has the same form as (2.15). The survival equation (2.6) is also a sum, with only one term. So we can express these in matrix notation by putting the survival and birth rates in the right places:

$$\begin{bmatrix} n_0(t+1) \\ n_1(t+1) \\ \vdots \\ \\ n_A(t+1) \end{bmatrix} = \begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_A \\ p_0 & 0 & 0 & \cdots & 0 \\ 0 & p_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & p_{A-1} & 0 \end{bmatrix} \begin{bmatrix} n_0(t) \\ n_1(t) \\ \vdots \\ \\ n_A(t) \end{bmatrix} \tag{2.16}$$

or simply

$$\mathbf{n}(t+1) = \mathbf{Ln}(t) \tag{2.17}$$

where $\mathbf{L}$ is the matrix in (2.16), and $\mathbf{n}(t)$ is the population vector $(n_0(t), n_1(t), \cdots, n_A(t))$. The top row of the matrix contains the births, and the other nonzero entries are survival.

For example, consider a plant with a maximum age of 2 – this might be a monocarpic plant that flowers once, at either age 1 or age 2, and then dies. Suppose that newborn offspring (age 0) have a 50% chance of surviving to age 1; age-1 plants produce f_1 offspring each on average, and have a 25% chance of surviving to age 2, and age-2 individuals have f_2 offspring each on average. The Leslie matrix is then

$$\mathbf{L} = \begin{bmatrix} 0 & f_1 & f_2 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \tag{2.18}$$

Definition (2.14) tells us how to "read" a matrix like (2.18). Since $n_i(t)$ multiplies the ith column of the Leslie matrix, the ith column of $\mathbf{L}$ gives the individuals of each age "next year" resulting from a single age-i individual "this year", as a consequence of their survival and fecundity. So the first column says that for each age-0 individual this year, there will be (on average) half a 1-year old next year – the ones that survive. The second says that for each age-1 individual this year, the population next year will have f_1 age-0 individuals (the offspring of age-1 individuals)

and 0.25 age-2 individuals (the age-1 individuals who survive to next year). Age-2 individuals all die, so their only contribution to next year's population is their offspring (f_2 per 2-year-old). In the same way, the jth column of the general Leslie matrix (2.16) says that for each j-year-old "this year", next year's population will have f_j offspring (age 0) and p_j survivors (age $j+1$).

Note that it doesn't matter how survivorship and breeding at a given age are related to each other. For example, it could be the case in (2.18) that half the age-1 individuals in reproduce and then die (having $2f_1$ offspring each, on average) while those that don't reproduce have a 50% chance of surviving to age 2. Or it could be the case that all age-1 individuals reproduce, and all have a 25% chance of surviving to age 2. Either way the matrix is the same.

Exercise 2.5 Verify that the following are correct in two different ways: using (2.14) and again using (2.15).

(a) $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0.5F & F \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 10 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 20F \\ 5 \\ 5 \end{bmatrix}$

Exercise 2.6 Write a script file to run simulations of the biennial plant model (2.18) with $f_1 = 1, f_2 = 5$, starting from a single age-1 individual at time 0. Have the script plot as functions of time (1) the log of the total population size $N(t) = n_0(t) + n_1(t) + n_2(t)$, and (2) the fraction of individuals of each age, $w_i(t) = n_i(t)/N(t), \quad i = 1, 2, 3$, for t=1 to 50. What long-term properties of the population do you see in your simulation results?

Exercise 2.7 Write down the Euler-Lotka equation (2.9) for the biennial plant model (2.18) with $f_1 = 1, f_2 = 5$, and numerically solve it for the value of λ. How does this compare to the rate of population growth that you saw in your simulations? [You can find λ approximately by having your script compute the value of the sum in the Euler-Lotka equation at a finely-spaced set of λ values, and finding one at which the sum is closest to 1. Or if you're adventurous, find a function in your scripting language that finds the roots of univariate functions, e.g. `uniroot` in **R** or `fzero` in Matlab.]

2.2.2 Warning: Pre-breeding versus post-breeding models

The interpretation of the f's in the age-structured model depends on how births relate to the annual census time. Suppose we census the population on January 1 each year. In humans, births occur year-round, so f_a should be the average number of births over the coming calendar year, to an individual whose age was between a and $a+1$ on January 1, but only counting offspring that survive until January 1 of next year. So f_a is the sum over all such individuals of

(average number of births on Jan 1) × (survival from this Jan 1 to next Jan 1)
+ (average number of births on Jan 2) × (survival from this Jan 2 to next Jan 1)

and so on.

In other cases it is more accurate to assume a once-per-year seasonal "pulse" of births, as if all offspring for the year were born in a single instant. Let m_a be the number of offspring that

an a year old has in the current birth pulse. If we census the population immediately after the pulse (*post-breeding* census) then

$$\begin{aligned} f_a = \quad & p_a \text{ [survival to next year]} \\ & \times m_{a+1} \text{ [\# offspring in next year's birth pulse].} \end{aligned}$$

But if we census just before the pulse (*pre-breeding* census), then

$$\begin{aligned} f_a = \quad & m_a \text{ [\# offspring now - but not counted until next year]} \\ & \times p_0 \text{ [fraction of offspring who survive to be counted].} \end{aligned}$$

Both of these are valid under their assumptions about census timing, and both are used. As a result, formulas for things like life expectancy, population growth rate, etc. exist in two different versions for pre-breeding and post-breeding models. An additional complication is that some authors (e.g. Caswell 2001) number age-classes starting at 1 rather than 0, so that their $n_1(t)$ is equivalent to our $n_0(t)$ – individuals who have not yet reached their first birthday. Even experts get confused by all these options, and many books and papers include a mix of formulas based on incompatible assumptions. So when you see f_4 in a book or paper, it's important to check what the author intends it to represent.

2.3 Stage-classified matrix models

In most applications to non-human organisms, the oldest age A really consists of individuals aged A *or older* due to lack of data. After all, the meaning of "extreme old age" is that most individuals die before they get there, so there are relatively few observations of what happens to extremely old individuals. For example, suppose during your study of a lizard population, there is one hardy 4-year old who lives to be 5, then lives to be 6, and then dies, while all other individuals die before they reach the age of 4. So would you then take $p_4 = 1, p_5 = 1, p_6 = 0$? This sort of thing has been done – and published in reputable journals – but a better option is to assume that all individuals above some age are identical so that you can get a reasonable estimate of their average survival probability.

Collapsing all ages above some cutoff into one "age-class" is our first example of the tradeoff between *model error* and *parameter error*. Model error means errors due to incorrect assumptions, where the model simplifies or omits known aspects of reality. Combining all individuals of age 3 or above is likely to create model error, because we have no grounds for believing that there really are no systematic differences between a 3-year-old and an 5-year-old. Parameter error means errors due to parameters being estimated inexactly from a limited set of data. By combining all 3+ year olds, we avoid the parameter errors that would result from estimating p_4, p_5 and p_6 from a sample of size 1 in which the only possible values are 0 or 1. The resulting model has a category of individuals who are likely to be fairly similar, with parameters for the category being "average" or "typical" values for members of the category. It is always possible to reduce model error by making a model more complex, but parameter error usually goes up because you have to somehow estimate more parameters from the same amount of data. We discuss this tradeoff more fully in Chapter 9.

Category	Carapace length (mm)	Category	Plant area (cm^2)
Yearling		Seedling	0.5 - 5
Juvenile 1	<60	Juvenile	5.1 - 35
Juvenile 2	60-99	Adult 1*	35.1-200
Immature 1	100-139	Adult 2*	200.1-600
Immature 2	140-179	Adult 3*	> 600
Subadult*	180-207		
Adult 1*	208-239		
Adult 2*	>240		

Table 2.1: Two examples of stage classifications based in part on individual size. Asterisks indicate reproductive categories. The two left columns give the categories used by Doak et al. (1994) for modeling population viability of desert tortoise in the western Mojave desert, which were the same as those used by the Bureau of Land Management in the population monitoring program that provided the data for the model. The two right columns give the categories defined by Valverde and Silvertown for the forest herb *Primula vulgaris*.

2.3.1 General stage classifications

More generally, individuals can be classified by their **stage** in the life cycle. Sometimes there really are discrete life stages, such as caterpillar-cocoon-butterfly (or more generally larva-pupa-adult in insects). However, sometimes it's just a group of individuals defined by some measurable feature such as length or weight, that is the best available attribute for predicting their fate over the next period of time. As in our example above, we strive to define groups of individuals similar enough that it is safe to describe them by a single representative value for each parameter.

The most commonly used attribute for defining categories is some measure of individual size. These have long been popular in the forestry literature, because size is generally much better than age for predicting tree growth and mortality. Now size is used also for animal populations, with recent examples including sea turtles, desert tortoise, geese, corals, copepods, and fish (Caswell 2001). Typically, size categories are defined so that between one census and the next individuals can grow or shrink by at most one category, and all newborns are in the smallest size category, but this is not always the case. For example, Valverde and Silvertown (1998) used size-classified matrix models for the woodland herbaceous plant *Primula vulgaris* in which individuals could grow by 2 categories, in order to study how *Primula* population growth was affected by the degree of forest canopy closure. For one of their study sites (Woburn Wood), the projection matrix for 1993 to 1994 was

$$\begin{bmatrix} 0 & 0 & .03 & .10 & .18 \\ .25 & .35 & .12 & .02 & 0 \\ .04 & .45 & .65 & .33 & .19 \\ 0 & 0 & .16 & .58 & .38 \\ 0 & 0 & 0 & .05 & .38 \end{bmatrix}$$

with the categories being defined by plant area (see the right-hand side of Table 2.1). As with our hypothetical model (**??**) we "read" this matrix by recognizing that each column specifies the contribution of one category to next year's population.

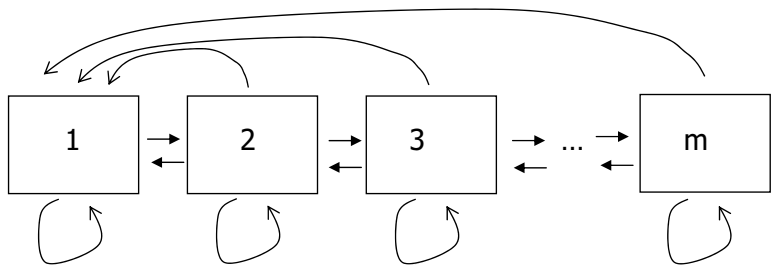

Figure 2.1: The standard size-class model. Size categories are broad enough that individuals can't change by more than one category between population censuses, and all newborn individuals are in the smallest size class. These all look the same apart from the number of "stages".

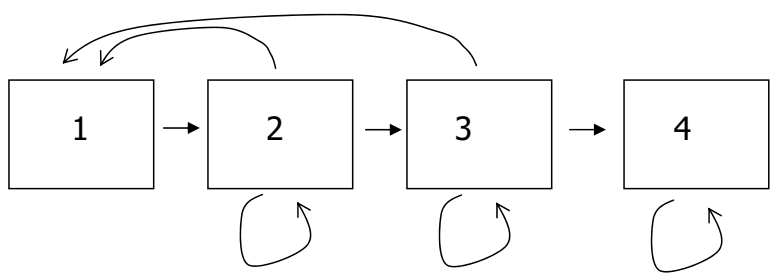

Figure 2.2: Stage-structured model for killer whales (Brault and Caswell 1993). The stages recognized were 1=Yearling, 2=Juvenile, 3=Mature female, and 4=Post-reproductive female.

A useful way to graphically represent a stage-classified matrix model is the *life cycle graph* in which each "node" represents a stage, and arrows show possible changes in stage for individuals between one time interval and the next. Figures (2.1) and (2.2) show two examples. By convention, staying put in a stage is drawn as an arrow, but deaths are omitted on the assumption that no life stage is invulnerable. Also note that there is no distinction in the diagram between survival and fecundity. Since a basic premise of the model is that an individual's stage-classification provides complete information about their future prospects, it doesn't matter (in the model) if a small individual is a newborn or an older individual who shrank back down to newborn size. The life cycle diagram represents your idea of a good way of classifying individuals. If there are discrete stages, it's probably a good idea to use those in the model. Otherwise, experience suggests that the most important issue is selecting which trait to use as the basis for classifying individuals (e.g., age versus size). The trait used for classifying individuals is sometimes called (somewhat confusingly) the *individual state* variable or *i-state* variable.

Given the right data, alternative choices of i-state variable can be compared objectively. For example: if you know an individual's size, can you predict her fecundity more accurately if you also know her age? Caswell (2001, section 3.3) presents several examples of this kind of

comparison. However, the classification is often dictated by circumstances. For example Doak et al. (1994) based their model on data that had already been collected by the Bureau of Land Management. They had no choice but to use size as their i-state variable, with the categories used in the BLM surveys (Table 2.1). Valverde and Silvertown (1998) based their classification on knowledge of the species natural history, with class boundaries chosen so that each category had sufficient sample size for estimating matrix entries.

Having chosen a stage classification, the model is completed by specifying the projection matrix entries a_{ij},

$$a_{ij} = \text{number of type-}i\text{ individuals at time } t+1\text{, per type-}j\text{ individual at time } t.$$

Contributions from j to i may be survival, fecundity or a combination of these. As in the age-structured model, we assume (for now) that the a_{ij} are constant. The fundamental balance law is then

$$\begin{aligned} n_i(t+1) &= a_{i1}n_1(t) + a_{i2}n_2(t) + \cdots \\ &= \sum_{j=1}^{m} a_{ij}n_j(t) \end{aligned} \tag{2.19}$$

The *projection matrix* $\mathbf{A}$ is defined to be the $n \times n$ matrix with entries a_{ij}, and the model then becomes

$$\mathbf{n}(t+1) = \mathbf{A}\mathbf{n}(t) \tag{2.20}$$

Estimating the value of matrix entries is a subject in itself. Entire careers (and entire books, e.g., Williams et al. 2002) are devoted to methods for analyzing census data on populations in order to estimate demographic rates. Morris and Doak (2002, Chapter 6) give some guidelines on how to conduct field studies to estimate demographic rates. The ideal situation is if individuals can be given a unique tag or mark (or come with unique markings), and you come back later to see what happened to them. Plants sit still and wait to be counted, but with animals it is often hard to distinguish between death and emigration out of the study area.

Similarly for fecundities, the best situation is if you can link newborn individuals to their parents, obtaining directly the fecundity of each individual. This is often possible with large animals, or animals that live in family groups. For plants, a common approach is to count seeds while they are still on the parent plant. Then assuming that (once released) a seed is a seed is a seed, you can estimate

$$\begin{aligned} F_i = &\ (\text{average number of seeds produced by a class-}i\text{ plant}) \\ &\times (\text{fraction of all seeds that survive to be seedlings at the next census}) \end{aligned}$$

The same idea has also been used for estimating fecundity in birds: do a post-breeding count of offspring while they are still in the nest, and decrement those numbers by a single estimate of the fraction that survive until the next time the population is censused.

There are also ways to indirectly estimate parameters from population count data. Indirect methods begin by assuming that the model is valid, and then asking: what must the survival (or growth or fecundity) parameters be, in order to generate the population that I observed? This is more difficult than direct estimates and less secure, because of the *a priori* assumption that the model is valid. Indirect methods for structured population models are reviewed by Wood (1997) and Caswell (2001, section 6.2).

Exercise 2.8 State in words the meaning of the second and fourth columns of the projection matrix given above for *Primula vulgaris* in Woburn Wood.

Exercise 2.9 Draw the life-cycle graph for *Primula vulgaris* in Woburn Wood.

2.4 Matrices and matrix operations

Our goal now is to derive general properties of matrix models that allow us to make connections between the matrix entries and the long-term fate of a population governed by (2.20). For example, in conservation planning it is important to know which matrix entries have the greatest impact on whether the population is growing or shrinking, so that those can be targets for remediation efforts (e.g., striving to increase the survival during particularly important stages of the life cycle).

Rewriting the balance equations (2.19) in matrix form (2.20) is more than a convenience, because the algebra of matrices (called *linear algebra*) has a lot to tell us about the equations. This section reviews some concepts and results from linear algebra that will give us much insight into the balance equations for populations. Moreover, the *same* results will be employed in Chapter XX in an entirely different setting, to model the gating of membrane channels in neurons. The universality of mathematics is astonishing!

2.4.1 Review of matrix operations

Addition and subtraction of matrices are done element-by-element, and are therefore only defined for matrices of the same size:

$$\mathbf{C} = \mathbf{A} \pm \mathbf{B} \text{ has entries } c_{ij} = a_{ij} \pm b_{ij}. \tag{2.21}$$

Multiplication of a matrix by a scalar (real number) is also element-by-element:

$$\mathbf{B} = c\mathbf{A} \text{ has entries } b_{ij} = ca_{ij}. \tag{2.22}$$

Examples:

$$\begin{gathered}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 6 & 10 \\ 4 & 8 & 12 \end{bmatrix} = \begin{bmatrix} 1+2 & 2+6 & 3+10 \\ 4+4 & 5+8 & 6+12 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 13 \\ 8 & 13 & 18 \end{bmatrix} \\ 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}\end{gathered} \tag{2.23}$$

Matrix multiplication is more complicated. The product $\mathbf{C} = \mathbf{AB}$ is defined if the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$. If $\mathbf{A} = a_{ij}$ has size $m \times n$ and $\mathbf{B} = b_{ij}$ has size $n \times r$, then $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ has size $m \times r$ and

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \tag{2.24}$$

Note that if $\mathbf{B}$ has only one column, this reduces to the definition of matrix-vector multiplication (2.15). Thus, another definition of matrix multiplication is the following:

$$j^{th} \text{ column of } \mathbf{AB} = \mathbf{A} \times (j^{th} \text{ column of } \mathbf{B}) \tag{2.25}$$

Our attitude is that matrix multiplication is usually best done on the computer. It is important to understand the conceptual definition (2.25) and the algebraic formula (2.24), but when working with actual numbers, it is easier and less error-prone to use a computer language which includes matrices and matrix multiplication.

An important special case of matrix multiplication is when $\mathbf{A}$ is $1 \times n$, i.e. a row vector $\mathbf{v}$, and $\mathbf{B}$ is $n \times 1$, i.e. a column vector $\mathbf{x}$. The produce $\mathbf{AB}$ then reduces to the *inner product* of two vectors,

$$\mathbf{v} \cdot \mathbf{x} = v_1 x_1 + v_2 x_2 + \cdots + v_n x_n \tag{2.26}$$

This expression is sometimes called the *dot product*, and the alternate notations $\langle \mathbf{v}, \mathbf{x} \rangle$ or $(\mathbf{v}, \mathbf{x})$ are also used.

Matrix operations share many properties with the familiar arithmetic of real numbers. For example,

- Matrix addition is associative $[\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}]$ and commutative $[\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}]$.
- Matrix multiplication is associative $[\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}]$ and distributive over addition $[\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}, (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}]$.

However, *matrix multiplication is not commutative*: typically $\mathbf{AB} \neq \mathbf{BA}$. Indeed, unless $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same size, either one of the products $\mathbf{AB}$ and $\mathbf{BA}$ will be undefined, or the two products will be matrices of different sizes. But even in the case of square matrices commutativity typically does not hold. Here is a simple example:

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

but

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

However, scalar multiplication is commutative in the sense that $\mathbf{A}(c\mathbf{B}) = c(\mathbf{AB})$.

2.4.2 Solution of the matrix model

Having defined matrix multiplication, we can now easily write down the solution to the matrix model (2.20), in the same way that we solved the unstructured model (2.2). Starting from some initial population vector $\mathbf{n}(0)$ we get

$$\begin{aligned} &\mathbf{n}(1) = \mathbf{An}(0) \\ &\mathbf{n}(2) = \mathbf{An}(1) = \mathbf{A}(\mathbf{An}(0)) = \mathbf{A}^2\mathbf{n}(0) \\ &\mathbf{n}(3) = \mathbf{An}(2) = \mathbf{A}(\mathbf{A}^2\mathbf{n}(0)) = \mathbf{A}^3\mathbf{n}(0) \\ &\vdots \end{aligned} \tag{2.27}$$

leading to the general solution

$$n(t) = \mathbf{A}^t \mathbf{n}(0). \tag{2.28}$$

where $\mathbf{A}^t$ denotes the product of $\mathbf{A}$ with itself t times. We can form these products because $\mathbf{A}$ is a square matrix, and the order of operations in computing the products doesn't matter because matrix multiplication is associative.

2.5 Eigenvalues and a second solution of the model

The most basic question we can ask about a population is whether it will grow or become extinct after a long time. The solution of the matrix model shows that the answer depends on the behavior of $\mathbf{A}^t$, the powers of the projection matrix as t increases. We study the properties of $\mathbf{A}^t$ through the eigenvalues and eigenvectors of the matrix $\mathbf{A}$. A number (possibly complex) λ is an *eigenvalue* of $\mathbf{A}$ if there is a non-zero vector $\mathbf{w}$ such that $\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$, and $\mathbf{w}$ is called the corresponding eigenvector. Eigenvectors are only defined up to scaling factors: if $\mathbf{w}$ is an eigenvector for λ then so is $c\mathbf{w}$ for any number c. An $n \times n$ matrix $\mathbf{A}$ must have at least 1 eigenvalue-eigenvector pair, and it can have up to n [see this chapter's Appendix for an explanation of why this is true]. The typical situation is to have n distinct eigenvalues each with a corresponding eigenvector – this is typical in the sense that if matrix entries are chosen at random according to some smooth probability distribution, the probability of the resulting matrix having n distinct eigenvalues is 1.

There is a useful formula for the eigenvalues of a 2×2 matrix $\mathbf{A}$. If $T = a_{11} + a_{22}$ is the *trace* (sum of diagonal elements) and $\Delta = a_{11}a_{22} - a_{12}a_{21}$ is the *determinant* then the eigenvalues are

$$\lambda_{1,2} = \frac{1}{2}\left(T \pm \sqrt{T^2 - 4\Delta}\right) \tag{2.29}$$

Back to equation (2.20). Assuming there are n distinct eigenvalues, the corresponding $\mathbf{w}_i$ are *linearly independent*, which means that it is possible to find constants c_i such that

$$\mathbf{n}(0) = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n = \sum_i^n c_i\mathbf{w}_i. \tag{2.30}$$

Then

$$\begin{aligned} \mathbf{n}(1) = \mathbf{A}n(0) &= \mathbf{A}(c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n) \\ &= c_1\mathbf{A}\mathbf{w}_1 + c_2\mathbf{A}\mathbf{w}_2 + \cdots + c_nA\mathbf{w}_n \\ &= c_1\lambda_1\mathbf{w}_1 + c_2\lambda_2\mathbf{w}_2 + \cdots + c_n\lambda_n\mathbf{w}_n \end{aligned} \tag{2.31}$$

Comparing (2.30) with (2.31) we see that going forward one step in time corresponds to multiplying all the coefficients c_i by the corresponding eigenvalue λ_i. We can go from $t = 1$ to $t = 2$ in the same way, getting

$$\mathbf{n}(2) = c_1\lambda_1^2\mathbf{w}_1 + c_2\lambda_2^2\mathbf{w}_2 + \cdots + c_n\lambda_n^2\mathbf{w}_n \tag{2.32}$$

and so forth. Thus the solution of the matrix model is

$$\mathbf{n}(t) = \sum_i^n c_i\lambda_i^t\mathbf{w}_i. \tag{2.33}$$

An eigenvalue λ_1 is called *dominant* if $|\lambda_i| < |\lambda_1|$ for all other eigenvalues of A. If so, it follows from (2.33) that the long-run behavior of the population is determined by the dominant eigenvalue and its eigenvector:

$$\mathbf{n}(t) \sim c_1 \lambda_1^t \mathbf{w}_1. \tag{2.34}$$

The meaning of $\sim$ in equation (2.34) is that as $t \to \infty$ the relative error goes to zero.

Equation (2.34) tells us two things about the population. First, in the long run the total population size grows exponentially at rate λ_1, just like the unstructured model (2.2). Second, the population vector becomes becomes proportional to $\mathbf{w}_1$; in particular the relative numbers in each stage become constant. For that reason, $\mathbf{w}_1$ is called the *stable stage distribution.*

The Perron-Frobenius Theorem from linear algebra provides an easy-to-check condition which guarantees existence of a dominant eigenvalue. A matrix $\mathbf{A}$ is called *power-positive* if there is an integer $m > 0$ such that all entries of the matrix $\mathbf{A}^m$ are strictly positive. The most important result is the following:

If a non-negative, square matrix $\mathbf{A}$ *is power-positive, then then* $\mathbf{A}$ *has a unique dominant eigenvalue* λ *which is real and positive, and the eigenvector* $\mathbf{w}$ *corresponding to* λ *has all positive entries.*

This criterion for existence of a dominant eigenvalue is especially useful because power-positivity only depends on which elements in the matrix are positive, *not on their numerical values.* It is also useful that there is a simple test to determine if a non-negative matrix is power-positive (Horn and Johnson 1985, p. 520), which is easy to implement on the computer:

If $\mathbf{A}$ *is a non-negative square matrix with* n *rows and columns, then* $\mathbf{A}$ *is power-positive if and only if all entries of* $\mathbf{A}^{n^2-2n+2}$ *are positive.*

Because eigenvectors are only defined up to multiplication by a constant, the statement that the dominant eigenvector has all positive entries really means that all entries in any dominant eigenvector have the same sign. A software package may give you an eigenvector (call it $\mathbf{w}^*$) with all *negative* entries, in which case $\mathbf{w} = -\mathbf{w}^*$ is the strictly positive eigenvector guaranteed by Perron-Frobenius[1].

Exercise 2.10 Write a script to verify that the following projection matrix is power positive:

$$\begin{bmatrix} 0 & 0 & 0 & 4 \\ .8 & .1 & 0 & 0 \\ 0 & .8 & .6 & .3 \\ 0 & .1 & .3 & .6 \end{bmatrix}$$

Exercise 2.11 . Find a 4×4 Leslie matrix $\mathbf{L}_1$ that is power-positive and second $\mathbf{L}_2$ that is not power-positive. In the latter case, verify your conclusion by writing a script that computes and prints the smallest value in the matrix $\mathbf{L}_2^j$ for each $j = 1, 2, \cdots, n^2 - 2n + 2$. What happens to the age structure, starting from a single newborn, in your non-power-positive example? [Note: a Leslie matrix is a matrix of the form (2.16), in which all of the p's are positive, and at least one of the f's must be positive.]

[1]Power-positivity does not guarantee that there are n distinct eigenvalues, but so long as there is a dominant eigenvalue, equation (2.34) still holds – this is not hard to show using the Jordan Canonical Form for matrices.

Exercise 2.12 According to Lande (Oecologia 75 (1988): 601-607) females of the Northern Spotted Owl begin breeding at age $a = 3$, and are estimated to have an average of 0.24 female offspring per year until they die ($f_a = 0.24$ for $a \geq 3$). The survival probability from birth to age 3 is estimated to be 0.0722, and the annual survival probability of adults (age 3 or older) is 0.942 (these values refer to the notational conventions that we used in the age-structured model, so that a newborn individual is 0). This owl has been controversial, because of the conflict between the need to preserve old-growth forests as habitat for spotted owl, and the interest of logging companies in harvesting those forests.
(a) We told you that $l_3 = p_0 p_1 p_2 = 0.0722$ but not the values of the individual p's. That's because any choice of p's with this product will result in the same population growth rate. Why is that true? (Note: the answer to this question should be verbal: no formulas are needed).
(b) Construct a projection matrix for the population based on the estimates above.
(c) Compute the owl's long-term growth rate λ from the projection matrix. Does it appear that the population is safe, or in danger of extinction?

Exercise 2.13 Killer whales (*Orcinus orca*) are long-lived marine mammals that live in stable social groups called "pods". Their stable social structure and the fact that individual whales can be photo-identified makes them especially well suited to scientific study. Demographic data on killer whale populations in the coastal waters of British Columbia and Washington state have been collected since 1973. Brault and Caswell (1993) used the 1973-1987 data and a stage-structured matrix model to investigated several demographic questions concerning the whales. They model the females with a mixed age-stage classification: yearlings, juveniles (past the first year, but not mature), mature females, and post-reproductive females. The life-cycle graph is shown in Figure (2.2) and the projection matrix $\mathbf{A}$ is given below:

$$\mathbf{A} = \begin{bmatrix} 0 & 0.0043 & 0.1132 & 0 \\ 0.9775 & 0.9111 & 0 & 0 \\ 0 & 0.0736 & 0.9534 & 0 \\ 0 & 0 & 0.0452 & 0.9804 \end{bmatrix} \tag{2.35}$$

Write a script file that
(a) Computes the dominant eigenvalue λ and stable stage distribution $\mathbf{w}$ for the whale population.
(b) Projects the population dynamics for the next 50 years assuming that the current population vector is $x_0 = (10, 60, 110, 70)$.
(c) Plots on 3 separate graphs the projected changes over time in
(i) $N(t)$ = total population size in year t,
(ii) the annual population growth rate $\lambda(t) = N(t+1)/N(t)$,
(iii) the proportion of individuals in each stage.
Does the population structure become stable? How does it change over time? How quickly does the annual growth rate $\lambda(t)$ converge to the dominant eigenvalue λ?

Exercise 2.14 Re-run your script for killer whale population dynamics with the following initial population vectors: $x_0 = (250, 0, 0, 0), (0, 250, 0, 0), (0, 0, 250, 0)$, and $(0, 0, 0, 250)$. Compare and contrast the four population projections – for example, (a) consider the stage distribution and its stability; (b) which stage seems to be the most important in terms of the future growth of the population?

Exercise 2.15 Consider a possible harvest from the killer whale population, consisting of

individuals from a single stage, e.g. all juveniles or all reproductive adults. Suppose that the initial population structure is the stable distribution $\mathbf{w}$ with a total of 250 individuals. What is the maximum number of juveniles that can be taken each year such that the population is not driven to extinction? What is the maximum number of reproductive adults? Note: Assume that harvest will take place after the breeding season, so that the model becomes $\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) - \mathbf{h}$ where $\mathbf{h} = (h_1; h_2; h_3; h_4)$ is a vector of the number of individuals harvested from each stage each year, h_i. *Assume that* $\mathbf{h}$ *is constant: same harvest each year.*

2.5.1 Left eigenvectors

There is a definition of left eigenvalues and eigenvectors analogous to that for right eigenvalues and eigenvectors: if $\mathbf{v}\mathbf{A} = \lambda\mathbf{v}$ (where $\mathbf{v}$ is a row vector of length n) then $\mathbf{v}$ is a left eigenvector and λ is the corresponding eigenvalues. There are three key properties:

- The left eigenvalues of a matrix $\mathbf{A}$ are the same as the right eigenvalues.
- The left eigenvectors of $\mathbf{A}$ are the right eigenvectors of its transpose $\mathbf{A}^T$. ($\mathbf{A}^T$ is the matrix whose $(i,j)^{th}$ element is a_{ji}. That is, the rows of $\mathbf{A}$ become the columns of $\mathbf{A}^T$, for example,
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$
- If A is power-positive so is $\mathbf{A}^T$, and hence the dominant left eigenvalue has all positive entries.

Left eigenvectors are important for the long-term dynamics of matrix models. First, they determine eigenvalue sensitivity: the extent to which each matrix entry affects the dominant eigenvalue. Let $\mathbf{v}, \mathbf{w}$ be the left and right eigenvectors corresponding to the dominant eigenvalue λ of a power-positive projection matrix $\mathbf{A}$. Then

$$\frac{\partial \lambda}{\partial a_{ij}} = \frac{v_i w_j}{\mathbf{v} \cdot \mathbf{w}} \tag{2.36}$$

where $\mathbf{v} \cdot \mathbf{w}$ is the dot product defined above (equation 2.26).

Second, the dominant left eigenvector (when it exists) has a biological interpretation as the "reproductive value" of different stages, a concept due to R. A. Fisher. Think of different stages as alternate "investments" in long-term population growth. If you could put one dollar into any one of these investments ($\sim$ one individual in any of the stages) what is their relative payoff in the long run (relative size of the resulting population in the distant future)? The answer is that the "payoff" from a stage-j individual is proportional to v_j (see the Appendix). In age-structured models it is conventional to scale $\mathbf{v}$ so that $v_0 = 1$, i.e., so that its entries give age-specific reproductive value relative to that of a newborn.

Exercise 2.16 Compute the reproductive value vector $\mathbf{v}$ for the killer whale model (2.35), scaled so that $v_1 = 1$. How does this relate to your conclusions from simulating the population starting from different initial populations?

Model	$\mathbf{n}(t+1) = \mathbf{A}\mathbf{n}(t), \mathbf{n}(0) = \mathbf{n}_0$
Solution	$\mathbf{n}(t) = \mathbf{A}^t\mathbf{n}_0$
Eigenmode Expansion	$\mathbf{n}(t) = \sum c_i\lambda_i^t\mathbf{w}_i$
Long-term exponential growth	$\sum_i n_i(t) \sim c\lambda^t$ as $t \to \infty$
Stable Stage Distribution	$\mathbf{n}(t) \sim c\lambda^t\mathbf{w}$
Stage-specific reproductive value	Proportional to $\mathbf{v}$
Eigenvalue sensitivity formula	$\frac{\partial\lambda}{\partial a_{ij}} = \frac{v_i w_j}{\mathbf{v}\cdot\mathbf{w}}$

Table 2.2: Main properties of a matrix model with power-positive projection matrix $\mathbf{A}$.

Exercise 2.17 Construct the projection matrix $\mathbf{A}$, and then find $\lambda, \mathbf{v}$, and $\mathbf{w}$ for an age-structured model with the following survival and fecundity parameters. Age-classes 0-5 are genuine age classes with survival probabilities $[p_0\ p_1 \cdots p_5]$ =[0.3, 0.4, 0.5, 0.6, 0.6, 0.7] Note that $p_j = a_{j+1,j}$, the chance of surviving from age j to age $j+1$, for these ages. Age-class 6 are adults (age 6 or older), with survival 0.9 and fecundity 12.

2.6 Some applications of matrix models

Table 2.2 summarizes the main theoretical results for the case of a power-positive projection matrix $\mathbf{A}$. From here on $\lambda, \mathbf{v}, \mathbf{w}$ without subscripts will refer to the dominant eigenvalue (formerly λ_1) and corresponding left and right eigenvectors, whose existence is guaranteed by the Perron-Frobenius Theorem. We now present two applications of these results to biological questions that have been addressed using matrix population models.

2.6.1 Why do we age?

Evolutionary biologists distinguish between *proximate* and *ultimate* explanations for phenomena. A proximate explanation tells us *how* the phenomenon occurs - for example, the physical and biochemical processes involved in meiotic cell division as part of sexual reproduction. An ultimate explanation attempts to say *why* the phenomenon occurs – for example, why some species have evolved to have sexual reproduction while others have not.

Models are important for developing ultimate explanations, because they let us consider the consequences of the alternatives that aren't seen in nature. As R. A. Fisher observed, if we want to understand why humans have two sexes rather than three or more, we must "work out the detailed consequences experienced by organisms having three or more sexes" (Fisher 1930). To understand the "why" of traits molded by evolution, we first need to compare things as they are with the other ways things might have been, and only then can we start to hypothesize why evolution produced one outcome rather than the other.

For modeling evolution of the life cycle, the growth-rate λ can be identified with Darwinian *fitness*: the contribution of offspring to future generations. On the reasonable assumption that matrix entries are determined by the organism's genotype, and if multiple genotypes are present within a population, then population genetic models predict (with some caveats) that

the genotype with the largest λ for its matrix becomes fixed in the population (Charlesworth 1982). The main caveat is a standard one in population genetics theory: if the most fit genotype is a heterozygote, then a stable polymorphism is maintained.

An ultimate explanation for sexual reproduction still eludes us: theories abound, and the rate at which old ones are rejected is balanced by the rate at which new ones take their place. But for aging (technically called senescence), there is general agreement on a simple explanation based on eigenvalue sensitivity. This explanation is derived by modeling a life cycle in which there is no aging, and then asking whether a little bit of aging would lead to increased Darwinian fitness. A "no aging" life cycle means that females start reproducing at some age m (for "maturity"), and thereafter have constant fecundity $f_j = f$ and survival $p_j = p < 1$ for all ages $j \geq m$. We have $p < 1$ to represent an age-independent rate of accidental deaths unrelated to aging.

The eigenvalue sensitivity formula lets us compute the relative eigenvalue sensitivities at different ages for this life cycle without any hard calculations, so long as $\lambda = 1$. Populations can't grow or decline without limit, so λ must be near 1. The reproductive value of adults $(v_i, i \geq m)$ is independent of age because all adults have exactly the same future prospects and therefore make the same long-term contribution to future generations. On the other hand, the stable age distribution w_j goes down with age. With $\lambda = 1$ the number of m year olds is constant, so we can compute $n_{m+k}(t) = n_m(t-k)p^k = n_m(t)p^k$. That is, in order to be age $(m+k)$ now, you must have been m years old k years ago, and you must have survived for the k years between then and now. Therefore

$$w_j \propto p^{j-m} \text{ for } j \geq m.$$

Consequently, the relative sensitivity of λ to changes in either the fecundity $a_{1,j}$ or survival $a_{j+1,j}$ of age-j females, is proportional to p^{j-m}. In both cases, as j changes the relevant w_j is proportional to p^{j-m} while the reproductive value v stays the same. This has two consequences:

1. The strength of selection against deleterious mutations acting late in life is weaker than selection against deleterious mutations acting early in life.

2. Mutations that increase survival or fecundity early in life, at the expense of an equal decrease later in life, will be favored by natural selection.

These are known, respectively, as the *Mutation Accumulation* and *Antagonistic Pleiotropy* theories of aging. In addition, there is a particular form of Antagonistic Pleiotropy, the *Disposable Soma* hypothesis, which posits that the connection between early and late vigor is mediated by investment in maintenance and repair mechanisms at the cellular and molecular levels, such as DNA repair and antioxidant systems.

Distinguishing between these theories is difficult because they agree on the fundamental prediction: if the level of unavoidable extrinsic mortality is high then the organism is predicted to be short-lived even in a protected environment, while low levels of extrinsic mortality should lead to potentially long-lived organisms. Experiments – mainly with *Drosophila* – have uniformly supported this prediction (Kirkwood and Austad 2000). In addition, there is some direct support for each mechanism:

Antagonistic pleiotropy. There is abundant evidence for antagonistic tradeoffs (e.g., Roff

2000 Chapter 3). The ideal organism would mature instantly, live forever, breed often, and have many offspring each time. We don't see such organisms because improvements on one front are paid for on another. For example, early maturation typically entails smaller adult body size and hence lower fecundity. Conversely, experimental selection against early fecundity in *Drosophila* led to increased fecundity later in life. Similar results have been obtained on other insects, birds, and mice (Roff 2000).

Disposable soma. A unique prediction of this theory, also supported by numerous studies (Kirkwood and Austad 2000) is that intrinsically long-lived organisms should have higher levels of cellular-level maintenance and repair processes: humans need to keep our liver going for decades, but a mouse does not. For example, DNA repair capacity correlates with lifespan in mammals, as does the level of poly(ADP-ribose) polymerase, an enzyme that is important in maintaining DNA integrity. Intrinsic longevity also correlates with levels of defense against thermal extremes and chemical toxins (Kirkwood and Austad 2000).

Mutation accumulation. A unique prediction of this theory is that genetic variability should increase with age. There is some evidence in *Drosophila* for genetic variance in male mating success and mortality rate, but experiments on other traits have found no evidence of mutation buildup with age.

So it seems likely that all three hypothesis play some role in actual patterns of aging, and the fundamental message of the theory is solidly supported: *In a proximate sense, most of us will die from effects of aging. In an ultimate sense, the reason we undergo aging is that we would die anyway even if we didn't age.*

Since this is a book about modeling, we add a second message: structured population models provide a framework for understanding the life cycles of organisms as adaptations for maximizing fitness subject to tradeoffs and constraints. This topic, called *life history theory*, has been an active research area since the 1960s. Stearns (2000) gives a good short overview, and Roff (2001) is a recent comprehensive text.

2.6.2 Elasticity analysis and conservation biology

The dominant eigenvalue-eigenvector pair summarize what will happen to the population if nothing changes. A value of $\lambda > 1$ implies a growing population, and $\lambda < 1$ means that the population is predicted to decline to extinction.

In the latter case, the practical issue is: what can we do to improve things? One approach to that question was based on using the eigenvalue sensitivity formula to identify matrix entries with the biggest effect on the value of λ. Fairly soon, the objection was raised that survival and fecundity entries are on intrinsically different scales: a survival must lie between 0 and 1 while fecundities can be enormous (balanced by high mortality between birth and maturation). As a result, survival rates often have higher sensitivity than fecundity: changing newborn survival from 0.1 to 0.4 would probably have a large impact, but changing adult fecundity from 1000.1 to 1000.4 won't do much at all. As this example indicates, a better measure is the proportional

sensitivity or *elasticity*, defined as

$$\begin{aligned} e_{ij} &= \frac{\text{fractional change in } \lambda}{\text{fractional change in } a_{ij}} = \frac{\partial\lambda/\lambda}{\partial a_{ij}/a_{ij}} \\ &= \frac{a_{ij}}{\lambda}\frac{\partial\lambda}{\partial a_{ij}} \end{aligned} \tag{2.37}$$

The value of e_{ij} says nothing about which matrix entries actually could be changed, or by how much, but it does identify potential targets of opportunity. So in many applications of matrix population models, the main goal of building the model is to compute the elasticities. Note also that (2.37) only applies to small changes in matrix entries, and effects of large changes have to be computed directly, by modifying the matrix and recomputing λ.

Desert tortoise

A structured population model for desert tortoise *Gopherus agassizii* (Doak et al. 1994) illustrates how an imperfect model can still be valuable because its relative predictions are robust in the face of uncertainty about parameter values.

Desert tortoise was listed as endangered in 1989 and a draft recovery plan was issued in 1993. Particularly severe declines were occurring in the western Mojave desert. Direct human impacts on the tortoise include

- habitat degradation by off-road vehicles
- habitat loss to urban or agricultural uses
- deliberate hunting (up to 14% of mortality in some areas)
- getting run over by cars or off-road vehicles.

There are also indirect impacts, including:

- habitat degradation by sheep or cattle grazing
- predation by ravens (which are associated with human presence and attack yearlings and juveniles)
- an upper respiratory tract infection that may have been introduced by release of pet tortoises into the wild.

Doak et al. (1994) had two goals. The first was to asses the potential threat to the tortoise posed by the U.S. Army's proposed expansion of Fort Irwin. The second was to compare two management scenarios being considered or implemented: reducing human disturbance and removing ravens. Human disturbance mainly affects larger individuals, while raven predation is limited to smaller ones.

The model was based mainly on government reports and previously unanalyzed mark-recapture data at 8 Bureau of Land Management permanent study plots in the western Mojave. Individuals were classified based on size and life stage (Table 2.1). The data included multiple site

× year combinations for which stage-specific growth or survival rates could be estimated (6-18 combinations for the different stages and rates). However, data on fecundity were limited. For the Mojave there were no direct observations of individual fecundity. Instead, the modelers divided yearling counts by the number of females censused at the time the yearlings would have been born. Because yearlings are much harder to find than adults, this value was regarded this as an underestimate and Doak et al. (1994) applied an arbitrary 10-fold factor to compensate for undercounting of yearlings. In addition, they considered fecundity estimates based on direct observations of egg production and two estimates of survival to hatching at a different site in the eastern Mojave, where tortoise populations were not in decline. This gave a total of 4 fecundity estimates for the breeding classes, so the overall matrix model is:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & f_6 & f_7 & f_8 \\ .716 & .567 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .149 & .567 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .149 & .604 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .235 & .560 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .225 & .678 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .249 & .851 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .016 & .860 \end{bmatrix} \tag{2.38}$$

$$\begin{aligned} (f_6, f_7, f_8) = & [0.042, 0.069, 0.069] \\ & \text{or } [0.42, 0.69, 0.69] \\ & \text{or } [1.30, 1.98, 2.57] \\ & \text{or } [2.22, 3.38, 4.38] \end{aligned}$$

Even this enormous range of possible fecundity estimates is not necessarily catastrophic, because the model predictions that you care about may not be affected. Beginning modelers often doubt such claims, but sometimes you get lucky and the parameters you know the least about, turn out to be the least important (of course, sometimes you don't get lucky). In this case, Figure 2.3 shows that the eigenvalue elasticities are consistently highest for survival of larger individuals. The management implication is to forget about ravens, and concentrate on reducing the impacts of humans on larger individuals.

The recovery plan for desert tortoise, summarized by Berry (1997), accepted this conclusion and proposed to create 14 reserves where the tortoise would be protected from detrimental human activities, including cattle grazing. Grazing limits have been hard to implement because much of the proposed reserve area was on federal lands that had long been used for grazing. Some of the proposed reserve areas have been established, while others (as of this writing) remain in dispute because court-ordered grazing restrictions have not been implemented.

Loggerhead sea turtles

These studies were also intended to evaluate two different management strategies, in this case for Loggerhead sea turtles *Caretta caretta* in the southeastern US (Heppell et al. 1998). Loggerheads are listed as threatened under the US Endangered Species Act. Conservation efforts for

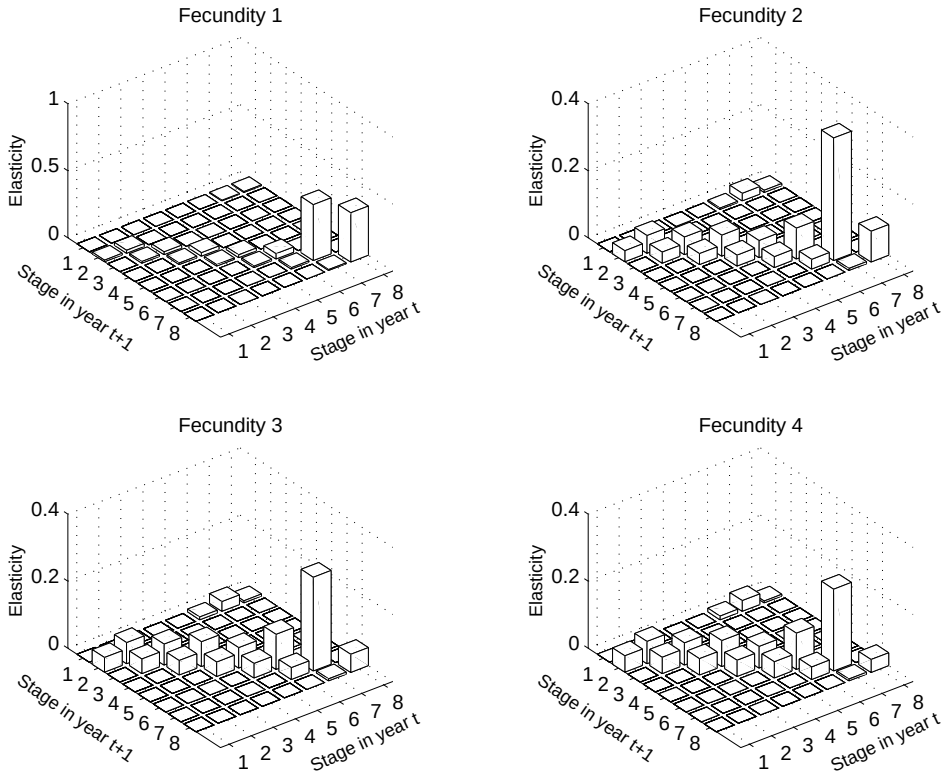

Figure 2.3: Eigenvalue elasticities for the Doak et al. (1994) stage-structured model for desert tortoise. The stages are 1=yearling, 2,3=juveniles, 4,5=immature, 6=subadult, 7=smaller adult, 8=larger adult.

marine turtles had focused on reducing egg mortality on human-impacted beaches, but after 20 or 30 years of effort the numbers of nesting turtles were not showing any increases. In addition, incidental trapping and drowning of sea-turtles in commercial fishing gear, especially shrimp trawlers, led the National Marine Fisheries Service (NMFS) to develop a turtle excluder device (TED) that released 97% of trapped turtles while keeping most shrimp in the net. However, the shrimping industry complained that TEDs led to loss of valuable harvest, damage to their gear and crew injuries, whereas nest protection projects were yielding large increases in hatchling production at very low cost.

Crouse et al. (1987) used very rough estimates of age-specific survival, growth, and fecundity rates to derive a seven-class structured model for Loggerheads. This was later revised to a

5-class structured model (Crowder et al. 1994).

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 4.665 & 61.896 \\ 0.675 & 0.703 & 0 & 0 & 0 \\ 0 & 0.047 & 0.657 & 0 & 0 \\ 0 & 0 & 0.019 & 0.682 & 0 \\ 0 & 0 & 0 & 0.061 & 0.8091 \end{bmatrix}$$

$\sigma_i =$ survival probability

$\gamma_i =$ growth probability, conditional on survival

On diagonal: $P_i = \sigma_i(1 - \gamma_i)$ survive & remain in stage

Sub-diagonal: $G_i = \sigma_i \gamma_i$ survive & grow to next stage

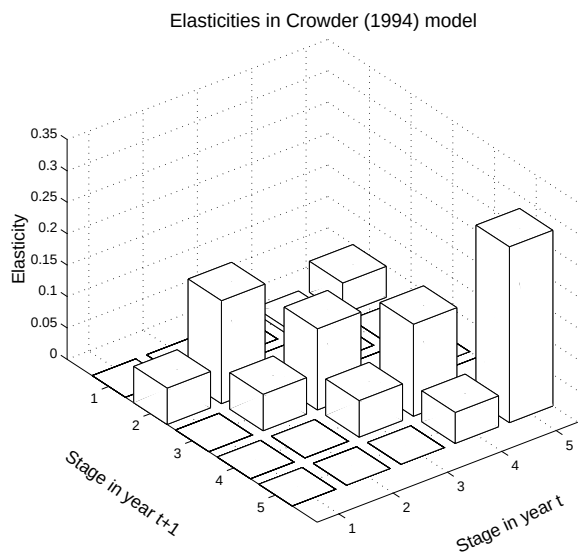

Figure 2.4: Eigenvalue elasticities for the Crowder et al. (1994) stage-structured model for Loggerhead sea turtles. Stages [stage durations] in the model are 1=egg and hatchling[1 yr], 2=small juvenile[7 yr], 3=large juvenile[8 yr], 4=subadult[6 yr], 5=adult[indefinite].

The eigenvalue elasticities for the 5-stage model, shown in Figure 2.4, provide an explanation for the ineffectiveness of strategies aimed at eggs and nestlings, and suggest that TEDs will be far more effective. Note that an increase in a stage-specific survival (σ_i) will increase both P_i and G_i values by the same proportional amount. If TEDs were to increase the annual survival of all individuals in stages 2-5, then the proportional sensitivity would be given by the sum of the diagonal and sub-diagonal bar heights, which is 0.88. That is, a 10% increase in annual survival applied to all these stages would lead to an $0.88 \times 10\% = 8.8\%$ increase in λ, which would be sufficient to bring the models λ above 1. In contrast, even if the stage-1 survival were increased to 100%, λ would still be less than 1. So nest protection is helpful but not sufficient to reverse the population decline.

Based in large part on the original analysis, the National Academy of Sciences recommended requiring TEDs, and the National Marine Fisheries Service expanded seasonal TED requirements

to all southeastern shrimp trawls starting in December 1994. By 1998, loggerhead populations were found to be stable or increasing on most monitored nesting beaches (Heppell et al. 1998). Because this was an "uncontrolled experiment" we cannot have full scientific certainty, but it is highly suggestive that a change in management plans prescribed on the basis of population models allowed the population to rebound.

Biological control and pest management

Matrix models can also be used to compare alternative options for controlling an undesired species such as a weed or an exotic invader. For some reason such applications have been rare until very recently. Rockwell et al. (1997) developed a matrix model for the lesser snow goose, which has overgrazed and damaged areas of salt marsh in Canada large enough to be clearly visible in satellite images. As one author of that study puts it, "If you can see it from space, it's a real problem" (Evan Cooch, *personal communication*). Rockwell et al. (1997) found that the elasticity of adult survival was 87%, which implies that the only means for controlling the geese is to reduce adult survival. As a result of this and other analyses, the US Fish and Wildlife Service relaxed its restrictions on goose hunting in order to increase the harvest of migrating adults. In contrast Parker (2000) found that elasticities for *Cytisus scoparius*, an invasive shrub on the west coast of North America, were similar for all life stages in populations at the outer edge of the species' range expansion. This suggests that biological control would require a suite of control agents, rather than a single agent targeting a specific stage. Other applications of matrix models for control of undesired plants are reviewed by McEvoy and Coombs (1999).

Exercise 2.18 Johnson and Braun (1999) constructed a matrix model for an exploited population of sage grouse; in this and subsequent exercises we consider a simplified version of their model. The mean matrix estimated from 23 years of field survey data is

$$\mathbf{A}_0 = \begin{bmatrix} 0.18 & 0.60 & 0.60 \\ 0.33 & 0 & 0 \\ 0 & 0.73 & 0.73 \end{bmatrix} \tag{2.39}$$

with the stages being age 0-1, age 1-2, and age$>$ 2. For this matrix, compute the dominant eigenvalue, and the matrix of elasticities e_{ij}. As the manager responsible for survival of this population, which vital rates would you be trying to improve? [Note: the yearling and adult stages could be combined for projecting population growth, but Johnson and Braun (1999) kept them separate to consider possible age-selective harvesting policies].

2.6.3 How much should we trust these models?

Heppell et al. (1996, p. 558) note that simple matrix models omit many potentially important factors for population persistence: emigration and immigration, density dependence in fecundity and survival rates, environmental variability, and *demographic stochasticity* (effects of finite population size). The use of these models to guide conservation policy, despite their limitations, brings to the fore the issue of their reliability.

Our examples focused on predicting the relative effectiveness of possible actions, but models are also asked to make absolute predictions. For example, Kareiva et al. (2000) used a matrix model

to predict whether dam removals (a proposed but highly controversial action) would be sufficient to reverse the precipitous decline in salmon stocks in the Columbia river basin. Absolute predictions are also among the listing criteria for the IUCN Red List of Threatened Species (IUCN 2001; also at `www.redlist.org`), which plays a major role in guiding conservation efforts worldwide. For example, the criteria for listing as Critically Endangered include "Quantitative analysis showing the probability of extinction in the wild is at least 50% within 10 years or three generations, whichever is the longer (up to a maximum of 100 years)". Appreciable risk of extinction is also the main criterion for listing under the US Endangered Species Act. Consequently, one of the main uses for models in conservation is to make quantitative predictions about extinction risk (Morris and Doak 2002).

We still have little evidence as to the accuracy of either relative or absolute predictions. In management decisions predicated on relative predictions, we only know what happened under the one policy that was actually implemented. Only a few studies have tried to compare absolute predictions with actual outcomes. The most comprehensive is Brook et al. (2000), who compared observed and predicted population growth rates and risk of extinction (more precisely "quasi-extinction", meaning decline below some threshold density at which extinction is considered to be inevitable). They compiled population studies for 21 animal species, of sufficient duration that they could use only the first half of the population record for each species to parameterize a simple structured population model. They then used the model to predict population changes over the time period covered by the second half of the data, and compared these with the actual data. Brook et al. (2000) found remarkably close agreement between observed and predicted total numbers of extinctions in the second halves of the data series, and no tendency for the models to systematically under- or over-estimate the growth rate. Fagan and Holmes (2002) performed a similar test, with similarly good results, for US Pacific Northwest salmon stocks.

These tests indicate that predictions referring to groups of related species or populations (like "we expect x% of amphibian species to go extinct within y years"), are in fact trustworthy, if they are made by experienced modelers with access to a few decades of good data. Unfortunately, these comparisons do not tell us about the reliability of predictions for each individual species (Ellner et al. 2002) – so long as the errors are neither optimistic nor pessimistic *on average* they will average out across a large group of species, even if the individual errors are large. Theoretical analysis (Fieberg and Ellner 2000) and simulation studies (Ellner et al. 2002) indicate that species-by-species predictions of long-term extinction risk will not be very accurate given the amounts of data generally available.

On the other hand, the Doak et al. (1994) analysis illustrates that the important predictions may still be robust, especially comparisons of options within a single population. The only way to find out if this is true, is to build the model, and then quantify how much uncertainty in the relevant predictions is produced by your uncertainty about parameter values and other aspects of the model (e.g., whether or not there is density dependence). Some general computational methods for quantifying the uncertainty in predictions are reviewed by Ellner and Fieberg (2003), and illustrated on models for salmon stocks in the Northwest US. Decisions have to be made, and an imperfect model based on limited data is better than none at all, so long as you examine model predictions across the range of plausible parameter values and model assumptions (Morris and Doak 2002, Reed et al. 2002).

2.7 Generalizing the matrix model

Before moving on to other models we briefly mention three important ways in which matrix models can be made more realistic. The first two involve dropping the assumption that matrix entries are constant.

2.7.1 Stochastic matrix models

First, we can allow random variation in matrix entries. In natural animal and plant populations, fecundity has been observed to vary enormously between years (by factors of up to 333 in plants, 38 in terrestrial vertebrates, and 2200 in birds; Hairston et al. 1996). When sufficient data are available, random variability in transition rates can be incorporated into the model. There are two main ways of doing so. A nonparametric "bootstrap" approach can be used if a population has been studied for a series of years, resulting in a series of different estimates of the matrix, $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_k$. The model is then

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

where $\mathbf{A}(t)$ is drawn at random from $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_k$. Some of the variability among the estimated $\mathbf{A}$'s will typically be sampling error rather than real variation in vital rates, so this approach will tend to over-estimate variability. The "parametric" approach is to fit statistical distributions to the observed patterns of variation and covariation in matrix entries, and simulate the model by drawing an $\mathbf{A}$ matrix for each year from the fitted distributions. Fieberg and Ellner (2001) review applications of stochastic matrix models in conservation biology.

The theory of random matrix models requires mathematics beyond that presented in this chapter. We refer interested readers to Caswell (2001) for a readable and practically-oriented summary, and to Tuljapurkar (1990) for a comprehensive review of the theory. One very important result is that there is typically still a long-term growth rate, analogous to the dominant eigenvector λ. Knowing that this rate exists, it is then possible to compute its numerical value by simulation, and also elasticities and the like.

Exercise 2.19. This exercise continues our study of the (simplified) Johnson-Braun (1999) sage grouse model. Johnson and Braun also estimated the variability over time in matrix entries; the matrix $\mathbf{S}$ below gives the standard deviations.

$$\mathbf{S} = \begin{bmatrix} 0.04 & 0.13 & 0.13 \\ 0.09 & 0 & 0 \\ 0 & 0.15 & 0.15 \end{bmatrix}$$

Write a script to simulate the sage grouse model with random variability in the vital rates, using the standard deviations above and assuming a normal distribution for each entry, and initial population vector $\mathbf{n}(0) = (430, 140, 430)$. Since each simulation run will have a different outcome, have your program do 1000 simulations, record (in a vector) the minimum population size (total number of individuals in all size classes) over the course of 100 years in each run, and plot a histogram of the minimum population sizes in the 1000 runs.

Exercise 2.20 When you computed elasticities for the deterministic version of the sage grouse model, you should have found that the highest elasticity is for adult (age >2) survival. Therefore, a 20% increase in the mean adult survival has a larger beneficial impact than a 20% increase in mean adult fecundity, based on the mean matrix $\mathbf{A}_0$. Use simulations to check if this is also true in the stochastic model of the previous exercise, by comparing (in some informative way) the results from 1000 simulations each of the two scenarios (1: 20% higher mean adult fecundity, 2: 20% higher mean adult survival).

2.7.2 Density-dependent matrix models

Second, we can allow density-dependent limits to population growth. Here the changes in vital rates are assumed to occur due to endogenous feedbacks, such as decreased survival when there is stiff competition for resources, or difficulty finding mates when the population is too sparse. In contrast to random variation, these feedbacks can give rise to completely new kinds of dynamic behavior. We will soon explore these behaviors in a different context - differential equation models - so for the moment we give only one illustrative example. Through a series of elegant experiments, a simple 3-stage model has been shown to give remarkably accurate predictions for laboratory populations of flour beetles *Tribolium castaneum* (Cushing et al. 2001, Dennis et al. 1997, 2001, Henson et al. 2001). The model is

$$\begin{aligned} L(t+1) &= bA(t)\exp(-c_{el}L(t) - c_{ea}A(t)) \\ P(t+1) &= (1-\mu_l)L(t) \\ A(t+1) &= P(t)\exp(-c_{pa}A(t)) + (1-\mu_a)A(t) \end{aligned} \tag{2.40}$$

Here L, P, A are the numbers in the larval, pupal, and adult stages of the beetle life cycle. A time-step of the model corresponds to 2 weeks of real time, which is the approximate duration of the larval and pupal stages under laboratory conditions. Population growth is limited by cannibalism. Foraging adults consume eggs and pupae (which are immobile), and foraging larvae consume eggs. The model parameters are the birth rate b in the absence of cannibalism, intrinsic larval mortality μ_l and adult mortality μ_a, and cannibalism rate parameters c_{el}, c_{al}, c_{pa}.

Figure 2.5 shows some of the possible dynamical behaviors. We plot the number of larvae, which shows the patterns most clearly; the total population size has the same qualitative dynamics in each case. In panel (A), the population quickly reaches a steady state. Matrix entries are then constant (because L, P and A are all constant), and such that the dominant eigenvalue is $\lambda = 1$ exactly: neither growth nor decline. The population self-regulates to a state where each individual (on average) replaces itself exactly. In panel (B) the population oscillates: it grows to such a high density that almost all new eggs are cannibalized, and consequently then drops to a lower density at which higher egg survival allows the population to rebound. Panels (C) and (D) show more complicated patterns of overgrowth, crash, and recovery, with periodic oscillations in (C) and aperiodic *chaotic* oscillations in (D). Note that there is nothing random in this model: the erratic behavior in (D) is entirely due to the nonlinear feedback of population density on population growth rate mediated by cannibalism.

The model has been extensively validated by experiments in which parameter values of the laboratory population are manipulated (e.g., removing or adding adults to alter the effective values of μ_a and c_{pa}), and experimental data are compared with model predictions. Panels

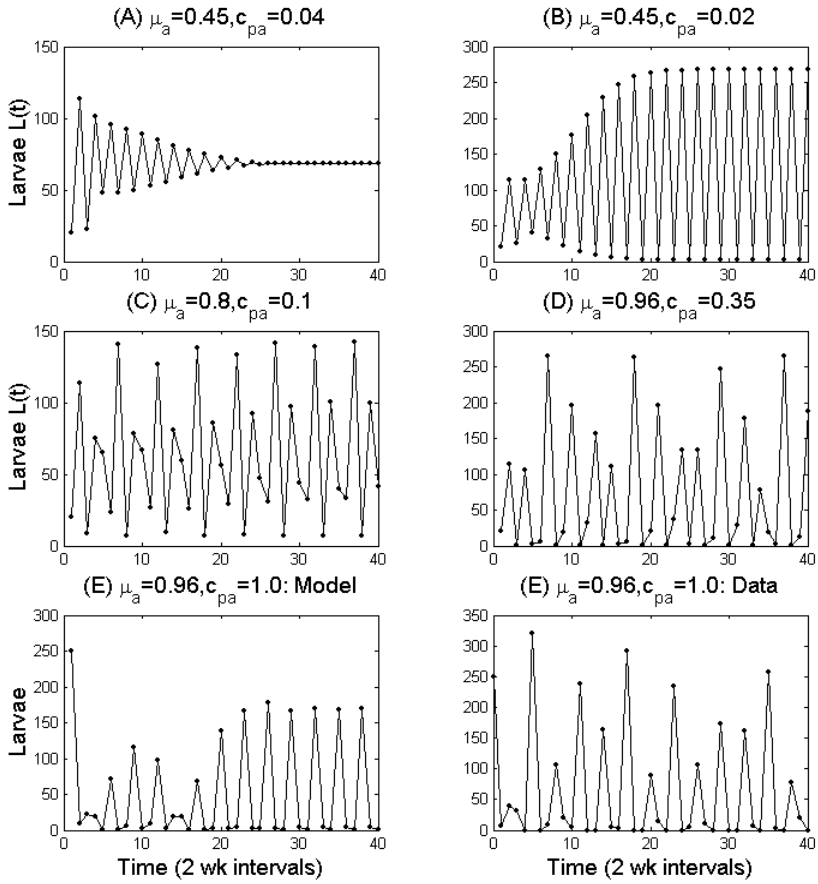

Figure 2.5: The LPA model. Panels (A)-(E) show simulations started with 20 individuals in each life stage and running for 40 two-week time steps of the model. Panels (E) and (F) compare model simulations with experimental results for high values of μ_a and $c_p a$; model simulations used the same initial conditions as the experiments, $[L(0), P(0), A(0)] = [250, 5, 100]$.

(E) and (F) show one example out of many (Cushing et al. 2002). Setting high values of μ_a and $c_p a$ and using other parameter values estimated from experimental data, cycles of period 3 with an *on/off/off* sequence of egg laying are predicted in model simulations (panel E) and observed in experiments (panel F). The *on*'s are variable in the data, rather than constant as predicted by the model. To capture this variability we would need a more complex model that simulated variability in egg laying rates, and the actual process of cannibalism upon random encounters of beetles burrowing through flour. But clearly the simple model (2.40) has captured the essentials of the experimental dynamics.

The exercises below are an introduction to how nonlinearity can affect population dynamics. First, consider a population model for discrete, non-overlapping generations defined by

$$x(t+1) = bx(t)\exp(-cx(t)) = f(x(t)), \quad b, c, > 0 \tag{2.41}$$

Models of this kind have been used for insects like gypsy moths that have one generation each

year, with adults that do not survive from one year to the next and eggs that overwinter. The parameter b is the "intrinsic" birth rate that holds when the population size is small, and the term $\exp(-cx(t))$ represents density dependent decreases in the birth rate. A *fixed point* x for the model is a value of x for which $f(x) = x$. If $x(t)$ is a fixed point, then the population size remains constant: $x(t+1) = x(t)$.

Exercise 2.21 (a) For which values of the parameters b and c does the model have a positive fixed point? (b) Can the model have more than one positive fixed point? [There is a fixed point for this model where $x = bx\exp(-cx)$; this occurs if $x = 0$ or if $\cdots$].

A fixed point x_e is *stable* if nearby values of the population evolve to the fixed point: if $x = x(0)$ is close to x_e, then the sequence $x(t)$ has x_e as a limit. Then a fixed point is unstable, the behavior of the population model is more complicated.

Exercise 2.22 . Write a script file to simulate model (2.41). For parameter values $c = 0.01$ and $b = 4, 8, 12, 16, 20, 24$, and initial population size $x(1) = 1$, have the script compute and graph $x(t)$ for $1 \leq t \leq 100$. Have all six graphs appear in a single window, and put the value of b in the title of each graph.

Exercise 2.23 Write a script file to simulate the LPA model and replicate the results in Figure 2.5.

2.7.3 Continuous size distributions

In many applications of matrix models individuals are categorized based on a continuously varying attribute such as body size, rather than by discrete life-stages. These modeler-defined "stages" are an artifice imposed to allow the convenience of using of a matrix model. Our final case study in this chapter is a cautionary tale about the limits of this approach and a possible solution.

Northern Monkshood *Aconitum noveboracense* is an herbaceous perennial plant listed as threatened under the U.S. Endangered Species Act (Dixon and Cook 1990). Transition rates were estimated by repeated census of marked individuals in a series of populations in the Catskill mountains, 1990) with stem diameter and number of leaves recorded as measures of size. Dixon and Cook (1990) and Dixon et al. (1997) analyzed the data using matrix models with a small number of size classes. Figure 2.6 shows elasticity analysis for two matrix models differing only in their choice of stage-class boundaries. The top panel shows the elasticities using the "stages" selected by Dixon and Cook (1990), with parameters estimated from 3 years of data at one of their sites (Easterling et al. 2000). The bottom panel shows elasticities based on exactly the same data, but with boundaries set so that matrix entry estimates are based on the same number of observations for each stage class.

The elasticities and their implications for managing the population appear to very different depending on the choice of stage boundaries. This discrepancy occurs because the size categories used to build the population model are also used for sensitivity analysis. These objectives may conflict. To predict future population trends, size categories should contain individuals who are similar in survival and fecundity under current conditions. To predict the effect of management

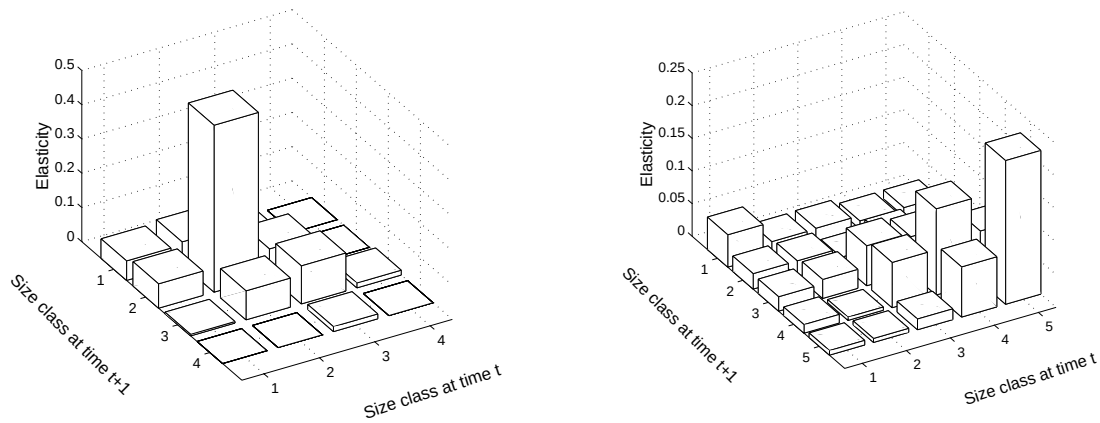

Figure 2.6: Top panel: Eigenvalue elasticities for a matrix model of Northern Monkshood using the size-based classification of Dixon and Cook: seedling, juvenile, small (<2mm), medium (2-4mm), and large (>4mm). The seedling stage is omitted because the subset of populations used to parameterize the model did not have any reproduction by seed during the study period. Bottom panel: Eigenvalue elasticities when class boundaries are set based on stem diameter, so that each class contained the same number of censused individuals.

actions, size categories should contain individuals who are similar in their response to possible actions. In principle we could achieve both goals by using a large number of small size categories, so individuals within a category are necessarily similar. But then we would have very little data on each category and therefore poor parameter estimates – another example of the tradeoff between model error and parameter error.

To avoid these problems, Easterling et al. (2000) proposed that the matrix model should be replaced by an integral projection model (IPM) in which size is a continuous variable, and the population state is described by a continuous size distribution $n(y,t)$ such that $\int_a^b n(y,t)dy$ is the number of individuals whose size is between a and b at time t. In place of a projection matrix $\mathbf{A}$, the IPM has a "projection kernel" function K, defined so that

$$n(y,t+1) = \int_s^S K(y,x)n(x,t)dx \tag{2.42}$$

where s and S are the minimum and maximum possible sizes of individuals. This integral model does the same thing as the matrix model (equation 2.19), computing $n(y,t+1)$ as the total contribution of size-y individuals "now" from individuals of any size x "last year". As in the matrix model, K includes the sum of contributions from survival and fecundity.

Easterling et al. (2000) describe how a projection kernel can be estimated from the same data on monkshood that were used to estimate a projection matrix. The elasticity surface can then be calculated from the kernel in much the same way that it is calculated from the matrix entries in a matrix model (Easterling et al. 2000). The elasticity surface for monkshood (Figure 2.7)

indicates that there are regions of high elasticity that could be the focus of management actions. The narrow, high ridge occurs at the typical size of newborn plants and thus corresponds to survival from age 1 to age 2. The broader diagonal mound corresponds to survival of mid-size individuals, with stem diameter up to 2 or 2.5mm.

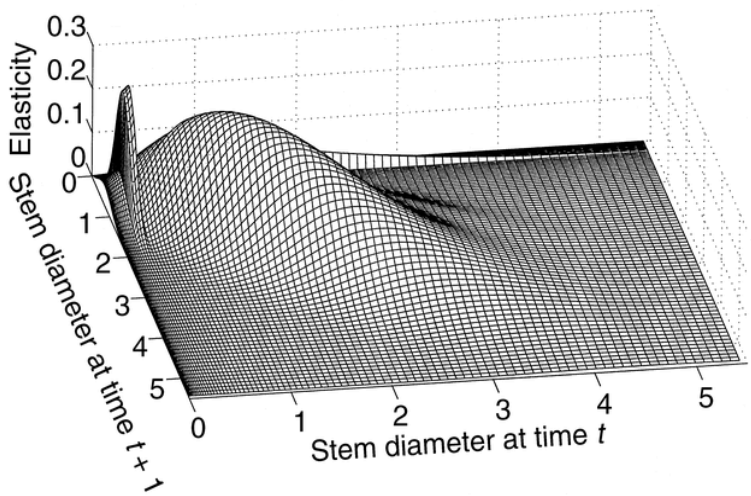

Figure 2.7: Elasticity surface for the integral projection model for Northern Monkshood.

Mark Rees and collaborators (Rees and Rose 2002, Rees et al. 2004, Childs et al. 2003,2004) have recently extended the integral model to include size- and age-dependent demography, as well as stochastic variation. An advantage of the integral model in these cases is that age and size dependence are described by smooth functions that may involve only a few parameters. For example in the Childs et al. (2004) model for the monocarpic thistle *Carlina vulgaris*, data on the probability of flowering as a function of age a and size x (the log-transformed length of the longest leaf) was fitted by the function

$$p_f(a,x) = \frac{\exp(\beta_0 + \beta_a a + \beta_x x)}{1 + \exp(\beta_0 + \beta_a a + \beta_x x)}$$

while the seed production of flowering plants was size-dependent and fitted by $s(x) = \exp(A + Bx)$. The complete pattern of age- and size-dependent fecundity is specified by 5 parameters, whereas a matrix model would require cross-classifying individuals by age and size and a separate fecundity parameter for each age-size class. Then by using the integral model to predict how natural selection acts on the value of the flowering function parameters, it was possible to show that the observed strategy, where flowering depends on both age and size, is an adaptation to random variation in mortality and growth – models ignoring the random variation gave incorrect predictions of the distribution of plant size at flowering, whereas a model incorporating observed levels of variation gave very accurate predictions (Childs et al. 2004).

2.8 References

Berry, K.H. 1997. The Desert Tortoise Recovery Plan: An Ambitious Effort to Conserve Biodiversity in the Mojave and Colorado Deserts of the United States. pp.430-440 in: Proceedings: Conservation, Restoration, and Management of Tortoises and TurtlesAn International Conference. New York Turtle and Tortoise Society.

Brault, S. and H. Caswell. 1993. Pod-specific demography of killer whales (*Orcinus orca*). Ecology 74: 1444-1454.

Brook, B. W., J. J. O'Grady, A. P. Chapman, M. A. Burgman, and R. Akcakaya, and R. Frankham. 2000. Predictive accuracy of population viability analysis in conservation biology. Nature 404: 385-387.

Caswell, H. 2001. Matrix Population Models: Construction, Analysis, and Interpretation, $2^n d$ edition. Sinauer Associates, Sunderland, MA.

Charlesworth, B.C. 1994. Evolution in age-structured populations. 2nd edition. Cambridge University Press, Cambridge UK. (1st edition 1980).

Childs, D.Z., M. Rees, K.E. Rose, P.J. Grubb, and S.P. Ellner. 2003. Evolution of complex flowering strategies: an age and size-structured integral projection model. Proceedings of the Royal Society B 270: 1829-1839.

Childs, D.Z., M. Rees, K.E. Rose, P.J. Grubb, and S.P. Ellner. 2004. Evolution of size dependent flowering in a variable environment: Construction and analysis of a stochastic integral projection model. Proceedings of the Royal Society B 271: 425-434.

Crouse, D.T., L.B. Crowder, and H. Caswell. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology 68: 1412-1423.

Crowder, L.B., D.T. Crouse, S.S. Heppell, and T.H. Martin. 1994. Predicting the impact of turtle excluder devices on loggerhead sea turtle populations. Ecological Applications 4: 437 - 445.

Dennis, B., R.A. Desharnais, J.M. Cushing, and R.F. Costantino. 1997. Transitions in population dynamics: equilibria to periodic cycles to aperiodic cycles. Journal of Animal Ecology 66: 704-729.

Dennis, B. R.A. Desharnais, J.M. Cushing, S.M. Henson and R.F. Costantino. 2001. Estimating chaos and complex dynamics in an insect population. Ecological Monographs 71: 277303.

Dixon, P. M. and R. E. Cook. 1990. Life history and Demography of Northern Monkshood (*Aconitum noveboracense*) in New York State. Cornell Plantations, Ithaca, NY.

Dixon, P.M., N. Friday, P. Ang, S. Heppell, M. Kshatriya. 1997. Sensitivity analysis of structured-population models for management and conservation. Pp. 471-513 in: S. Tuljapurkar and H. Caswell (eds.) Structured Population Models in Marine, Terrestrial and Freshwater Systems. Chapman and Hall, London.

Doak, D., P. Kareiva, and B. Klepetka. 1994. Modeling population viability for the desert tortoise in the Western Mojave desert. Ecological Applications 4, 446-460.

Easterling, M.R., S.P. Ellner, and P. Dixon. 2000. Size-specific sensitivity: applying a new structured population model. Ecology 81: 694-708.

Ellner, S.P, J. Fieberg, D. Ludwig, and C. Wilcox. 2002. Precision of population viability analysis. Conservation Biology 16: 258-261.

Ellner, S.P. and J. Fieberg. Using PVA for management despite uncertainty: effects of habitat, hatcheries, and harvest on salmon. Ecology, *in press*.

Fieberg, J. and S.P. Ellner. 2000. When is it meaningful to estimate an extinction probability? Ecology 81: 2040-2047.

Fieberg, J. and S.P. Ellner. 2001. Stochastic matrix models for conservation and management: a comparative review of methods. Ecology Letters 4: 244-266.

Hairston, N.G., S. Ellner, and C.M. Kearns. 1996. Overlapping generations: the storage effect and the maintenance of biotic diversity. pp. 109-145 in: O.E. Rhodes, R.K. Chesser, and M.H. Smith (eds). Population Dynamics in Ecological Space and Time. University of Chicago Press, Chicago.

Henson, S.M., R.F. Costantino, J.M. Cushing, R.A. Desharnais, B. Dennis, B., and A.A. King, 2001. Lattice effects observed in chaotic dynamics of experimental populations. Science 294: 602605.

Heppell, S.S., L.B. Crowder, and D.T. Crouse. 1996. Model to evaluate headstarting as a management tool for long-lived turtles. Ecological Applications 6: 556-565.

Heppell, S.S., D.R. Crouse, and L.B. Crowder. 1998. Using matrix models to focus research and management efforts in conservation. pp. 148-168 in: S. Ferson and M. Burgman (eds.) Quantitative Methods for Conservation Biology. Springer, New York.

Holmes, E. E. and W. F. Fagan. 2002. Validating population viability analysis for corrupted data sets. Ecology 83: 23792386.

Horn, R.A. and C.R. Johnson. 1985. Matrix Analysis. Cambridge University Press, Cambridge UK.

IUCN 2001. IUCN Red List Categories and Criteria : Version 3.1. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK.

K. H. Johnson and C.E. Braun. 1999. Viability and conservation of an exploited sage grouse population. Conservation Biology 13: 77-84.

Kareiva, P., M. Marvier, and M. McClure. 2000. Recovery and management options for spring/summer Chinook salmon in the Columbia River basin. Science 290: 977-979.

Kirkwood, T.B.L. and S. N. Austad. 2000. Why do we age? Nature 408: 233-238.

McEvoy, P.B. and E.M. Coombs. 1999. Biological control of plant invaders: regional patterns,

field experiments, and structured population models. Ecological Applications 9: 387401.

Menges, E.S. 2000. Population viability analyses in plants: challenges and opportunities. Trends in Ecology and Evolution 15: 51-56.

Morris, W.F. and D.F. Doak. 2002. Quantitative Conservation Biology: Theory and Practice of Population Viability Analysis. Sinauer, Sunderland MA.

Parker, I.M. 2000. Invasion dynamics of *Cytisus scoparius*: a matrix model approach. Ecological Applications 10: 726-743.

Reed, J.M., L. S. Mills, J. B. Dunning Jr., E. S. Menges, K. S. McKelvey, R. Frye, S. R. Beissinger, M.-C. Anstett, P. Miller. 2002. Emerging Issues in Population Viability Analysis. Conservation Biology 16: 7-19.

Rees, M., and K.E. Rose. 2002. Evolution of flowering strategies in *Oenothera glazioviana*: an integral projection model approach. Proceedings of the Royal Society B 269: 1509-1515.

Rees, M., D.Z. Childs, K.E. Rose, and P.J. Grubb. 2004. Evolution of size dependent flowering in a variable environment: Partitioning the effects of fluctuating selection. Proceedings of the Royal Society B 271: 471- 475.

Rockwell, R.F., E.G. Cooch, and S. Brault. 1997, Dynamics of the mid-continent population of lesser snow geese: projected impacts of reductions in survival and fertility on population growth rates. pp. 73-100 in: B. Batt (ed.) Arctic Goose Habitat Working Group Report. US Fish and Wildlife Service, Washington DC and Canadian Wildlife Service, Ottowa, Ontario.

Roff, D.A. 2001. Life History Evolution. Sinauer Associates, Sunderland MA.

Stearns, S.C. 2000. Life history theory: successes, limitations, and prospects. Naturwissenschaften 87: 476-486.

Tuljapurkar, S. 1990. Population Dynamics in Variable Environments. Springer-Verlag, New York.

USFWS. 1988. Endangered Species Act of 1973 as amended through the 100th Congress. US Department of Interior, Fish and Wildlife Service, Washington DC.

USFWS. 1994. The desert tortoise (Mojave population) recovery plan. U.S. Fish and Wildlife Service, Region 1-Lead Region, Portland, Oregon.

Williams, B.K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and Management of Animal Populations. Academic Press, NY.

Wood, S.N. 1997. Inverse problems and structured population dynamics. pp. 555-586 in: S. Tuljapurkar and H. Caswell (ed.s) Structured Population Models in Marine, Terrestrial and Freshwater Systems. Chapman and Hall, NY.

2.9 Appendix

Here we fill in some mathematical details for readers who have had a course in linear algebra.

1. Existence and number of eigenvalues
λ is an eigenvalue of $\mathbf{A}$ if and only if it satisfies the *characteristic equation* $\det(\mathbf{A} - \lambda\mathbf{I})$=0, since this is equivalent to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{w} = 0$ having a nonzero solution $\mathbf{w}$. Here $\mathbf{I}$ is the *identity matrix* having 1's on its diagonal (running from top left to bottom right) and all other elements zero. It is so named because $\mathbf{AI} = \mathbf{A}$ and $\mathbf{IB} = \mathbf{B}$ for any matrices $\mathbf{A}, \mathbf{B}$. The characteristic equation is a polynomial of degree n, so it has has at least 1 and at most n distinct solutions in the set of complex numbers).

2. Reproductive value
Here we show that reproductive values v_i from the left eigenvector of the projection matrix give the long-term relative sizes of populations descended from a single founding individual in stage i. First, observe that $\mathbf{v}$ defines an exponentially growing weighted sum of the population vector $\mathbf{n}(t)$. That is,

$$V(t+1) \equiv \mathbf{v} \cdot \mathbf{n}(t+1) = \mathbf{v} \cdot \mathbf{An}(t) = \mathbf{v} \cdot \lambda\mathbf{n}(t) = \lambda\mathbf{v} \cdot \mathbf{n}(t) = \lambda V(t)$$

So if $V_j(t)$ is the value of $V(t)$ when the population is started from a single stage-j individual at time 0, we have

$$V_j(t) = \lambda^t v_j.$$

But we also have

$$V_j(t) = \mathbf{v} \cdot \mathbf{n}(t) \sim \mathbf{v} \cdot C_j\lambda^t\mathbf{w} = C_j\lambda^t\mathbf{v} \cdot \mathbf{w}$$

where the constant C_j depends on the stage j of the founding individual. Equating the two expressions above for $V_j(t)$ we see that C_j is proportional to v_j, as claimed.

Another way of expressing this result is that for large t,

$$\mathbf{A}^t \sim C\lambda^t(\mathbf{w} * \mathbf{v}) \tag{2.43}$$

regarding $\mathbf{w}$ as a column vector (size $n \times 1$) and $\mathbf{v}$ as a row vector (size $1 \times n$)) so that $\mathbf{w} * \mathbf{v}$ is a size-n square matrix. From the definition of matrix multiplication we see that $\mathbf{w} * \mathbf{v}$ is the matrix whose j^{th} column is $\mathbf{w} \times v_j$. If we start with a single single stage-j individual, the population at time t is the jth column of $\mathbf{A}^t$. Since the population structure converges to $\mathbf{w}$ for any initial population, and the total number of individuals is proportional to $v_j\lambda^t$, the conclusion is that A^t is given by (2.43).