

Chapter 6

Age x Stage-Classified Models



6.1 Introduction

The first step in developing any kind of structured population model is choosing one or more variables in terms of which to describe the population structure. The job of these *i*-state variables is to encapsulate all the information about the past experience of an individual that is relevant to its future behavior (Metz and Diekmann 1986; Caswell 2001). Classical demography (for both humans and for non-human animals and plants) uses age as a *i*-state, but other, more biologically relevant criteria (e.g., size, developmental stage, parity, physiological condition, etc.) are now widely used in ecology, with age-classified models viewed as a special case.

However, it has long been recognized that cases exist where it is important to classify individuals by *both* age and stage.

1. Even in a stage-classified model, age still exists; every individual becomes older, by one unit of age, with the passage of each unit of time (e.g., Feichtinger 1971a; Caswell 2001, 2006, 2009; Tuljapurkar and Horvitz 2006; Horvitz and Tuljapurkar 2008). In these analyses, age dependence is implicit in the stage-classified model (see Chap. 5). Models that include both age and stage provide information that goes beyond this implicit dependence.
2. If the vital rates depend on both age and stage, only a model that includes both can reveal the joint action of age-and stage-specific processes (e.g., Goodman 1969). Such models, of course, require information on the joint age-dependence and stage-dependence of the vital rates, and thus are challenging to construct. A special case that has been extensively explored is the multi-regional case, in

This chapter is modified, under a Creative Commons Attribution License, from Caswell, H. 2012. Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theoretical Ecology* 5:403–417.

which the stage variable describes spatial location (e.g., Rogers 1966; Lebreton 1996). Models that combine age and some measure of health or disability status are an important part of health demography (e.g., Willekens 2014; Peeters et al. 2002; Wu et al. 2006; Zhou et al. 2016).

This chapter presents a model framework in which individuals are classified by age and stage, using the vec-permutation matrix approach (so-called for the role that the vec-permutation matrix plays in rearranging age and stage categories in the population vector). This formalism was introduced by Hunter and Caswell (2005) for populations classified by stage and location, was used in Chap. 5 to classify individuals by stage and environmental state; it has also been applied to stage and infection status (Klepac and Caswell 2011), stage and age (Caswell 2012; Caswell et al. 2018), and age and frailty (Caswell 2014). Megamatrix models (e.g., Pascarella and Horvitz 1998; Horvitz and Tuljapurkar 2008) can be written using this approach, as can block-structured multiregional models (e.g., Rogers 1975; Lebreton 1996). Matrix models can describe both population dynamics and cohort dynamics. Population dynamics (population growth, age and stage structure, reproductive value) depend on both the transitions of extant individuals and the production of new individuals by reproduction. In contrast, cohort dynamics (survivorship, life expectancy, age at death, generation time) depend only on the fates of already existing individuals. This chapter describes both kinds of analysis. For a more complete review and treatment, see Caswell et al. (2018).

6.2 Model Construction

The construction and analysis of these models requires a number of different matrices and operators (some of the notation is collected in Table 6.1). Individuals are classified into stages $1, \dots, s$ and age classes $1, \dots, \omega$. The model treats the processes of moving among stages and moving among age classes as alternating. First, stage-specific demography operates to move individuals among stages and to produce new offspring, with rates appropriate to their ages. Then aging acts to move individuals to the next older age, and the process repeats.

Define a stage-classified projection matrix \mathbf{A}_i , of dimension $s \times s$, for each age class, $i = 1, \dots, \omega$. Decompose \mathbf{A}_i into

$$\mathbf{A}_i = \mathbf{U}_i + \mathbf{F}_i \tag{6.1}$$

where \mathbf{U}_i contains the transition probabilities of extant individuals and \mathbf{F}_i describes the generation of new individuals by reproduction.

Table 6.1 Mathematical notation used in this chapter. Dimensions are shown, where relevant, for matrices and vectors; s denotes the number of stages and ω the number of age classes

Quantity	Description	Dimension
A_i, F_i, U_i	Stage-classified projection, fertility, and transition matrices for age class i	$s \times s$
D_U, D_F	Age transition matrices for individuals already present in the population and for new individuals produced by reproduction	$\omega \times \omega$
$\underline{A}, \underline{F}, \underline{U}, \underline{D}$	Block-diagonal matrices	$s\omega \times s\omega$
\tilde{A}, \tilde{U} , etc.	Age \times stage matrices constructed from block-diagonal matrices using the vec-permutation matrix	$s\omega \times s\omega$
$K_{s,\omega}, K$	Vec-permutation matrix	$s\omega \times s\omega$
I_s	Identity matrix	$s \times s$
$\mathbf{1}_s$	Vector of ones	$s \times 1$
e_i	The i th unit vector, with a 1 in the i th entry and zeros elsewhere	Various
E_{ij}	A matrix with a 1 in the (i, j) position, and zeros elsewhere	Various
\otimes	Kronecker product	
\circ	Hadamard, or element-by-element, product	
vec \mathbf{X}	The vec operator, which stacks the columns of a $m \times n$ matrix \mathbf{X} into a $mn \times 1$ vector	
$\mathcal{D}(\mathbf{x})$	A diagonal matrix with \mathbf{x} on the diagonal and zeros elsewhere	

Aging is described by two matrices, each of dimension $\omega \times \omega$ (shown here for 3×3 , but easily generalized),

$$D_U = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{dimension } \omega \times \omega \tag{6.2}$$

$$D_F = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \omega \times \omega \tag{6.3}$$

The matrix D_U applies to extant individuals; such an individual advances to the next age class. I have set the (ω, ω) entry of D_U to 1, so that the last age class contains individuals of age ω and older. If this entry were set to 0, all individuals in the last age class would die. The matrix D_F applies to individuals newly created by reproduction; such newborn individuals are placed in the first age class, regardless of the age of their parents.

Using the matrices \mathbf{A}_i , \mathbf{U}_i , \mathbf{F}_i , \mathbf{D}_U , and \mathbf{D}_F , construct block-diagonal matrices, each of dimension $s\omega \times s\omega$. For example,

$$\mathbb{A} = \begin{pmatrix} \mathbf{A}_1 & & \\ & \ddots & \\ & & \mathbf{A}_\omega \end{pmatrix} \quad (6.4)$$

with similar structures for \mathbb{U} , \mathbb{F} , \mathbb{D}_U , and \mathbb{D}_F . These block-diagonal matrices can be written

$$\mathbb{A} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{A}_i) \quad (6.5)$$

$$\mathbb{U} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{U}_i) \quad (6.6)$$

$$\mathbb{F} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{F}_i) \quad (6.7)$$

$$\mathbb{D}_U = \mathbf{I}_s \otimes \mathbf{D}_U \quad (6.8)$$

$$\mathbb{D}_F = \mathbf{I}_s \otimes \mathbf{D}_F \quad (6.9)$$

where \mathbf{E}_{ii} is of dimension $\omega \times \omega$.

If the demography is strictly stage-dependent, so that $\mathbf{A}_i = \mathbf{A}$, for $i = 1, \dots, \omega$, then the block-diagonal matrices \mathbb{A} , \mathbb{F} , and \mathbb{U} reduce to, e.g.,

$$\mathbb{A} = \mathbf{I}_\omega \otimes \mathbf{A} \quad (6.10)$$

with corresponding expressions for \mathbb{F} and \mathbb{U} .

The state of the population at time t could be described by a 2-dimensional array

$$\mathcal{N}(t) = \begin{pmatrix} n_{11} & \cdots & n_{1\omega} \\ \vdots & & \vdots \\ n_{s1} & \cdots & n_{s\omega} \end{pmatrix} (t) \quad s \times \omega \quad (6.11)$$

where rows correspond to stages and columns to age classes. However, such a 2-dimensional array cannot be projected directly; instead, it is transformed to a vector,

$$\mathbf{n}(t) = \text{vec } \mathcal{N}(t) = \begin{pmatrix} n_{11} \\ \vdots \\ \hline n_{s1} \\ \vdots \\ \hline n_{1\omega} \\ \vdots \\ n_{s\omega} \end{pmatrix} (t) \quad s\omega \times 1 \quad (6.12)$$

using the vec operator, which stacks the columns of the matrix one above the next. The vector $\mathbf{n}(t)$ created in this way contains the stages arranged within age classes. An alternative configuration, with ages arranged within stages, is obtained by applying the vec operator to \mathcal{N}^T :

$$\text{vec } \mathcal{N}^T(t) = \begin{pmatrix} n_{11} \\ \vdots \\ n_{1\omega} \\ \vdots \\ n_{s1} \\ \vdots \\ n_{s\omega} \end{pmatrix} (t) \quad s\omega \times 1 \quad (6.13)$$

The two vectors $\text{vec } \mathcal{N}$ and $\text{vec } \mathcal{N}^T$ are related by the vec -permutation matrix, or commutation matrix, \mathbf{K} , (Henderson and Searle 1981),

$$\text{vec } \mathcal{N}^T = \mathbf{K}_{s,\omega} \text{vec } \mathcal{N} \quad (6.14)$$

(see Sect. 2.2.3). Where no confusion seems likely to arise, we will suppress the subscripts and write $\mathbf{K}_{s,\omega}$ as \mathbf{K} . As with any permutation matrix, $\mathbf{K}^T = \mathbf{K}^{-1}$.

The goal of the model is to project the age-stage vector $\mathbf{n} = \text{vec } \mathcal{N}$ from t to $t + 1$. The complete projection is given by

$$\mathbf{n}(t + 1) = \left(\mathbf{K}^T \mathbb{D}_U \mathbf{K} \mathbf{U} + \mathbf{K}^T \mathbb{D}_F \mathbf{K} \mathbf{F} \right) \mathbf{n}(t) \quad (6.15)$$

This deserves some explanation. Consider the first term on the right hand side, $\mathbf{K}^T \mathbb{D}_U \mathbf{K} \mathbf{U}$. Reading from right to left, it first operates on the vector $\mathbf{n}(t)$ with the block diagonal matrix \mathbf{U} , which moves surviving extant individuals among stages without changing their age. Then the resulting vector is rearranged by the vec -permutation matrix \mathbf{K} to group individuals by age classes within each stage. The block diagonal matrix \mathbb{D}_U then moves each surviving individual to the next older age class. Finally, \mathbf{K}^T rearranges the vector back to the stage-within-age arrangement of $\mathbf{n}(t)$.

The second term in (6.15), $\mathbf{K}^T \mathbb{D}_F \mathbf{K} \mathbf{F}$, carries out a similar sequence of transformations for the generation of new individuals. First, newborn individuals are produced according to the block-diagonal fertility matrix \mathbf{F} . The resulting vector is rearranged by the vec -permutation matrix, and then the matrix \mathbb{D}_F places all the newborn individuals into the first age class. Finally, \mathbf{K}^T rearranges the vector to the stage-within-age arrangement.

I will write the age \times stage projection matrix in (6.15) as

$$\tilde{\mathbf{A}} = \left(\mathbf{K}^T \mathbb{D}_U \mathbf{K} \mathbf{U} + \mathbf{K}^T \mathbb{D}_F \mathbf{K} \mathbf{F} \right) \quad (6.16)$$

$$= \left(\tilde{\mathbf{U}} + \tilde{\mathbf{F}} \right) \quad (6.17)$$

The matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{U}}$, and $\tilde{\mathbf{F}}$ that operate on the age-stage vector \mathbf{n} are denoted with a tilde ($\tilde{\mathbf{A}}$, $\tilde{\mathbf{U}}$, $\tilde{\mathbf{F}}$); these matrices define the age \times stage-classified model and can be subjected to all the usual demographic analyses.

6.3 Sensitivity Analysis

Age-stage models pose particular challenges for perturbation analysis, because interest naturally focuses on changes in the matrices \mathbf{F}_i and \mathbf{U}_i ($i = 1, \dots, \omega$), which are deeply embedded within $\tilde{\mathbf{F}}$, $\tilde{\mathbf{U}}$, and $\tilde{\mathbf{A}}$.

Consider a generic dependent variable ξ , which is a scalar- or vector-valued function of $\tilde{\mathbf{A}}$. In the examples to follow, ξ will be either the population growth rate λ or the joint distribution of age and stage at death in a cohort, but it could be any variable calculated from $\tilde{\mathbf{A}}$. Let θ be a vector of parameters; these could be entries of the matrices, or lower-level parameters determining those entries. The goal of perturbation analysis is to obtain the derivative of ξ with respect to θ ,

$$\frac{d\xi}{d\theta^T} = \frac{d\xi}{d\text{vec}^T \tilde{\mathbf{A}}} \frac{d\text{vec} \tilde{\mathbf{A}}}{d\theta^T}. \quad (6.18)$$

The first term in (6.18) is the derivative of ξ with respect to the matrix $\tilde{\mathbf{A}}$. If, for example, ξ was the dominant eigenvalue λ , then this term would be the matrix calculus version of the well-known eigenvalue sensitivity equation.

The second term in (6.18) requires differentiating $\tilde{\mathbf{A}}$ with respect to the parameters that determine it. From (6.16), write

$$\tilde{\mathbf{A}} = \mathbf{Q}_U \mathbf{U} + \mathbf{Q}_F \mathbf{F} \quad (6.19)$$

where $\mathbf{Q}_U = \mathbf{K}^T \mathbb{D}_U \mathbf{K}$ and $\mathbf{Q}_F = \mathbf{K}^T \mathbb{D}_F \mathbf{K}$ are the (constant) matrix products appearing in the definition of $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{F}}$ in (6.16).

Differentiating $\tilde{\mathbf{A}}$ in (6.19) gives

$$d\text{vec} \tilde{\mathbf{A}} = (\mathbf{I}_{sw} \otimes \mathbf{Q}_U) d\text{vec} \mathbf{U} + (\mathbf{I}_{sw} \otimes \mathbf{Q}_F) d\text{vec} \mathbf{F} \quad (6.20)$$

This requires the differentials of \mathbb{U} and \mathbb{F} . Differentiating \mathbb{U} in (6.6) gives

$$d\mathbb{U} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes d\mathbf{U}_i) \quad (6.21)$$

Applying the vec operator to $d\mathbb{U}$ gives

$$d\text{vec } \mathbb{U} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec } \mathbf{U}_i \quad (6.22)$$

using the results of Magnus and Neudecker (1985, Theorem 11); see also Klepac and Caswell (2011, Appendix B) on the derivative of the Kronecker product. Differentiation of \mathbb{F} proceeds in the same fashion, yielding

$$d\text{vec } \mathbb{F} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec } \mathbf{F}_i \quad (6.23)$$

In the special case where \mathbb{U} and \mathbb{F} are constructed from single stage-classified matrices \mathbf{U} and \mathbf{F} , as in (6.10), Eqs. (6.22) and (6.23) simplify even further to

$$d\text{vec } \mathbb{U} = (\mathbf{I}_{\omega} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec } \mathbf{U} \quad (6.24)$$

$$d\text{vec } \mathbb{F} = (\mathbf{I}_{\omega} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec } \mathbf{F} \quad (6.25)$$

Substituting (6.22) and (6.23) into (6.20) and then substituting (6.20) into (6.18) yields the general result for the derivative

$$\begin{aligned} \frac{d\xi}{d\boldsymbol{\theta}^T} &= \frac{d\xi}{d\text{vec }^T \tilde{\mathbf{A}}} \left[(\mathbf{I}_{s\omega} \otimes \mathbf{Q}_U) \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) \frac{d\text{vec } \mathbf{U}_i}{d\boldsymbol{\theta}^T} \right] \\ &+ \frac{d\xi}{d\text{vec }^T \tilde{\mathbf{A}}} \left[(\mathbf{I}_{s\omega} \otimes \mathbf{Q}_F) \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) \frac{d\text{vec } \mathbf{F}_i}{d\boldsymbol{\theta}^T} \right] \end{aligned} \quad (6.26)$$

Notice that (6.26) requires only three pieces of demographic information: the derivatives of \mathbf{U}_i and \mathbf{F}_i with respect to the parameters (whatever those may be in the case at hand) and the sensitivity of the dependent variable ξ (whatever that may be) to the elements of the matrix $\tilde{\mathbf{A}}$ from which it is calculated. All the other pieces of (6.26) are constants. Some of these constant matrices may be large, depending on s and ω , but they are very sparse; the sparse matrix technology available in MATLAB can be extremely useful in implementation. An alternative formulation of the differentials of the block matrices \mathbb{U} and \mathbb{F} is given in Caswell and van Daalen (2016).

6.4 Examples

Here we consider two examples of the sensitivity analysis of age-stage model to extract age-classified information from a stage-classified model. The first example will derive the sensitivity of the population growth rate λ , obtaining the sensitivity of λ to both age- and stage-specific survival, permitting examination of how selection pressures on senescence-inducing traits would vary from stage to stage. The second example is an analysis of the joint distribution of age and stage at death.

These examples are based on a stage-classified model (Parker 2000) for Scotch broom (*Cytisus scoparius*). Scotch broom is a large (up to 4 m tall) leguminous shrub, introduced into North America from Europe in the late nineteenth century. It is an invasive plant, considered a pest in the northwestern parts of North America. Stage-classified demographic models have been used to evaluate potential management policies for the plant (Parker 2000) and to investigate its potential for spatial spread (Neubert and Parker 2004).

The model contains seven stages (stage 1 = seeds, 2 = seedlings, 3 = juveniles, 4 = small adults, 5 = medium adults, 6 = large adults, 7 = extra-large adults), and parameters were estimated at a number of locations in Washington State. As is typical with many perennial plant species, survival is low for seeds and seedlings, but increases dramatically in larger stages. Parker's study presented estimated projection matrices for plants at the edge, at intermediate locations, and at the center of an invading stand. Plants near the center experience more crowding, with resulting reduced rates of survival, growth, and fertility.

6.4.1 Population Growth Rate and Selection Gradients

The population growth rate λ , the stable age or stage distribution \mathbf{w} , and age or stage-specific reproductive value vector \mathbf{v} are given by the dominant eigenvalue and corresponding right and left eigenvectors of the population projection matrix, respectively. In evolutionary demography, λ measures the fitness of a phenotype, in that it gives the eventual rate at which descendants of an individual with that phenotype will increase. The selection gradient on a vector of traits $\boldsymbol{\theta}$ is given by

$$\frac{d\lambda}{d\boldsymbol{\theta}^T} \quad (6.27)$$

These gradients play a fundamental role in evolutionary biodemography, whether evolution is conceived of in terms of population genetics, quantitative genetics, adaptive dynamics, or mutation accumulation (e.g., Metz et al. 1992; Dercole and Rinaldi 2008; Rice 2004; Barfield et al. 2011). If the gradient is positive, selection favors an increase in the trait, and vice-versa.

In this application, ξ in (6.18) is the dominant eigenvalue λ . Let \mathbf{w} and \mathbf{v} be the right and left eigenvectors corresponding to λ , scaled so that $\mathbf{v}^\top \mathbf{w} = 1$. Then, in (6.26),

$$\frac{d\lambda}{d\text{vec } \tilde{\mathbf{A}}} = \mathbf{w}^\top \otimes \mathbf{v}^\top. \quad (6.28)$$

See Chap. 3 and Caswell (2010).

In this model, the vital rates are functions only of stage; the phenotype is blind to the age of the individual. However, the terms in the summations in (6.26) give the selection gradients on traits that would modify the phenotype at each age. That is,

$$\begin{aligned} \frac{d\lambda}{d\theta^\top} \Big|_{\text{age}=i} &= \frac{d\lambda}{d\text{vec } \tilde{\mathbf{A}}} \left[(\mathbf{I}_{sw} \otimes \mathbf{Q}_U) (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_\omega \otimes \mathbf{I}_{s^2}) \frac{d\text{vec } \mathbf{U}_i}{d\theta^\top} \right] \\ &+ \frac{d\lambda}{d\text{vec } \tilde{\mathbf{A}}} \left[(\mathbf{I}_{sw} \otimes \mathbf{Q}_F) (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec } \mathbf{I}_\omega \otimes \mathbf{I}_{s^2}) \frac{d\text{vec } \mathbf{F}_i}{d\theta^\top} \right] \end{aligned} \quad (6.29)$$

Thus, these terms reveal the selection patterns that would operate on a mutation that was able to detect the age of an individual within a given stage, or that affected age differentially depending on the stage of the individual.

To examine the selection gradients on survival, it is necessary to separate survival from inter-stage transitions in \mathbf{U} . Let σ be the vector of stage-specific survival probabilities. The matrix \mathbf{U} can be written as the product of a matrix $\Sigma = \mathbf{1}_s \sigma^\top$ containing the survival probabilities on the diagonal and a matrix \mathbf{G} of transition probabilities, conditional on survival;

$$\mathbf{U} = \mathbf{G}\Sigma. \quad (6.30)$$

(cf. Chap. 8). If \mathbf{F} is independent¹ of σ , then

$$d\mathbf{U} = \mathbf{G} d\Sigma. \quad (6.31)$$

Applying the vec operator gives

$$\begin{aligned} d\text{vec } \mathbf{U} &= (\mathbf{I}_s \otimes \mathbf{G}) \text{vec } \mathcal{D} (\mathbf{1}_s d\sigma^\top) \\ &= (\mathbf{I}_s \otimes \mathbf{G}) \mathcal{D} (\text{vec } \mathbf{I}_s) (\mathbf{I}_s \otimes \mathbf{1}_s) d\sigma \end{aligned} \quad (6.32)$$

¹By assuming that \mathbf{F} does not depend on σ , I am in effect choosing a pre-breeding census and excluding neonatal mortality from σ .

which implies that

$$\frac{d\text{vec } \mathbf{U}}{d\sigma^\top} = (\mathbf{I}_s \otimes \mathbf{G}) \mathcal{D}(\text{vec } \mathbf{I}_s) (\mathbf{I}_s \otimes \mathbf{1}_s) \quad (6.33)$$

Setting $\theta = \sigma$ and substituting (6.33) and (6.28) into (6.18) gives the selection gradient on σ . Substituting (6.33) and (6.28) into (6.29), with $d\text{vec } \mathbf{F}/d\theta^\top = \mathbf{0}$, gives the selection gradient on σ as a function of age and stage.

Results The projection matrix \mathbf{A} for Scotch broom² is

$$\mathbf{A} = \begin{pmatrix} 0.740 & 0 & 3.400 & 47.1 & 108.700 & 1120.0 & 3339.0 \\ 0.001 & 0.310 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0.310 & 0 & 0 & 0 & 0 \\ 0 & 0.038 & 0.290 & 0.024 & 0 & 0 & 0 \\ 0 & 0 & 0.069 & 0.390 & 0.320 & 0 & 0.091 \\ 0 & 0 & 0 & 0.440 & 0.440 & 0.530 & 0.091 \\ 0 & 0 & 0 & 0 & 0.029 & 0.400 & 0.730 \end{pmatrix} \quad (6.34)$$

The matrix \mathbf{U} is obtained from \mathbf{A} by setting all elements in the first row, except for a_{11} , to zero. The matrix \mathbf{F} is a 7×7 matrix with the elements of row 1, columns 2–7 of \mathbf{A} in the corresponding positions, and zeros elsewhere. The maximum age was set to $\omega = 30$. The aging matrices \mathbf{D}_U and \mathbf{D}_F are given by (6.2) and (6.3) with $\omega = 30$. Because the vital rates do not depend on age, the dominant eigenvalues of \mathbf{A} and $\hat{\mathbf{A}}$ should be identical, and they are; $\lambda = 1.268$.

The selection gradients on stage-specific survival (i.e., sensitivities of λ to σ) are shown in Fig. 6.1. There is a steady decline with increasing stage, from seeds to medium-sized adults, but then an increase for large and extra-large adults. A quite different pattern emerges when the selection gradients are calculated as functions of both age and stage, using (6.29). These results are shown in Fig. 6.2. The age-specific selection gradients on survival in stages 1–3 are strictly decreasing with age. But the age-specific selection gradients on survival in the adult stages 4–7 increase with age, level off, and then decline. The increase is longer and more pronounced in the larger adult stages.

It is now known that this pattern is widespread in plant populations. It appears in all eight of the Scotch broom populations studied by Parker (2000), and in almost all of 36 species of plants examined by Caswell and Salguero-Gómez (2013). It has important implications for the evolution of senescence. Hamilton (1966) showed that the selection gradient on age-specific mortality is always decreases with age, and argued that this implied that selection would always lead to senescence. Incorporating stage-dependence as well as age-dependence of the vital

²This is the matrix for the Discovery Park population, 1993–1994, edge conditions; taken from the Appendix of Parker (2000).

Fig. 6.1 Sensitivity of population growth rate λ to stage-specific survival probabilities. Calculated for the stage-classified model of Scotch broom (*Cytisus scoparius*) using data from Parker (2000). Stages: 1 = seeds, 2 = seedlings, 3 = juveniles, 4 = small adults, 5 = medium adults, 6 = large adults, 7 = extra-large adults

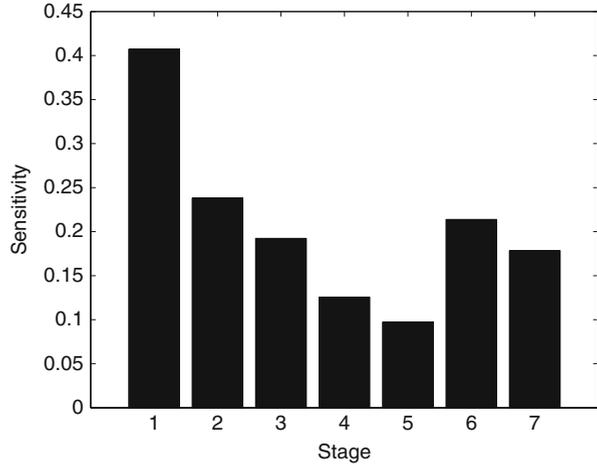
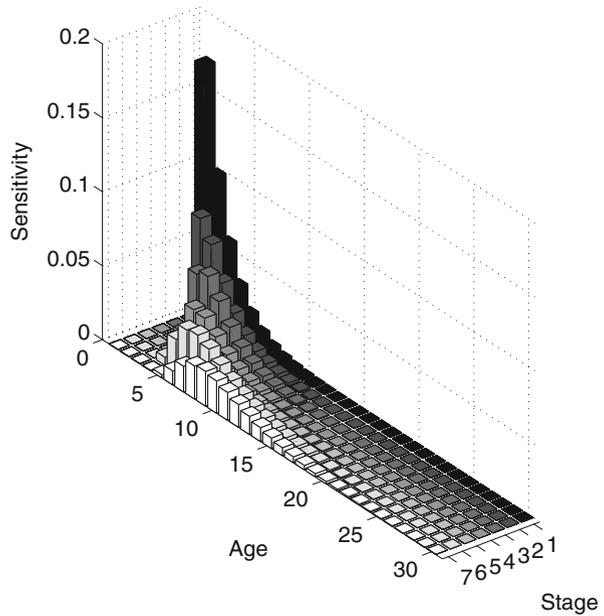


Fig. 6.2 Sensitivity of population growth rate λ to stage-specific survival as a function of age, for Scotch broom. Stages defined as in Fig. 6.1



rates means that, over some range of ages, the selection gradient increases (*contra-senescent* selection in the terminology of Caswell and Salguero-Gómez 2013). Thus conclusions that follow from the general decline in selection gradients with age may not apply to traits that affect age-specific survival differentially depending on developmental stage. Traits that affect survival in adult stages should postpone senescence for at least some time.

6.4.2 Distributions of Age and Stage at Death

The pattern of longevity within a population is captured by the probability distribution of the age at death, one of the standard results of age-classified life table analysis. The moments of the age at death and their sensitivity can also be calculated directly from stage-classified models using Markov chain methods (Feichtinger 1971b; Caswell 2001, 2006, 2009; Tuljapurkar and Horvitz 2006; Horvitz and Tuljapurkar 2008); see Chaps. 4 and 5. Here we can go beyond that and get the full joint distribution of stage and age at death, along with the marginal distributions of age at death and stage at death, implied by an age \times stage classified model.

To do this, note that the cohort projection matrix $\tilde{\mathbf{U}}$ describes movement of individuals among transient states of an absorbing Markov chain, where the absorbing state is death, or death classified by stage or age at death. The transition matrix of the chain is

$$\tilde{\mathbf{P}} = \left(\begin{array}{c|c} \tilde{\mathbf{U}} & \mathbf{0} \\ \hline \tilde{\mathbf{M}} & \mathbf{I} \end{array} \right) \quad (6.35)$$

By properly structuring \mathbf{M} , the model can give information about the age, stage, or the joint distribution of age and stage at death.³ Each row of $\tilde{\mathbf{M}}$ corresponds to an absorbing state, and \tilde{m}_{ij} is the probability of a transition from transient state j to absorbing state i . To compute the distribution of age and stage at death, we define the absorbing states to correspond to the age \times stage combination at death. Thus $\tilde{\mathbf{M}}$ contains probabilities of death on the diagonal and zeros elsewhere,

$$\tilde{\mathbf{M}} = \mathbf{I}_{s\omega} - \mathcal{D} \left(\mathbf{1}_{s\omega}^T \tilde{\mathbf{U}} \right). \quad (6.36)$$

The fundamental matrix of the Markov chain in (6.35) is

$$\tilde{\mathbf{N}} = \left(\mathbf{I} - \tilde{\mathbf{U}} \right)^{-1} \quad (6.37)$$

The (i, j) element of $\tilde{\mathbf{N}}$ is the expected number of visits that an individual in state j will make to transient state i before death.

Consider the eventual fate of an individual starting in transient state j . Let

$$\tilde{b}_{ij} = P \left[\text{eventual absorption in } i \mid \text{starting in } j \right] \quad (6.38)$$

³This also leads to a powerful approach, including sensitivity analysis, for cause of death calculations (Caswell and Ouellette 2016, 2018).

The \tilde{b}_{ij} are the elements of the matrix $\tilde{\mathbf{B}}$ ($s\omega \times s\omega$) given by

$$\tilde{\mathbf{B}} = \tilde{\mathbf{M}}\tilde{\mathbf{N}} \quad (6.39)$$

(Iosifescu 1980, Theorem 3.3; see also Caswell 2001, Section 5.1). Since the absorbing states (the rows of $\tilde{\mathbf{M}}$) correspond to combinations of age and stage at death, column j of $\tilde{\mathbf{B}}$ gives the joint distribution of age and stage at death, starting from state (i.e., age \times stage combination) j :

$$\tilde{\mathbf{B}}(:, j) = \tilde{\mathbf{B}}\mathbf{e}_j \quad (6.40)$$

using MATLAB notation in which $\mathbf{X}(:, j)$ is column j of \mathbf{X} , and where \mathbf{e}_j is a vector of length $s\omega$ with a 1 in the j th entry and zeros elsewhere. The rows of $\tilde{\mathbf{B}}$ correspond to combinations of stage and age at death. Summing the rows over stages gives the marginal distribution of *age* at death, starting in column j of $\tilde{\mathbf{B}}$, as

$$\mathbf{g}_j = \left(\mathbf{I}_\omega \otimes \mathbf{1}_s^\top \right) \tilde{\mathbf{B}}(:, j) \quad \text{marginal age distribution} \quad \omega \times 1 \quad (6.41)$$

Similarly, summing over ages gives the marginal distribution of *stage* at death:

$$\mathbf{h}_j = \left(\mathbf{1}_\omega^\top \otimes \mathbf{I}_s \right) \tilde{\mathbf{B}}(:, j) \quad \text{marginal stage distribution} \quad s \times 1 \quad (6.42)$$

6.4.2.1 Perturbation Analysis

In the general sensitivity equation (6.18), the dependent variable $\xi = \tilde{\mathbf{B}}(:, j)$. This depends only on $\tilde{\mathbf{U}}$, so the first term in (6.18) can be shown to be

$$\begin{aligned} \frac{d\xi}{d\text{vec } \tilde{\mathbf{A}}} &= \frac{d\tilde{\mathbf{B}}(:, j)}{d\text{vec } \tilde{\mathbf{U}}} \quad (6.43) \\ &= - \left(\mathbf{e}_j^\top \tilde{\mathbf{N}}^\top \otimes \mathbf{I}_{s\omega} \right) \mathcal{D} \left(\text{vec } \mathbf{I}_{s\omega} \right) \left(\mathbf{I}_{s\omega} \otimes \mathbf{I}_{s\omega} \mathbf{1}_{s\omega}^\top \right) + \left(\mathbf{e}_j^\top \tilde{\mathbf{N}}^\top \otimes \tilde{\mathbf{B}} \right) \end{aligned} \quad (6.44)$$

The desired derivative $d\tilde{\mathbf{B}}(:, j)/d\boldsymbol{\theta}^\top$ is obtained by substituting (6.44) for $d\xi/d\text{vec } \tilde{\mathbf{A}}$ in (6.26), setting $d\text{vec } \mathbf{F}_i/d\boldsymbol{\theta}^\top = 0$.

The sensitivities of the marginal distributions of age and stage at death are then given by

$$\frac{d\mathbf{g}_j}{d\boldsymbol{\theta}^\top} = \left(\mathbf{I}_\omega \otimes \mathbf{1}_s^\top \right) \frac{d\tilde{\mathbf{B}}(:, j)}{d\boldsymbol{\theta}^\top} \quad (6.45)$$

$$\frac{d\mathbf{h}_j}{d\boldsymbol{\theta}^\top} = \left(\mathbf{1}_\omega^\top \otimes \mathbf{I}_s \right) \frac{d\tilde{\mathbf{B}}(:, j)}{d\boldsymbol{\theta}^\top} \quad (6.46)$$

Derivation To derive the sensitivity of the joint distribution of age and stage at death, conditional on some starting age \times stage combination, we start by differentiating equation (6.40) for column j of $\tilde{\mathbf{B}}$ and applying the vec operator,

$$d\tilde{\mathbf{B}}(:, j) = \left(\mathbf{e}_j^T \otimes \mathbf{I}_{s\omega} \right) d\text{vec } \tilde{\mathbf{B}}. \quad (6.47)$$

However, from (6.39), $\tilde{\mathbf{B}} = \tilde{\mathbf{M}}\tilde{\mathbf{N}}$, so

$$d\tilde{\mathbf{B}} = \left(d\tilde{\mathbf{M}} \right) \tilde{\mathbf{N}} + \tilde{\mathbf{M}} \left(d\tilde{\mathbf{N}} \right). \quad (6.48)$$

and

$$d\text{vec } \tilde{\mathbf{B}} = \left(\tilde{\mathbf{N}}^T \otimes \mathbf{I}_{s\omega} \right) d\text{vec } \tilde{\mathbf{M}} + \left(\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{M}} \right) d\text{vec } \tilde{\mathbf{N}}. \quad (6.49)$$

The differential of the fundamental matrix $\tilde{\mathbf{N}}$ is

$$d\text{vec } \tilde{\mathbf{N}} = \left(\tilde{\mathbf{N}}^T \otimes \tilde{\mathbf{N}} \right) d\text{vec } \tilde{\mathbf{U}} \quad (6.50)$$

(Caswell 2006; see Chap. 5). The differential of $\tilde{\mathbf{M}}$ is obtained by rewriting (6.36) as

$$\tilde{\mathbf{M}} = \mathbf{I}_{s\omega} - \mathbf{I}_{s\omega} \circ \left(\mathbf{1}_{s\omega} \mathbf{1}_{s\omega}^T \tilde{\mathbf{U}} \right), \quad (6.51)$$

differentiating,

$$d\tilde{\mathbf{M}} = -\mathbf{I}_{s\omega} \circ \left[\mathbf{1}_{s\omega} \mathbf{1}_{s\omega}^T \left(d\tilde{\mathbf{U}} \right) \right], \quad (6.52)$$

and applying the vec operator to obtain

$$d\text{vec } \tilde{\mathbf{M}} = -\mathcal{D} \left(\text{vec } \mathbf{I}_{s\omega} \right) \left(\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega} \mathbf{1}_{s\omega}^T \right) d\text{vec } \tilde{\mathbf{U}} \quad (6.53)$$

Substituting (6.50) and (6.53) into (6.49) gives

$$\begin{aligned} d\text{vec } \tilde{\mathbf{B}} = & \left[- \left(\tilde{\mathbf{N}}^T \otimes \mathbf{I}_{s\omega} \right) \mathcal{D} \left(\text{vec } \mathbf{I}_{s\omega} \right) \left(\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega} \mathbf{1}_{s\omega}^T \right) \right. \\ & \left. + \left(\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{M}} \right) \left(\tilde{\mathbf{N}}^T \otimes \tilde{\mathbf{N}} \right) \right] d\text{vec } \tilde{\mathbf{U}} \end{aligned} \quad (6.54)$$

Substituting this into (6.47) gives

$$\begin{aligned} d\tilde{\mathbf{B}}(:, j) = & \left[- \left(\mathbf{e}_j^T \otimes \mathbf{I}_{s\omega} \right) \left(\tilde{\mathbf{N}}^T \otimes \mathbf{I}_{s\omega} \right) \mathcal{D} \left(\text{vec } \mathbf{I}_{s\omega} \right) \left(\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega} \mathbf{1}_{s\omega}^T \right) \right. \\ & \left. + \left(\mathbf{e}_j^T \otimes \mathbf{I}_{s\omega} \right) \left(\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{M}} \right) \left(\tilde{\mathbf{N}}^T \otimes \tilde{\mathbf{N}} \right) \right] d\text{vec } \tilde{\mathbf{U}} \end{aligned} \quad (6.55)$$

Equation (6.55) can be simplified to obtain (6.44), using the fact that

$$(A \otimes B)(C \otimes D) = (AC \otimes BD),$$

provided the products exist.

Results Figure 6.3 shows the joint distribution of age and stage at death for a seed of age 1 (one definition of “newborn” in this life cycle), with $\omega = 40$. Almost all seeds will die as seeds, because the germination probability is low, $a_{21} = 0.001$; see (6.34). The fates of seedlings (another possible choice for newborn status) are more diverse, and those of juveniles and small adults even moreso; the distributions show what proportion will die as seedlings, juveniles, etc., and at what ages (Fig. 6.3).

The marginal distribution of age at death, for individuals in each initial stage, is given in Fig. 6.4. Not surprisingly, larger stages have an age distribution of death shifted to later ages, including some probability of survival to age class $\omega (\geq 40$ years in this calculation).

The sensitivity of g_2 (the marginal distribution of age at death for a seedling) is shown in Fig. 6.5. Changes in the survival of seeds (σ_1) have no effect on this

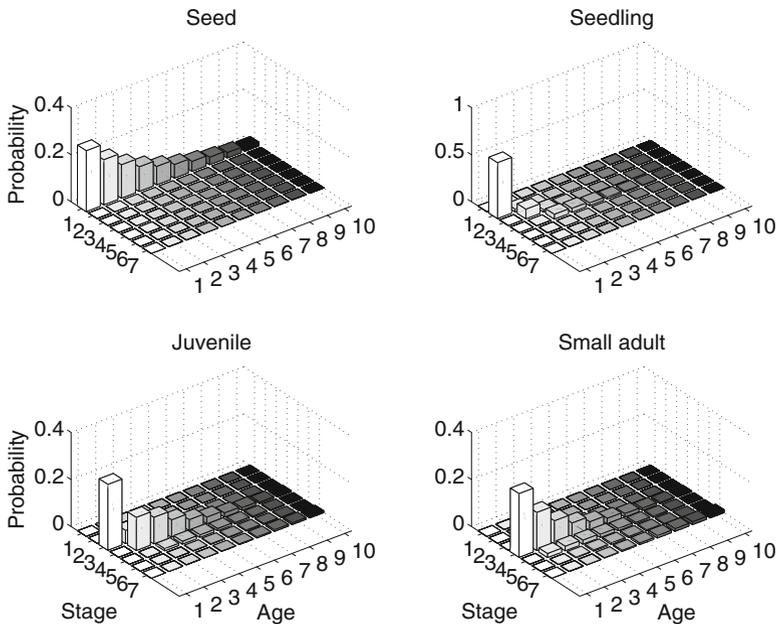


Fig. 6.3 The joint probability distribution of age (1, . . . , 10) and stage (1, . . . , 7) at death for an individual seed, seedling, juvenile, or small adult of Scotch broom. Stages as in Fig. 6.1

Fig. 6.4 The marginal distributions of age at death for individuals of Scotch broom in each stage. Maximum age is $\omega = 40$. Stages as in Fig. 6.1

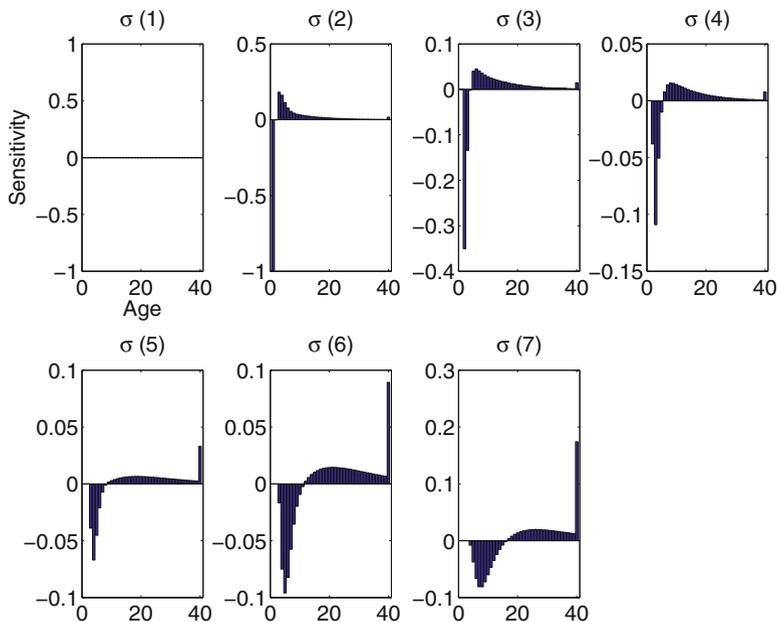
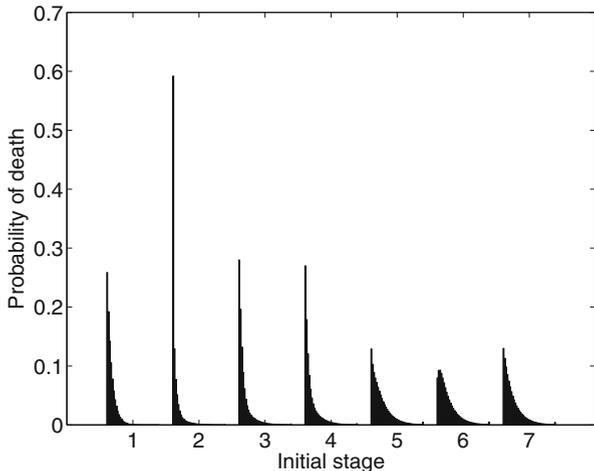


Fig. 6.5 Sensitivity of the marginal distribution of age at death, g_2 , to the survival probabilities of each stage, for an individual starting in stage 2 (seedlings). Stages as in Fig. 6.1

distribution, because seedlings have already left the seed stage. Changes in σ_2 – σ_7 shift the distribution to progressively older ages, by reducing the probability of death at young ages and increasing it at older ages.

6.5 Discussion

Models in which individuals are classified by both age and stage extend demographic analyses in several directions. They permit biodemographic analyses of aging to take advantage of the many stage-classified demographic analyses accumulated by ecologists (Salguero-Gómez et al. 2015, 2016). They also permit human demographers to take account of factors other than age in determining mortality, longevity, fertility, and population dynamics.

Age- and stage-specific demographic processes are often combined in demography using multistate life table methods (e.g., Rogers 1975; Willekens 2002, 2014). These are usually focused on cohort dynamics and associated survival statistics (but see Rogers 1975, Chap. 5 for an explicit consideration of population projection). Multistate life table models are written as continuous-parameter, discrete-state Markov chains, where the parameter represents age and the states represent stages. In order to solve the resulting equations, the dynamics must be approximated over a (usually short) finite age interval; this would correspond to the sequence of matrices \mathbf{A}_i in the model here. The age \times stage-classified model described by $\tilde{\mathbf{A}}$ is a way to solve the discretized equations in a single step, and makes possible a variety of analyses that are difficult or impossible in the usual life table formulation. Further investigation of the relation between continuous multistate life table methods and age \times stage-classified models will be interesting.

These analyses blur the distinction (Chap. 5) between implicit and explicit age dependence. If the \mathbf{A}_i are truly identical, by definition only implicit age dependence is revealed. But the structure of the age \times stage model separates all of the age-dependent \mathbf{A}_i , and thus is ready to include any degree of explicit joint dependence of the vital rates on age and stage.

Given sufficient longitudinal data on both age and stage, it is possible to estimate the stage-specific matrices \mathbf{A}_i as explicit functions of age; see Peeters et al. (2002) for an example of a study of human heart disease, and Lebreton et al. (2009) for a review of methods used in multistate capture-mark-recapture analysis in ecology. Needless to say, the data requirements for a full age \times stage parameterization are challenging. I suspect that the development of estimation methods at intermediate levels of detail will be an important step.

6.5.1 Reducibility and Ergodicity

The properties of $\tilde{\mathbf{A}}$ raise an important theoretical and technical issue regarding population growth, fitness, and selection gradients. The use of λ as a measure of fitness is usually justified by the strong ergodic theorem (Cohen 1979, Caswell 2001, Section 4.5.2), which guarantees the eventual convergence to the stable population structure and growth at a rate given by the dominant eigenvalue λ . A sufficient condition for this convergence is that the projection matrix be irreducible; i.e.,

that there exist a pathway connecting any two stages. Stott et al. (2010) surveyed published population projection matrices and found that reducible matrices were not uncommon, and explored the implications for ergodicity. Reducible matrices are not as bad as some people think, but it is important to understand their implications, especially for age \times stage models.

General results about the irreducibility of block-structured matrices are difficult; see Csetenyi and Logofet (1989), Logofet (1993, Chap. 3), and Logofet and Belova (2007) for some important graph-theoretical results. However, the age \times stage matrices developed here are unusual among population models in that they are (almost) always reducible, because they contain categories to which there are no possible pathways. This arises because age 1 individuals are produced only by reproduction. Hence there can never be age 1 individuals in any stage that is not produced by reproduction. For example, Scotch broom reproduces only by seeds, so age 1 seeds appear in the model. However, the matrix $\tilde{\mathbf{A}}$ also contains entries corresponding to age 1 seedlings, age 1 juveniles, age 1 adults, etc. These do not exist, and because there are no pathways to these stages from any other stages, the matrix $\tilde{\mathbf{A}}$ is reducible.

The Perron-Frobenius theorem guarantees that a reducible non-negative matrix will have a real, non-negative, dominant eigenvalue that is at least as large as any of the others. However, the asymptotic population growth rate and structure may depend on initial conditions (Caswell 2001, Section 4.5.4) This means that one must ascertain that the eigenvalues and eigenvectors under analysis correspond to initial conditions of interest.

Appendix A shows that a necessary and sufficient condition for population growth to be described by the dominant eigenvalue λ of $\tilde{\mathbf{A}}$, regardless of the (non-negative and non-zero) initial population vector, is that the left eigenvector \mathbf{v} be strictly positive, and that this corresponds to a particular block-triangular form of $\tilde{\mathbf{A}}$. This provides a simple check for the ergodicity of population growth, and justifies the use of λ as a population growth rate and measure of fitness.

Primitivity may be difficult to evaluate for an age \times stage matrix (but see Logofet 1993) but as with any projection matrix model, the long-term average growth rate of a primitive matrix is still given by the dominant real eigenvalue.

The matrix $\tilde{\mathbf{A}}$ for Scotch broom in (6.34) is reducible, as shown by calculating $(\mathbf{I}_{s\omega} + \tilde{\mathbf{A}})^{s\omega}$ and finding that this matrix contains zeros (Caswell 2001). However, the left eigenvector \mathbf{v} is strictly positive, so we know that the population eventually grows at the rate λ regardless of initial conditions.

6.5.2 A Protocol for Age \times Stage-Classified Models

The approach outlined here gives a step-by-step procedure for constructing and analyzing age \times stage-classified matrix population models.

1. Choose a question, and a corresponding demographic outcome. Are you interested in population dynamics (growth, structure, transients)? Or in cohort dynamics (survival, longevity)? Or in some combination of the two?
2. Obtain the stage-classified projection matrices \mathbf{A}_i for ages $i = 1, \dots, \omega$.
3. Decompose $\mathbf{A}_i = \mathbf{U}_i + \mathbf{F}_i$.
4. Construct the block-diagonal matrices \mathbb{A} , \mathbb{F} , \mathbb{U} , and \mathbb{D} , according to Eqs. (6.5)–(6.10).
5. Construct the age \times stage matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{F}}$, $\tilde{\mathbf{U}}$ using (6.16) and, if appropriate for the question at hand, also $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{P}}$ using (6.35) and (6.36).
6. Analyze the model, e.g., by computing eigenvalues, eigenvectors, the fundamental matrix, etc., as appropriate. If necessary, check for reducibility and ergodicity using the methods in Sect. 6.5.1.
7. For sensitivity analysis,
 - (a) choose a dependent variable ξ and a vector of parameters θ ,
 - (b) compute the sensitivity matrix $d\xi/d\text{vec}^T \tilde{\mathbf{A}}$,
 - (c) compute the matrices:

$$\frac{d\text{vec } \mathbf{A}_i}{d\theta^T}, \quad \frac{d\text{vec } \mathbf{U}_i}{d\theta^T}, \quad \text{and} \quad \frac{d\text{vec } \mathbf{F}_i}{d\theta^T}$$

- (d) compute $d\xi/d\theta^T$ according to (6.18).

The explicit connection between matrix population models and absorbing Markov chain theory makes it possible to analyze both population dynamics and cohort dynamics in a unified framework (cf. Feichtinger 1971a; Caswell 2001, 2006, 2009). Cohort dynamics are, in essence, the demography of individuals. It may seem paradoxical to speak of the demography of individuals, but that is what it is, because the *statistical* properties of a cohort (e.g., average lifespan) are *probabilistic* properties of an individual (e.g., life expectancy). Demography in general, and matrix population models in particular, provides the link between the individual and the population.

A Appendix: Population Growth and Reducible Matrices

Some ergodic properties of population growth under the action of reducible matrices are described by Caswell (2001, Section 4.5.4). Here we can extend the analysis.

Let \mathbf{A} be a reducible non-negative projection matrix. By permutation of its rows and columns (i.e., renumbering the stages in the life cycle), \mathbf{A} can be transformed to a block lower-triangular form. Here is an example:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B}_{11} & 0 & 0 & 0 \\ \mathbf{B}_{21} & \mathbf{B}_{22} & 0 & 0 \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} & 0 \\ \mathbf{B}_{41} & \mathbf{B}_{42} & \mathbf{B}_{43} & \mathbf{B}_{44} \end{pmatrix}. \tag{6.56}$$

In this form, all the diagonal blocks \mathbf{B}_{ii} are either irreducible matrices or 1×1 (i.e. scalar) zero matrices. The block triangular form is unique, up to a renumbering of the blocks and permutation of indices within blocks (Gantmacher 1959). It corresponds to a decomposition of the state space into a set of subspaces; let R_i be the subspace corresponding to the block \mathbf{B}_{ii} .

Some or all of the subdiagonal blocks in (6.56) may be zero. For reasons that will become apparent, consider an example where $\mathbf{B}_{21} = \mathbf{B}_{43} = \mathbf{0}$; i.e.,

$$\mathbf{A} = \begin{pmatrix} \mathbf{B}_{11} & 0 & 0 & 0 \\ 0 & \mathbf{B}_{22} & 0 & 0 \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} & 0 \\ \mathbf{B}_{41} & \mathbf{B}_{42} & 0 & \mathbf{B}_{44} \end{pmatrix} \quad (6.57)$$

Gantmacher (1959, Section 13.4) calls a block \mathbf{B}_{ii} *isolated* if there are no other non-zero blocks on its row, that is, if $\mathbf{B}_{ij} = 0$ for $j < i$. I will call such a block *row-isolated*, and introduce the term *column-isolated* to describe any block \mathbf{B}_{ii} with no other non-zero blocks in its column, that is, $\mathbf{B}_{ji} = 0$ for $j > i$. In the matrix in (6.57), the blocks \mathbf{B}_{11} and \mathbf{B}_{22} are row-isolated and the blocks \mathbf{B}_{33} and \mathbf{B}_{44} are column-isolated.

If \mathbf{B}_{ii} is row-isolated, then the life cycle graph contains no pathways from any state outside of the subspace R_i to any state inside R_i , and R_i is a source. If \mathbf{B}_{ii} is column-isolated, then the life cycle graph contains no pathways from any state in R_i to any state outside R_i , and R_i is a sink.

The eigenvalues of \mathbf{A} are the eigenvalues of the diagonal blocks \mathbf{B}_{ii} . Let λ_1 be the dominant eigenvalue of \mathbf{A} , with right and left eigenvectors \mathbf{w}_1 and \mathbf{v}_1 . The Perron-Frobenius theorem guarantees that λ_1 , \mathbf{w}_1 , and \mathbf{v}_1 are real and non-negative. Gantmacher (1959, Chap. 13, Theorem 6) proves that the eigenvector \mathbf{w}_1 is strictly positive if and only if λ_1 is an eigenvalue of every row-isolated block, and is not an eigenvalue of any of the non-row-isolated blocks. This makes it easy to demonstrate the following corollary.

Corollary: Positivity of \mathbf{v}_1 Let \mathbf{v}_1 be the left eigenvector corresponding to $\lambda_1[\mathbf{A}]$. Then \mathbf{v}_1 is strictly positive if and only if $\lambda_1[\mathbf{A}]$ is an eigenvalue of every column-isolated block, and is not an eigenvalue of any non-column-isolated block.

To see this, note that \mathbf{v}_1 is the right eigenvector of \mathbf{A}^\top . The column-isolated blocks of \mathbf{A} become row-isolated blocks of the block lower-triangular form of \mathbf{A}^\top , and application of Gantmacher's Theorem 6 proves the Corollary.

For example, transposing (6.57) gives

$$\mathbf{A}^\top = \begin{pmatrix} \mathbf{B}_{11}^\top & 0 & \mathbf{B}_{31}^\top & \mathbf{B}_{41}^\top \\ 0 & \mathbf{B}_{22}^\top & \mathbf{B}_{32}^\top & \mathbf{B}_{42}^\top \\ 0 & 0 & \mathbf{B}_{33}^\top & 0 \\ 0 & 0 & 0 & \mathbf{B}_{44}^\top \end{pmatrix} \quad (6.58)$$

Reversing the order of the rows and columns gives the block lower-triangular form

$$\begin{pmatrix} \mathbf{B}_{44}^\top & 0 & 0 & 0 \\ 0 & \mathbf{B}_{33}^\top & 0 & 0 \\ \mathbf{B}_{42}^\top & \mathbf{B}_{32}^\top & \mathbf{B}_{22}^\top & 0 \\ \mathbf{B}_{41}^\top & \mathbf{B}_{31}^\top & 0 & \mathbf{B}_{11}^\top \end{pmatrix} \quad (6.59)$$

The column-isolated blocks in \mathbf{A} (\mathbf{B}_{33} and \mathbf{B}_{44}) now appear as row-isolated blocks in \mathbf{A}^\top . Gantmacher’s result shows that the eigenvector \mathbf{v}_1 will be positive if and only if λ_1 is an eigenvalue of each of those blocks.

The usefulness of the Corollary follows from the population projection model

$$\mathbf{n}(t + 1) = \mathbf{A}\mathbf{n}(t) \quad \mathbf{n}(0) = \mathbf{n}_0 \quad (6.60)$$

and its solution⁴

$$\mathbf{n}(t) = \sum_{i=1}^s c_i \lambda_i^t \mathbf{w}_i \quad (6.61)$$

$$= \sum_{i=1}^s \left(\mathbf{v}_i^\top \mathbf{n}_0 \right) \lambda_i^t \mathbf{w}_i \quad (6.62)$$

Caswell (2001). If \mathbf{n}_0 is such that $c_1 = \mathbf{v}_1^\top \mathbf{n}_0$ is positive, then λ_1^t will eventually dominate all other terms in the solution and the population will grow at the rate λ_1 with stable structure \mathbf{w}_1 . We know the following about c_1 :

1. If \mathbf{A} is irreducible, then by the Perron-Frobenius theorem \mathbf{v}_1 is strictly positive, so any non-negative, non-zero initial population \mathbf{n}_0 leads to a positive value of c_1 and eventual growth at the rate λ_1 .
2. If \mathbf{A} is reducible and \mathbf{v}_1 is strictly positive, any non-negative, non-zero \mathbf{n}_0 leads to a positive value of c_1 and growth at the rate λ_1 .
3. If \mathbf{A} is reducible and \mathbf{v}_1 contains zero entries corresponding to a subspace R_i , then initial conditions with positive support only in R_i will lead to $c_1 = 0$, and λ_1 will make no contribution to population growth from those initial vectors.

In the first two cases, population growth is ergodic from any non-zero initial population. In the third case, there exists a basin of attraction leading to growth according to λ_1 , and a basin (or basins) of attraction for growth according to the dominant eigenvalues of the diagonal blocks \mathbf{B}_{ii} corresponding to the zero entries of \mathbf{v}_1 .

⁴This holds provided that \mathbf{A} is diagonalizable, which is a generic property for linear operators (Hirsch and Smale 1974, p. 157).

Bibliography

- Barfield, M., R. D. Holt, and R. Gomulkiewicz. 2011. Evolution in stage-structured populations. *American Naturalist* **177**:397–409.
- Caswell, H. 2001. *Matrix Population Models: Construction, Analysis, and Interpretation*. 2nd edition. Sinauer Associates, Sunderland, MA.
- Caswell, H., 2006. Applications of Markov chains in demography. Pages 319–334 in MAM2006: Markov Anniversary Meeting. Bosen Books, Raleigh, North Carolina.
- Caswell, H. 2009. Stage, age and individual stochasticity in demography. *Oikos* **118**:1763–1782.
- Caswell, H. 2010. Reproductive value, the stable stage distribution, and the sensitivity of the population growth rate to changes in vital rates. *Demographic Research* **23**:531–548.
- Caswell, H. 2012. Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theoretical Ecology* **5**:403–417.
- Caswell, H. 2014. A matrix approach to the statistics of longevity in heterogeneous frailty models. *Demographic Research* **31**:553–592.
- Caswell, H., C. de Vries, N. Hartemink, G. Roth, and S. F. van Daalen. 2018. Age \times stage-classified demographic analysis: a comprehensive approach. *Ecological Monographs* **88**:560–584.
- Caswell, H., and N. Ouellette, 2016. Mortality and causes of death: matrix formulation and sensitivity analysis. European Population Conference 2016, Mainz, Germany. URL <http://epc2016.princeton.edu/papers/160437>.
- Caswell, H., and N. Ouellette, 2018. Cause-of-death analysis: matrix formulation and sensitivity analysis. In prep.
- Caswell, H., and R. Salguero-Gómez. 2013. Age, stage and senescence in plants. *Journal of Ecology* **101**:585–595.
- Caswell, H., and S. F. van Daalen. 2016. A note on the vec operator applied to unbalanced block-structured matrices. *Journal of Applied Mathematics* **2016**:1–3.
- Cohen, J. E. 1979. Ergodic theorems in demography. *Bulletin of the American Mathematical Society* **1**:275–295.
- Csetenyi, A. I., and D. O. Logofet. 1989. Leslie model revisited: some generalizations to block structures. *Ecological Modelling* **48**:277–290.
- Dercole, F., and S. Rinaldi. 2008. *Analysis of evolutionary processes: the adaptive dynamics approach and its applications*. Princeton University Press, Princeton, NJ, USA.
- Feichtinger, G. 1971a. *Stochastische Modelle Demographischer Prozesse*. Lecture notes in operations research and mathematical systems, Springer-Verlag, Berlin, Germany.
- Feichtinger, G. 1971b. *Stochastische Modelle demographischer Prozesse*. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin.
- Gantmacher, F. R. 1959. *Matrix Theory*. Chelsea, New York, New York.
- Goodman, L. A. 1969. The analysis of population growth when the birth and death rates depend upon several factors. *Biometrics* **25**:659–681.
- Hamilton, W. D. 1966. The moulding of senescence by natural selection. *Journal of Theoretical Biology* **12**:12–45.
- Henderson, H. V., and S. R. Searle. 1981. The vec-permutation matrix, the vec operator and Kronecker products: a review. *Linear and Multilinear Algebra* **9**:271–288.
- Hirsch, M. W., and S. Smale. 1974. *Differential equations, dynamical systems, and linear algebra*. Academic Press, New York, NY, USA.
- Horvitz, C. C., and S. Tuljapurkar. 2008. Stage dynamics, period survival, and mortality plateaus. *American Naturalist* **172**:203–215.
- Hunter, C. M., and H. Caswell. 2005. The use of the vec-permutation matrix in spatial matrix population models. *Ecological Modelling* **188**:15–21.

- Iosifescu, M. 1980. *Finite Markov Processes and Their Applications*. Wiley, New York, New York.
- Klepac, P., and H. Caswell. 2011. The stage-structured epidemic: linking disease and demography with a multi-state matrix approach model. *Theoretical Ecology* **4**:301–319.
- Lebreton, J. 1996. Demographic models for subdivided populations: the renewal equation approach. *Theoretical Population Biology* **49**:291–313.
- Lebreton, J.-D., J. D. Nichols, R. J. Barker, R. Pradel, and J. A. Spendelov. 2009. Modeling individual animal histories with multistate capture-recapture models. *Advances in Ecological Research* **41**:87–173.
- Logofet, D. O. 1993. *Matrices and graphs: stability problems in mathematical ecology*. CRC Press, Boca Raton, FL, USA.
- Logofet, D. O., and I. N. Belova. 2007. Nonnegative matrices as a tool to model population dynamics: classical models and contemporary expansions (in Russian with English summary). *Journal of Mathematical Sciences* **155**:894–907.
- Magnus, J. R., and H. Neudecker. 1985. Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. *Journal of Mathematical Psychology* **29**:474–492.
- Metz, J. A. J., and O. Diekmann. 1986. *The dynamics of physiologically structured populations*. Springer-Verlag, Berlin, Germany.
- Metz, J. A. J., R. M. Nisbet, and S. A. H. Geritz. 1992. How should we define 'fitness' for general ecological scenarios? *Trends in Ecology and Evolution* **7**:198–202.
- Neubert, M. G., and I. M. Parker. 2004. Projecting rates of spread for invasive species. *Risk Analysis* **24**:817–831.
- Parker, I. M. 2000. Invasion dynamics of *Cytisus scoparius*: a matrix model approach. *Ecological Applications* **10**:726–743.
- Pascarella, J. B., and C. C. Horvitz. 1998. Hurricane disturbance and the population dynamics of a tropical understory shrub: megamatrix elasticity analysis. *Ecology* **79**:547–563.
- Peeters, A., A. A. Mamun, F. Willekens, and L. Bonneux. 2002. A cardiovascular life history: a life course analysis of the original Framingham Heart Study cohort. *European Heart Journal* **23**:458–466.
- Rice, S. H. 2004. *Evolutionary theory: mathematical and conceptual foundations*. Sinauer, Sunderland, MA, USA.
- Rogers, A. 1966. The multiregional matrix growth operator and the stable interregional age structure. *Demography* **3**:537–544.
- Rogers, A. 1975. *Introduction to multiregional mathematical demography*. Wiley, New York, NY, USA.
- Salguero-Gómez, R., O. R. Jones, C. R. Archer, C. Bein, H. Buhr, C. Farack, F. Gottschalk, A. Hartmann, A. Henning, G. Hoppe, G. Romer, T. Ruoff, V. Sommer, J. Wille, J. Voigt, S. Zeh, D. Vieregg, Y. M. Buckley, J. Che-Castaldo, D. Hodgson, A. Scheuerlein, H. Caswell, and J. W. Vaupel. 2016. COMADRE: a global data base of animal demography. *Journal of Animal Ecology* **85**:371–384.
- Salguero-Gómez, R., O. R. Jones, C. R. Archer, Y. M. Buckley, J. Che-Castaldo, H. Caswell, D. Hodgson, A. Scheuerlein, D. A. Conde, E. Brinks, et al. 2015. The COMPADRE Plant Matrix Database: an open online repository for plant demography. *Journal of Ecology* **103**:202–218.
- Stott, I., S. Townley, D. Carslake, and D. J. Hodgson. 2010. On reducibility and ergodicity of population projection matrix models. *Methods in Ecology and Evolution* **1**:242–252.
- Tuljapurkar, S., and C. C. Horvitz. 2006. From stage to age in variable environments: life expectancy and survivorship. *Ecology* **87**:1497–1509.
- Willekens, F. 2014. *Multistate Analysis of Life Histories with R*. Springer, New York, New York.
- Willekens, F. J., 2002. Multistate demography. in *Encyclopaedia of Population*. Macmillan Reference, New York, NY, USA.

- Wu, G.-M., Y.-M. Wang, M.-F. Yen, J.-M. Wong, H.-C. Lai, J. Warwick, and C. TH-H. 2006. Cost-effectiveness analysis of colorectal cancer screening with stool DNA testing in intermediate-incidence countries. *BMC Cancer* **6**:136.
- Zhou, Y., H. Putter, and G. Doblhammer. 2016. Years of life lost due to lower extremity injury in association with dementia, and care need: A 6-year follow-up population-based study using a multi-state approach among German elderly. *BMC Geriatrics* **16**:9.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

