Matrix Methods for Population Analysis

Hal Caswell

Matrix models for structured populations were introduced by P. H. Leslie in the 1940's (Leslie 1945, 1948). Although they are in some ways the simplest of the mathematical approaches to structured population modeling (see Chapter 1), their analysis requires computational power. For this reason, and because ecologists of the day viewed matrix algebra as an esoteric branch of advanced mathematics, they were largely neglected until the late 1960's, when they were rediscovered by ecologists (Lefkovitch 1965) and human demographers (Goodman 1967; Keyfitz 1967). In the 1970's, matrix models were adopted by plant ecologists, who discovered that they could easily handle the complexity of plant life cycles in which size or developmental stage was more important than chronological age in determining the fate of individuals (Sarukhán & Gadgil 1974; Hartshorn 1975; Werner & Caswell 1977).

This chapter introduces the construction and analysis of matrix population models. I will not try to be comprehensive; I have done that elsewhere in book form (Caswell 1989a) and twice in simplified form with a focus on particular taxa (Caswell 1986; McDonald & Caswell 1993). Instead, I try to convey the basics of matrix population models clearly and briefly. Wherever possible, I use different derivations than before (Caswell 1989a), so you may find some new ways to understand the source of some familiar results. I rarely cite my book (Caswell 1989a) (in spite of having done so three times in this paragraph); almost every topic presented here could be followed by the instruction, "see the book for more

information." My focus here is on methods; I am sparing in my use of examples, because they can be found in many of the other chapters in this volume.

A note about notation. I use boldface symbols to denote vectors (lower case, as in $\mathbf{n}$) and matrices (uppercase, as in $\mathbf{A}$). Entries of vectors and matrices are lowercase letters with subscripts, so the ith entry of $\mathbf{n}$ is n_i, and the element in the ith row and jth column of $\mathbf{A}$ is a_{ij}. Sometimes I use parenthetical superscripts to label matrices or vectors. Thus, $\mathbf{A}_m$ or $\mathbf{A}^{(m)}$ might both be used to denote the mth in a series of matrices; the ijth element of this matrix is written $a_{ij}^{(m)}$. The transpose of the matrix $\mathbf{A}$ is $\mathbf{A}^{\mathrm{T}}$. If $x = a + bi$ is a complex number, the complex conjugate is denoted by $\bar{x} = a - bi$. The complex-conjugate transpose of $\mathbf{A}$ is $\mathbf{A}^*$. The scalar product of two vectors is $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x}$.

1 Formulating Matrix Models

A matrix population model operates in discrete time, projecting a population from t to $t + 1$. The first step in formulating a matrix model is to define the time scale for the projection; this is called the *projection interval*. Models for the same population with different projection intervals may look quite different.

The second step is to choose a set of state variables for individuals (*i-state variables*); these provide the information necessary to determine the response of an individual to the environment, over a projection interval. Examples of i-state variables include age, size, developmental stage, and geographical location.

A matrix model uses discrete stages, so the third step is to define a set of discrete categories for each i-state variable. Some i-states are naturally discrete (e.g., instars), while others are naturally continuous and must be made discrete (e.g., size). Dividing continuous variables into discrete categories involves trade-offs. A model treats all individuals within a category as identical, so creating only a few large categories reduces the accuracy of the i-state dynamics. Creating many small categories, alternatively, leads to a large model and may make it hard to estimate parameter values because sample sizes in each category are small.

The stages describe the life cycle, or as much of it as we believe to be demographically important. The next step is to translate them into a model. The life-cycle graph is a useful tool for this translation.

The Life-Cycle Graph

A life-cycle graph describes the transitions an individual can make, during a projection interval, among the i-state categories that define its life cycle. To construct the graph, first draw a numbered point (a "node" in graph-theory terminology) for each i-state category. If, for example, size is the i-state variable, then the life-cycle graph contains a node for each size class. If age and size are both i-state variables, then the life-cycle graph contains a node for each age-size category. Draw arrows, "directed arcs," between nodes to indicate where it is possible for an individual in one stage to contribute individuals to another stage over a single projection interval. The head of the arrow shows the direction in which individuals move. If individuals can contribute in both directions between two stages, draw two arrows, rather than an arrow with a head on both ends. Contributions from one stage to another can result from the movement of individuals from one stage to another (e.g., by growth or aging) or from production of new individuals (e.g., by birth).

With each arrow is associated a coefficient; the coefficient on the arrow from stage j to stage i is denoted a_{ij} (the ordering of the subscripts is important; it corresponds to the arrangement of coefficients in the resulting matrix model). The coefficient a_{ij} gives the number of stage i individuals at $t + 1$ per stage j individual at time t.

So far, we have made no decisions about the nature of these coefficients; I return to this below.

Figure 1a shows a life-cycle graph for an age-classified model with the age interval equal to the projection interval. Individuals in one age class can contribute to another only by surviving to the next older age class or by reproduction to the first age class. Figure 1b shows the graph for a size-classified model in which individuals may grow to the next size class, remain in their own size class, and possibly reproduce new individuals in the first size class. Suppose some individuals in the first size class grow so rapidly that after one projection interval they are in the third size class. This would require the modification shown in Figure c, 1 where an arrow has been drawn from stage 1 to stage 3.

The interpretation of the coefficients depends on the identity of the stages and the processes involved in the transitions. In Figure 1a, the P_i are age-specific survival probabilities and the F_i are age-specific fertilities. In Figures 1b and 1c, the G_i are probabilities of surviving *and* growing, the P_i are probabilities of surviving
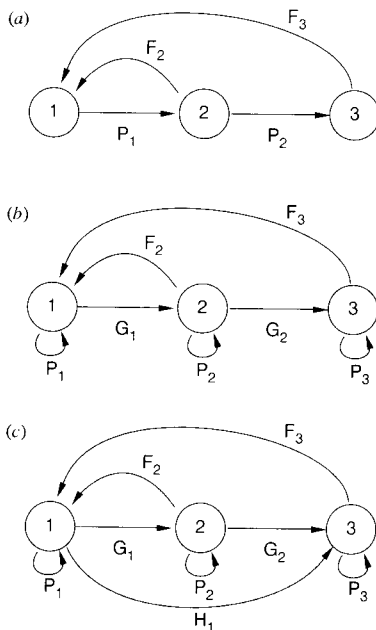
FIGURE 1. *Three life-cycle graphs.* (a) *An age-classifed model with three age classes; the P_i are age-specific survival probabilities, and the F_i are age-specific fertilities.* (b) *A size-classified model with three size classes; the G_i are size-specific probabilities of survival and growth, the P_i are size-specific probabilities of surviving and remaining in the same size class, and the F_i are size-specific fertilities.* (c) *The same life-cycle graph as in* (b), *but with an additional transition (H_1) from size class 1 to size class 3.*

and not growing, and the F_i are size-specific fertilities. In Figure 1c, the coefficient H_1 is the probability of surviving *and* growing enough to move from size class 1 to size class 3. Demographers use the term *vital rates* to refer collectively to the rates of survival, growth, reproduction, and any other important demographic processes.

A Set of Difference Equations

The life-cycle graph corresponds directly to a model written as a set of difference equations. For the size-classified graph in Figure 1b, remembering the definitions of the coefficients, the set of equations

describing the population is

$$n_1(t+1) = P_1 n_1(t) + F_2 n_2(t) + F_3 n_3(t) \,,$$
$$n_2(t+1) = G_1 n_1(t) + P_2 n_2(t) \,,$$
$$n_3(t+1) = G_2 n_2(t) + P_3 n_3(t) \,. \tag{1}$$

It is worth looking at these equations for a moment. The first states that the number of individuals in stage 1 at $t+1$ is the sum of those remaining in stage 1 from time t and those contributed by reproduction from stages 2 and 3. The second equation states that the number in stage 2 at $t+1$ is the sum of those growing into stage 2 from stage 1 and those remaining in stage 2 from time t. The third equation says the same thing for stage 3.

The equations corresponding to Figure 1c are

$$n_1(t+1) = \qquad P_1 n_1(t) + F_2 n_2(t) + F_3 n_3(t) \,,$$
$$n_2(t+1) = \qquad\qquad G_1 n_1(t) + P_2 n_2(t) \,, \qquad (2)$$
$$n_3(t+1) = \qquad H_1 n_1(t) + G_2 n_2(t) + P_3 n_3(t) \,.$$

It would be possible to write down these equations directly, without using the life-cycle graph, if we were clear about the nature of the possible transitions, which in turn depends on the definition of the stages. But using the life-cycle graph makes it easier, and helps to guard against mistakes in defining the stages and transitions.

The Matrix Model

The system of difference equations derived from the life-cycle graph can be written more simply in matrix form:

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

where

$$\mathbf{n}(t) = \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_k(t) \end{bmatrix} \tag{2}$$

is a *stage-distribution vector* and $\mathbf{A}$ is a *population-projection matrix*. The elements of this matrix can be obtained from the system of difference equations or directly from the life-cycle graph: the ijth entry of $\mathbf{A}$ is the coefficient on the arrow from stage j to

stage i. The reason for the order of the subscripts is to guarantee this correspondence.

Applying this rule to the life-cycle graphs in Figure 1 yields

$$
\mathbf{A}^{(a)} \;=\; \begin{bmatrix} 0 & F_2 & F_3 \\ P_1 & 0 & 0 \\ 0 & P_2 & 0 \end{bmatrix}, \tag{3}
$$

$$
\mathbf{A}^{(b)} \;=\; \begin{bmatrix} P_1 & F_2 & F_3 \\ G_1 & P_2 & 0 \\ 0 & G_2 & P_3 \end{bmatrix}, \tag{4}
$$

$$
\mathbf{A}^{(c)} \;=\; \begin{bmatrix} P_1 & F_2 & F_3 \\ G_1 & P_2 & 0 \\ H_1 & G_2 & P_3 \end{bmatrix}. \tag{5}
$$

The age-classified model produces a special matrix, with positive entries only on the first row (fertilities) and the subdiagonal (survival probabilities). Such a matrix is often called a Leslie matrix, in recognition of the early papers of Leslie (1945, 1948).

I have said nothing about how the numerical values of the coefficients a_{ij} are obtained. This obviously important question deserves its own chapter (see Chapter 19, by Wood, for one approach), but here I assume that the matrix is at hand and focus on how to analyze it.

Types of Matrix Models

The coefficient a_{ij} is the contribution of each individual in stage j to the number of individuals in stage i during one projection interval. What happens in the next projection interval? Depending on the answer to this question, matrix models fall into three classs, each with its own analytical approach.

Linear, constant-coefficient models. If the coefficients a_{ij} are constants, the resulting model is linear and time-invariant:

$$
\mathbf{n}(t+1) = \mathbf{A}\mathbf{n}(t). \tag{6}
$$

This is the simplest case; it can be analyzed in great detail, and it is widely used. But in reality the vital rates are not constants, so the biological interpretation of these results requires great care.

Nonlinear models. If the a_{ij} are not constant but depend on the current state of the population, the resulting model is nonlinear:

$$\mathbf{n}(t+1) = \mathbf{A}_n \mathbf{n}(t), \tag{7}$$

where $\mathbf{A}_n$ is the transition matrix evaluated at $\mathbf{n}$. The nonlinearity may result from density dependence (e.g., competition for resources), frequency dependence (e.g., competition for mates), or both.

Time-varying models. The coefficients may also change with time, independently of $\mathbf{n}(t)$, resulting in the model

$$\mathbf{n}(t+1) = \mathbf{A}_t \mathbf{n}(t). \tag{8}$$

Deterministic, periodic variation is often used to model seasonality or other kinds of environmental periodicity. Alternatively, the coefficients may vary stochastically, reflecting some random environmental process (see Chapter 3, by Tuljapurkar). Time-varying models may be either linear or nonlinear.

Objectives of Analysis

The analysis of each of these types of model, although requiring different mathematical tools, addresses a set of similar questions. Imagine that you are in possession of a matrix population model. What you should do with it depends on the question you want to answer.

1. *Transient analyses* describe the short-term dynamics resulting from specific initial conditions.

2. *Asymptotic analyses* describe the long-term dynamics of the population.

   (a) *Population growth rate*: what is the asymptotic rate of population growth or decline?

   (b) *Population structure*: what are the relative abundances of the different stages in the life cycle?

   (c) *Ergodicity*: are the dynamics, including the growth rate and the population structure, asymptotically independent of initial conditions?

   (d) *Attractors* (mainly in density-dependent models): what are the qualitative properties of the asymptotic dynamics (fixed point, cycle, quasiperiodicity, chaos, etc.)?

3. *Perturbation analyses* examine the effects of changes in param-
   eter values or initial conditions on the results of the models.
   Three questions are of particular importance:

   (a) *Sensitivity and elasticity analysis of population growth rate*:
       how does the growth rate respond to changes in vital rates?

   (b) *Stability analysis of equilibria*: if initial conditions are per-
       turbed slightly away from an equilibrium point, does the
       solution return to or depart from the neighborhood of the
       equilibrium?

   (c) *Bifurcation analysis*: what happens to the asymptotic be-
       havior of a nonlinear model as a parameter in the model is
       changed?

The methods used to address these questions depend on the nature
and sometimes on the details of the model, but any population-
modeling project that does not address short-term dynamics, long-
term dynamics, and the effects of perturbations on those dynamics
has left something out.

2 Analysis: The Linear Case

We begin with the linear time-invariant model (6), in which **A** is
a constant matrix. There are two justifications for spending time
on this model, in spite of the fact that the vital rates of any real
population are certainly not constant. The first is theoretical: un-
derstanding population dynamics in the simplest case is a funda-
mental step in understanding more-complicated cases. The second
is practical: when interpreted as a projection rather than a predic-
tion (Keyfitz 1968; Caswell 1989*a*), the results of a linear model
provide a valuable characterization of the current environment by
calculating the purely hypothetical consequences of maintaining
that environment forever. Linear matrix population models are
frequently used in this way, as a form of demographic analysis
of vital-rate data, rather than as a prediction of future population
dynamics.

Exponential Solutions and the Characteristic Equation

One approach to equation (6) is to conjecture that, like other linear
equations, it has an exponential solution(s),

$$\mathbf{n}(t) = \lambda^t \mathbf{w} \tag{9}$$

for some fixed vector $\mathbf{w}$. Substituting this into (6) gives

$$\lambda^{t+1}\mathbf{w} = \lambda^t \mathbf{A}\mathbf{w}.$$

A scalar λ and a vector $\mathbf{w}$ that satisfy this relation are called an eigenvalue and eigenvector of $\mathbf{A}$, respectively. They must satisfy

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{w} = 0,$$

which has a nonzero solution for $\mathbf{w}$ only if the determinant of the matrix $\mathbf{A} - \lambda\mathbf{I}$ equals zero. This is called the characteristic equation:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{10}$$

The Spectral Decomposition of $\mathbf{A}$

An alternative approach is to begin by solving equation (6), starting from a specified initial population $\mathbf{n}(t_0)$. By repeatedly applying (6), we see that $\mathbf{n}(t_0 + 1) = \mathbf{A}\mathbf{n}(t_0)$, $\mathbf{n}(t_0 + 2) = \mathbf{A}^2\mathbf{n}(t_0)$, and in general

$$\mathbf{n}(t_0 + t) = \mathbf{A}^t\mathbf{n}(t_0). \tag{11}$$

Thus, to understand population dynamics over time we need only understand the behavior of $\mathbf{A}^t$.

One approach to the problem is via the *spectral decomposition* of $\mathbf{A}$, which makes it possible to evaluate any function of $\mathbf{A}$, including $\mathbf{A}^t$. First, note a few facts about the eigenvalues and eigenvectors of a matrix. The vectors $\mathbf{w}$ and $\mathbf{v}$ are right and left eigenvectors of $\mathbf{A}$ if there is a (possibly complex) scalar λ such that

$$\mathbf{A}\mathbf{w} = \lambda\mathbf{w}, \tag{12}$$
$$\mathbf{v}^*\mathbf{A} = \lambda\mathbf{v}^*, \tag{13}$$

where the asterisk denotes the complex-conjugate transpose. A left eigenvector $\mathbf{v}$ of $\mathbf{A}$, corresponding to λ, is a right eigenvector of $\mathbf{A}^*$ corresponding to $\bar{\lambda}$; that is,

$$\mathbf{A}^*\mathbf{v} = \bar{\lambda}\mathbf{v}. \tag{14}$$

The eigenvalues are found as the solutions of the characteristic equation (10).

If $\mathbf{A}$ is a $k \times k$ matrix, the characteristic equation is a polynomial of degree k and has k solutions λ_i, $i = 1, 2, \ldots, k$. The corresponding eigenvectors are $\mathbf{w}_i$ and $\mathbf{v}_i$, $i = 1, 2, \ldots, k$. I assume that these eigenvalues are all distinct, as seems to be true in practice for population-projection matrices. This assumption guarantees that

the right eigenvectors and left eigenvectors, respectively, are linearly independent sets.

Let $\langle \mathbf{w}, \mathbf{v} \rangle = \mathbf{v}^* \mathbf{w}$ denote the scalar product of $\mathbf{w}$ and $\mathbf{v}$. The left and right eigenvectors can always be scaled so that $\langle \mathbf{v}_i, \mathbf{w}_i \rangle = 1$. In addition, the left and right eigenvectors corresponding to different eigenvalues are orthogonal, so that $\langle \mathbf{v}_i, \mathbf{w}_j \rangle = 0$ if $i \neq j$.

Any matrix $\mathbf{A}$ with distinct eigenvalues can be written in the form

$$\mathbf{A} = \lambda_1 \mathbf{Z}_1 + \cdots + \lambda_k \mathbf{Z}_k \,,$$

where the matrices $\mathbf{Z}_i$, known as the *constituent matrices* of $\mathbf{A}$, are given by

$$\mathbf{Z}_i = \mathbf{w}_i \mathbf{v}_i^* \,. \tag{15}$$

That is, $\mathbf{Z}_i$ is a matrix whose columns are all proportional to $\mathbf{w}_i$ and whose rows are all proportional to $\mathbf{v}_i^*$.

The constituent matrices have two important properties. First,

$$\begin{aligned}
\mathbf{Z}_i^2 &= \mathbf{w}_i \mathbf{v}_i^* \mathbf{w}_i \mathbf{v}_i^* \\
&= \mathbf{w}_i \langle \mathbf{w}_i, \mathbf{v}_i \rangle \mathbf{v}_i^* \\
&= \mathbf{Z}_i.
\end{aligned} \tag{16}$$

(Such matrices are called idempotent.) Second, multiplying two different constituent matrices yields a zero matrix:

$$\begin{aligned}
\mathbf{Z}_i \mathbf{Z}_j &= \mathbf{w}_i \mathbf{v}_i^* \mathbf{w}_j \mathbf{v}_j^* \\
&= \mathbf{w}_i \langle \mathbf{w}_j, \mathbf{v}_i \rangle \mathbf{v}_j^* \\
&= \mathbf{0} \,.
\end{aligned} \tag{17}$$

These properties are useful because, together, they imply that

$$\mathbf{A}^2 = \left(\sum_i \lambda_i \mathbf{Z}_i \right) \left(\sum_j \lambda_j \mathbf{Z}_j \right) = \sum_i \lambda_i^2 \mathbf{Z}_i \,. \tag{18}$$

Multiplying repeatedly by $\mathbf{A}$, it is not hard to see that

$$\mathbf{A}^t = \sum_i \lambda^t \mathbf{Z}_i \,. \tag{19}$$

This result, together with equation (13), yields our desired expression for the dynamics of a population described by (8):

$$\mathbf{n}(t_0 + t) = \sum_i \lambda_i^t \mathbf{Z}_i \mathbf{n}(t_0) \,. \tag{20}$$

The only parts of the right-hand side of (20) that vary with time are the factors λ_i^t. The behavior of λ_i^t depends on the sign of λ_i and on whether λ_i is real or complex. If λ_i is real and positive, λ_i^t grows or decays exponentially, depending on whether λ_i is greater or less than one. If λ_i is real and negative, λ_i^t oscillates between positive and negative values, growing or decaying in magnitude depending on whether $|\lambda_i|$ is greater or less than one. If λ_i is complex, λ_i^t oscillates in a sinusoidal pattern, growing or decaying in magnitude depending on whether $|\lambda_i|$ is greater or less than one.

Since the dynamic properties of the population are determined by the eigenvalues of $\mathbf{A}$, it behooves us to see what we can say, a priori and in general, about these eigenvalues.

Eigenvalues, Eigenvectors, and the Perron-Frobenius Theorem

We can safely assume that the elements of $\mathbf{A}$ are nonnegative. Negative elements in $\mathbf{A}$ imply the possibility of negative individuals, which I prefer not to deal with. Perhaps surprisingly, this simple assumption tells us almost everything we want to know about the eigenvalues and eigenvectors of $\mathbf{A}$, thanks to a mathematical result known as the Perron-Frobenius theorem. In order to state the theorem, we need two more properties of $\mathbf{A}$: *irreducibility* and *primitivity*.

A matrix $\mathbf{A}$ is irreducible if and only if its life-cycle graph is connected, that is, if there is a path, following the direction of the arrows, from every stage to every other stage. A matrix $\mathbf{A}$ is primitive if and only if there is some integer k such that every element of $\mathbf{A}^k$ is strictly greater than zero. A more biologically revealing criterion is based on the life-cycle graph. Define a *loop* as a sequence of arrows, traversed in the direction of the arrows, that begins and ends at the same node, without passing through any node twice. The matrix $\mathbf{A}$ is primitive if and only if the greatest common divisor of the lengths of the loops in the life-cycle graph is one. Any primitive matrix is also irreducible. Most population-projection matrices encountered in practice are both irreducible and primitive.

What about matrices that are reducible or imprimitive (i.e., not primitive)? A reducible matrix has some stages that make no contribution to some other stages; the life-cycle graph breaks into two (or more) pieces with only one-way communication. The most common example is a life cycle with post-reproductive stages; from
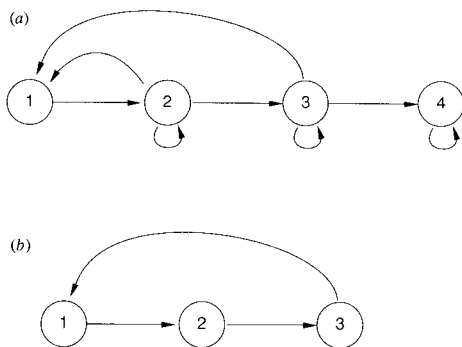
FIGURE 2. (a) A life-cycle graph corresponding to a reducible matrix. Stage 4 represents post-reproductive females; there is no pathway from this stage to any of the earlier stages. (b) A life-cycle graph corresponding to an imprimitive matrix. This is a semelparous age-classified model; individuals that survive to age class 3 reproduce and die.

such a stage there is no pathway back to the part of the life cycle that does reproduce. Figure 2a shows an example; a graph of this form appears in a stage-classified model for killer whales (Brault & Caswell 1993). An imprimitive life cycle has some underlying periodicity, so that the loops in the life-cycle graph are all multiples of some common loop length. Imprimitive matrices are sometimes called "cyclic" to reflect this fact. The most common example is a semelparous age-classified life cycle with a fixed age at reproduction (Fig. 2b). Only a single loop appears in such a life-cycle graph, with a length determined by the age at reproduction. Some kinds of seasonal models for annual organisms also produce imprimitive matrices, reflecting the periodicity imposed by the annual cycle of the seasons. The graphs in Figure 2 contain no coefficients because reducibility and primitivity depend on the form of the graph but not on the values of the coefficients.

The Perron-Frobenius theorem states that a nonnegative, irreducible, primitive matrix has three properties:

1. a simple (i.e., non-repeated) eigenvalue λ_1 that is real, positive, and strictly greater in magnitude than any of its other eigenvalues,

2. a right eigenvector $\mathbf{w}_1$ corresponding to λ_1, which is strictly positive (or can be made so by multiplying by a scalar) and is the only nonnegative right eigenvector, and

3. a left eigenvector $\mathbf{v}_1$ corresponding to λ_1, which is also strictly positive and is the only nonnegative left eigenvector.

The Perron-Frobenius theorem also describes the eigenvalues and eigenvectors of imprimitive and reducible matrices (see Caswell 1989*a* and, of course, many matrix texts, e.g., Seneta 1981; Horn & Johnson 1985).

Demographic Ergodicity

The eigenvalue λ_1 (the dominant eigenvalue) plays a central role in the asymptotic analysis of linear matrix models. In equation (20), the growth of $\mathbf{n}(t_0 + t)$ is given by a sum of terms involving the eigenvalues of $\mathbf{A}$ raised to higher and higher powers. Intuitively, as t gets large, λ_1^t increases more quickly, or decreases more slowly, than λ_i^t for $i \neq 1$. Asymptotically, we expect the growth of the population to be determined by λ_1, whereas all the eigenvalues contribute to short-term transient behavior. More precisely,

$$
\begin{aligned}
\lim_{t \to \infty} \frac{\mathbf{n}(t_0 + t)}{\lambda_1^t} &= \lim_{t \to \infty} \sum_i \left(\frac{\lambda_i}{\lambda_1} \right)^t \mathbf{Z}_i \mathbf{n}(t_0) \\
&= \mathbf{Z}_1 \mathbf{n}(t_0) \\
&= \mathbf{w}_1 \mathbf{v}_1^* \mathbf{n}(t_0) \,.
\end{aligned}
\tag{21}
$$

This gives the following results on asymptotic dynamics, conditional on the primitivity of $\mathbf{A}$.

1. The population eventually grows geometrically at a rate given by λ_1 (the *population growth rate* or *rate of increase*).

2. Population structure eventually becomes proportional to $\mathbf{w}_1$ (the *stable stage distribution*).

3. The constant of proportionality relating population structure and $\mathbf{w}$ is a weighted sum of the initial numbers in each stage ($\mathbf{v}_1^* \mathbf{n}(t_0)$). The weights are the elements in $\mathbf{v}_1$; the vector $\mathbf{v}_1$ thus gives the relative contributions of the stages to eventual population size (*not* population growth rate) and is called the *reproductive-value vector*.

The population eventually converges to the stable stage distribution, growing at a rate given by the dominant eigenvalue, regardless of the initial conditions (except, of course, the special case of a zero initial population). The property of forgetting the past and growing at a rate determined by the vital rates rather than by initial conditions is called *ergodicity*.

Because λ_1, $\mathbf{w}_1$, and $\mathbf{v}_1$ are properties of the vital rates rather than initial conditions, they are widely used as demographic statis-

tics. They can provide valuable insight into the vital rates and the environmental conditions that determine them, but they cannot predict actual population dynamics; everyone "knows" that populations cannot grow geometrically forever. These statistics must be interpreted as projections of what *would* happen if the vital rates were to remain constant, rather than as predictions of what *will* happen. They characterize the present environment, not the future of the population.

Similar ergodic results hold for stochastic models (see Chapter 3) and density-dependent models. In each case, the asymptotic properties provide demographic statistics that are determined by the vital rates rather than by the historical accidents of initial conditions. They can be used just as λ_1, $\mathbf{w}_1$, and $\mathbf{v}_1$ are used in the linear case.

3 Perturbation Analysis

Only rarely are we interested in one precisely specified model. We can usually imagine that the model might change in some way, and would like to know how such changes would affect the results of the analysis. Perturbation analyses address this problem. In density-independent models, perturbation analyses focus on the eigenvalues and eigenvectors, whereas in density-dependent models perturbation analyses focus on the stability and bifurcation of equilibria.

A perturbation analysis of the eigenvalues of a population-projection matrix can answer several questions.

1. What are the effects of potential changes in the vital rates, as might result from strategies designed to protect endangered species (by increasing λ) or control pest species (by reducing λ)?

2. Where should efforts to improve the estimates of the vital rates be focused in order to improve the accuracy of the estimate of λ? All else being equal, the biggest payoff comes from improving the estimates of the vital rates to which λ is most sensitive, since errors in those estimates have the biggest effect.

3. Genetic variation produces individuals whose vital rates are perturbed from the overall population values; from these, natural selection chooses those perturbations whose carriers increase most rapidly. Which vital rates are under the greatest selective pressure?

4. Suppose that some environmental differences (either natural or the result of experimental manipulation) have produced differences in the vital rates, and hence in λ, among two or more populations. How much do each of the vital-rate differences contribute to these observed differences in λ?

Fortunately, it is easy to calculate the sensitivity of λ to a change in any of the vital rates, once we know the eigenvectors. The next subsection presents these calculations. Formulas also exist for the sensitivities of the eigenvectors $\mathbf{w}$ and $\mathbf{v}$, for the sensitivities of λ for periodic time-varying models (Caswell & Trevisan 1994), and for the sensitivities of the sensitivities themselves (Caswell 1996b).

Sensitivity and Elasticity of Eigenvalues

The sensitivity of population growth rate to changes in the vital rates can be calculated as the derivative of λ to changes in the matrix elements a_{ij}. Suppose that λ, $\mathbf{w}$, and $\mathbf{v}$ satisfy

$$\mathbf{A}\mathbf{w} = \lambda\mathbf{w}, \tag{22}$$

$$\mathbf{v}^*\mathbf{A} = \lambda\mathbf{v}^*, \tag{23}$$

$$\langle\mathbf{w}, \mathbf{v}\rangle = \mathbf{v}^*\mathbf{w} = 1. \tag{24}$$

Now consider a perturbed matrix $\mathbf{A} + d\mathbf{A}$, where $d\mathbf{A}$ is a matrix of small perturbations da_{ij}. The eigenvalues and eigenvectors of the new matrix satisfy

$$(\mathbf{A} + d\mathbf{A})(\mathbf{w} + d\mathbf{w}) = (\lambda + d\lambda)(\mathbf{w} + d\mathbf{w}). \tag{25}$$

Expanding the products and eliminating second-order terms yields

$$\mathbf{A}\mathbf{w} + \mathbf{A}(d\mathbf{w}) + (d\mathbf{A})\mathbf{w} = \lambda\mathbf{w} + \lambda(d\mathbf{w}) + (d\lambda)\mathbf{w}, \tag{26}$$

which simplifies to

$$\mathbf{A}(d\mathbf{w}) + (d\mathbf{A})\mathbf{w} = \lambda(d\mathbf{w}) + (d\lambda)\mathbf{w}. \tag{27}$$

Multiplying both sides by $\mathbf{v}^*$ yields

$$\mathbf{v}^*\mathbf{A}(d\mathbf{w}) + \mathbf{v}^*(d\mathbf{A})\mathbf{w} = \lambda\mathbf{v}^*(d\mathbf{w}) + (d\lambda)\mathbf{v}^*\mathbf{w}. \tag{28}$$

The first term on the left-hand side is the same as the first term on the right-hand side (because of eq. 23), and the last term on the right-hand side simplifies to $d\lambda$ (because of eq. 24), leaving

$$\mathbf{v}^*d\mathbf{A}\mathbf{w} = d\lambda. \tag{29}$$

If $d\mathbf{A}$ contains only a single nonzero element da_{ij}, a change in only the ijth element of $\mathbf{A}$, we obtain the fundamental sensitivity

equation:

$$\frac{\partial \lambda}{\partial a_{ij}} = \bar{v}_i w_j. \tag{30}$$

(The bar over v_i has been ignored in most presentations of this formula, including mine. It is irrelevant to the case of the dominant eigenvalue of a population-projection matrix, which always has real eigenvectors, but it must be included for calculations involving other eigenvalues.)

Equation (30) says that the sensitivity of λ to changes in a_{ij} is proportional to the product of the reproductive value of stage i and the representation of stage j in the stable stage distribution.

The sensitivity of λ to changes in other parameters can be calculated using the chain rule: for some parameter x,

$$\frac{\partial \lambda}{\partial x} = \sum_{i,j} \frac{\partial \lambda}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial x}. \tag{31}$$

The sensitivity of λ gives the effect of a small additive change in one of the vital rates. The effect of a small *proportional* change in a vital rate is given by the elasticity of λ:

$$e_{ij} = \frac{a_{ij}}{\lambda} \frac{\partial \lambda}{\partial a_{ij}}. \tag{32}$$

In addition to giving the proportional change in λ resulting from a proportional change in the a_{ij}, the elasticities also measure the contribution of the a_{ij} to overall population growth rate. To be precise, $\sum_{i,j} e_{ij} = 1$ (for a simple proof, see Mesterton-Gibbons 1993), and e_{ij} can be interpreted as the proportion of λ contributed by a_{ij}.

Elasticities to other parameters can also be calculated:

$$\begin{aligned} e(x) &= \frac{x}{\lambda} \frac{\partial \lambda}{\partial x} \\ &= \frac{x}{\lambda} \sum_{i,j} \frac{\partial \lambda}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial x}. \end{aligned} \tag{33}$$

The elasticities of λ with respect to other parameters do not in general sum to one, and they cannot be interpreted as contributions to population growth rate.

Sensitivity or Elasticity?

Some authors seem to believe that sensitivities and elasticities are alternatives, that one is superior to the other, or that one or the other is biased in some way. This is not so; they provide accurate answers to different questions. The difference between them is comparable to the difference between plotting the same set of numbers on arithmetic (sensitivity) or logarithmic (elasticity) axes. Neither kind of graph is wrong, but one or the other may be better at revealing interesting patterns in the numbers. For more discussion, see Chapter 7, by Horvitz et al.

Life-Table-Response Experiments and Comparative Demography

Life-table-response experiments (LTRE's) are manipulative experiments or comparative observations in which the dependent variable is a complete set of vital rates (loosely speaking, a life table; Caswell 1989*b*). The different environmental conditions (the "treatments") cause changes in the vital rates, which in turn affect population dynamics. LTRE's are often summarized by using the rate of increase, λ, as a demographic statistic to integrate the treatment effects on survival and reproduction throughout the life cycle.

Knowing that a treatment produces a particular value of λ leaves unresolved the question of how the manifold changes in the vital rates contribute to the effect on growth rate. After all, some vital rates can be changed a great deal without affecting λ (e.g., the survival of a post-reproductive age class), whereas small changes in other vital rates produce large changes in λ. In addition, most environmental factors have differential effects on the different vital rates. A given treatment may affect survival, growth, and fertility differently, with different effects on those rates in different stages.

Treatment effects on λ can be decomposed into contributions from the effects on each of the vital rates (Caswell 1989*b*, 1996*a*). This decomposition makes it possible to pinpoint where in the life history the treatment has its greatest impact. The decomposition uses a first-order linear approximation to the effect on λ. I outline the simplest case here: a set of M treatments T_m, $m = 1, \ldots, M$, each of which produces its own matrix $\mathbf{A}_m$ and population growth rate $\lambda^{(m)}$. (I use parenthetical superscripts to denote treatments when subscripts distinguish matrix elements.)

Choose some condition as a reference treatment; this might be the mean of all the treatments, a control treatment, or any other condition of particular interest. The reference matrix is denoted $\mathbf{A}_r$, and treatment effects on λ are measured relative to $\lambda^{(r)}$.

To first order, we can write

$$\lambda^{(m)} - \lambda^{(r)} \approx \sum_{ij} \left(a_{ij}^{(m)} - a_{ij}^{(r)} \right) \frac{\partial \lambda}{\partial a_{ij}} \bigg|_{(\mathbf{A}_m + \mathbf{A}_r)/2} \qquad (34)$$

for $m = 1, \ldots, M$. Each term in the summation gives the contribution of the effect of treatment m on one of the vital rates to its overall effect on λ. That contribution is the product of the vital-rate effect and the sensitivity of λ to changes in that vital rate. If either of these terms is small—if the treatment doesn't effect a_{ij} or if λ is insensitive to a_{ij}—then the contribution of effects on a_{ij} to effects on λ is small. The converse is also true.

The sensitivities in (34) must be calculated from some particular matrix; here they are calculated from a matrix "halfway between" the two matrices ($\mathbf{A}_m$ and $\mathbf{A}_r$) being compared. There is some theoretical justification for this (Caswell 1989b), and it works well in practice.

The reason for using λ as a statistic to summarize the results of an LTRE is that it integrates the diverse and stage-specific effects of the treatments. The decomposition analysis complements this use; it pinpoints the source, within the life cycle, of the effects on λ. Experience with this kind of analysis shows that it is not safe to assume that the biggest changes in vital rates are responsible for the effects of a treatment on λ. Without some analysis like (34), half of the information contained in an LTRE is wasted.

Equation (34) describes a simple, one-way, fixed-effect experimental design. The approach has been extended to factorial designs, random designs, and regression designs (Caswell, in press). It can also be applied to statistics other than λ (as long as a perturbation theory is available for the statistic) and to parameters other than matrix elements (Caswell 1989c, 1996a, in press).

Prospective and Retrospective Analyses

The preceding subsections outline two ways of using perturbation analysis. Sensitivity and elasticity calculations are *prospective* analyses; they predict the results of perturbations of the vital rates before they happen. Indeed, they even show the results of perturbations that are biologically impossible. They tell nothing about

which vital rates are actually responsible for an observed change in λ. The LTRE decomposition analysis answers this kind of *retrospective* question. It does so by combining sensitivity analysis with information on the actual variance in the a_{ij}.

Don't confuse these two kinds of analysis, especially in ambiguous questions such as, "which of the vital rates is most important to population growth?" One way to answer this question is to find the rate with the biggest sensitivity (or elasticity); that rate is most important in the sense that if you were to change all the rates by the same amount (or same proportion), it would have the biggest impact. Another answer is based on an LTRE analysis; the most important vital rate is the one with the variation that makes the biggest contribution to the variability in λ. The two answers are usually different. Both are valid, but they answer different questions, the first prospective, the second retrospective. Chapter 7, by Horvitz et al., explores these issues further.