Calculating second derivatives of population growth rates for ecology and evolution

Esther Shyu[1,*] and Hal Caswell[1,2]

[1]*Biology Department MS-34, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA; and* [2]*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*

Summary

1. Second derivatives of the population growth rate measure the curvature of its response to demographic, physiological or environmental parameters. The second derivatives quantify the response of sensitivity results to perturbations, provide a classification of types of selection and provide one way to calculate sensitivities of the stochastic growth rate.

2. Using matrix calculus, we derive the second derivatives of three population growth rate measures: the discrete-time growth rate λ, the continuous-time growth rate $r = \log \lambda$ and the net reproductive rate R_0, which measures per-generation growth.

3. We present a suite of formulae for the second derivatives of each growth rate and show how to compute these derivatives with respect to projection matrix entries and to lower-level parameters affecting those matrix entries.

4. We also illustrate several ecological and evolutionary applications for these second derivative calculations with a case study for the tropical herb *Calathea ovandensis*.

Key-words: eigenvalues, Hessian matrix, invasion exponent, matrix population models, net reproductive rate, sensitivity analysis

Introduction

Using matrix population models, ecological indices can be calculated as functions of vital rates such as survival or fertility. Measures of population growth rate, including the discrete-time growth rate λ, the continuous-time growth rate $r = \log \lambda$ and the net reproductive rate R_0, are of particular interest. The discrete-time population growth rate λ is given by the dominant eigenvalue of the population projection matrix. Sensitivities (first partial derivatives) of λ with respect to relevant parameters quantify how population growth responds to vital rate perturbations. These first derivatives are used to project the effects of vital rate changes due to environmental or management perturbations, uncertainty in parameter estimates and phenotypic evolution (i.e. with λ as a fitness measure, the sensitivity of λ with respect to a parameter is the selection gradient on that parameter) (Caswell 2001).

APPLICATIONS OF SECOND DERIVATIVES OF GROWTH RATES

The second derivatives of growth rates have applications in both ecology (e.g. assessing and improving recommendations from sensitivity analysis, approximating the sensitivities of stochastic growth rates) and evolution (e.g. characterizing nonlinear selection gradients and evolutionary equilibria). Several of

these applications are summarized in Table 1 and described in the following sections.

Second-order sensitivity analysis and growth rate estimation

The sensitivity of growth rate provides insight into the population response to parameter perturbations. However, such perturbations also affect the sensitivity itself, that is, sensitivity is 'situational' (Stearns 1992). These second-order effects are quantified by the sensitivity, with respect to a parameter θ_j, of the sensitivity of λ to another parameter θ_i, that is, by the second derivatives $\frac{\partial^2 \lambda}{\partial \theta_j \partial \theta_i}$. The sensitivity of the elasticity of growth rate to parameters similarly depends on second derivatives (Caswell, 1996, 2001).

In conservation applications, attention is often focused on the vital rates to which population growth is particularly sensitive or elastic; these first-order results may change depending on parameter perturbations. First derivatives also provide a linear, first-order approximation to the response of the growth rate to changes in parameters. The linear approximation is guaranteed to be accurate for sufficiently small perturbations and is often very accurate even for quite large perturbations (Caswell 2001). If the response of λ to θ is nonlinear, it is tempting to use a second-order approximation for $\Delta\lambda$:

$$\Delta\lambda \approx \sum_i \frac{\partial \lambda}{\partial \theta_i} \Delta\theta_i + \sum_i \frac{1}{2} \frac{\partial^2 \lambda}{\partial \theta_i^2} (\Delta\theta_i)^2 + \sum_{i \neq j} \frac{\partial^2 \lambda}{\partial \theta_i \partial \theta_j} (\Delta\theta_i)(\Delta\theta_j)$$

eqn 1

*Correspondence author. E-mail: eshyu@whoi.edu

Table 1. Potential applications for the pure and mixed second derivatives of λ. Analogous interpretations apply to r or R_0 as alternative measures of growth or fitness

Second derivative	Sign	Interpretations
$\dfrac{\partial^2\lambda}{\partial\theta^2}$	=0	Sensitivity of λ to θ is independent of θ
		Linear selection on trait θ
	>0	Sensitivity of λ to θ increases with θ
		Convex selection on trait θ
		Evolutionarily unstable singular strategy
	<0	Sensitivity of λ to θ decreases with increases in θ
		Concave selection on trait θ
		Evolutionarily stable singular strategy
$\dfrac{\partial^2\lambda}{\partial\theta_j\partial\theta_i}$	>0	Sensitivity of λ to θ_i increases with θ_j
		Selection to increase correlation between traits θ_j and θ_i
	<0	Sensitivity of λ to θ_i decreases with increases in θ_j
		Selection to decrease correlation between traits θ_j and θ_i
$H[\lambda; \text{vec}\bar{\mathbf{A}}]$	N/A	Used to calculate sensitivity of the stochastic growth rate λ_s

We caution that although this may, in some cases, provide a more accurate calculation, this is not guaranteed. As shown in Fig. 1 of Carslake, Townley & Hodgson (2008), for example, adding the second-order terms may actually reduce the accuracy of the approximation.

Characterizing nonlinear selection processes

The second derivatives of fitness with respect to trait values have consequences for selection. The first derivatives of fitness are selection gradients (Lande 1982). When fitness is a linear function of a trait, its second derivatives are zero, and there is selection to shift the trait's mean value. When fitness is a nonlinear function of a trait, its second derivatives are nonzero and provide additional information on how selection affects the trait's higher moments (Lande & Arnold 1983, Phillips & Arnold 1989, Brodie, Moore & Janzen 1995). Such nonlinear selection can be classified as concave or convex depending on whether the second derivatives are negative or positive.

One can classify a selection process as linear, concave or convex using quadratic selection gradients, the local second derivatives of fitness with respect to trait value (Phillips & Arnold 1989). If fitness is measured as λ, these quadratic selection gradients are equivalent to $\partial^2\lambda/\partial\theta^2$, the pure second derivatives of λ with respect to trait θ (e.g. the second derivatives with respect to stage-specific survival in *C. ovandensis*, as shown in Fig. 3a). Concave, linear and convex selection correspond to negative, zero and positive second derivatives, respectively.

Concave selection reduces the variance in the trait, and convex selection increases it; Lande & Arnold (1983, p.1216) equate this to a more sophisticated version of the concepts of stabilizing and disruptive selection. Brodie, Moore & Janzen (1995) provide further analysis of the curvature of the fitness surface and its effects on selection.

Selection operating on pairs of traits is said to be correlational if the cross second derivatives are nonzero. Thus, if the pure second derivatives of two different traits, θ_i and θ_j, are both nonzero, their mixed second derivative $\partial^2\lambda/\partial\theta_j\partial\theta_i$ is a measure of correlational selection. If $\partial^2\lambda/\partial\theta_j\partial\theta_i < 0$, there is selection to decrease the phenotypic correlation between the two traits; if $\partial^2\lambda/\partial\theta_j\partial\theta_i > 0$, there is selection to increase their correlation. The concepts of nonlinear selection are powerful, but require the second derivatives of fitness to be applied.

Stability of evolutionary singular strategies

Second derivatives play a role in adaptive dynamic analyses. Evolutionary singular strategies (SSs) are phenotypes for which the selection gradient is locally zero (e.g. Geritz *et al*. 1998). SSs are classified as stable, attracting or repelling, and by whether they can invade or coexist with other nearby phenotypes (Geritz *et al*. 1998, Diekmann 2004, Waxman & Gavrilets 2005, Doebeli 2011).

These classifications depend on the local second derivatives of invasion fitness, the growth rate of a rare mutant in an equilibrium resident environment. For example, the second derivative of the mutant growth rate λ to the mutant trait y determines whether a SS is evolutionarily stable ($\partial^2\lambda/\partial y^2 < 0$) or evolutionarily unstable ($\partial^2\lambda/\partial y^2 > 0$). Evolutionarily stable strategies, once established, are unbeatable phenotypes against which no nearby mutants can increase under selection and are thus long-term evolutionary endpoints. Evolutionarily unstable strategies, on the other hand, are branching points open to phenotypic divergence and may ultimately become sources of sympatric speciation (Geritz *et al*. 1998).

Sensitivity of the stochastic growth rate

Second derivatives provide a way to calculate the sensitivity of the stochastic growth rate in some cases. The stochastic growth rate is

$$\log\lambda_s = \lim_{t\to\infty}\frac{1}{t}\log N(t) \qquad \text{eqn 2}$$

where $N(t)$ is the population size at time t. Tuljapurkar (1982) derived a small-noise approximation for $\log\lambda_s$ in the absence of temporal autocorrelation. As shown by Caswell (2001 Section 14.3.6), this approximation can be written in terms of the first derivatives of λ, the dominant eigenvalue of the mean

projection matrix $\bar{\mathbf{A}}$. Thus, the derivatives of this approximation can be written in terms of the second derivatives of that eigenvalue (Caswell 2001, Section 14.3.6). We discuss this application further in the section 'Sensitivity analysis of stochastic growth rates'.

CALCULATING SECOND DERIVATIVES OF GROWTH RATES

The second derivatives of λ with respect to matrix elements were introduced by Caswell (1996); see also Caswell (2001, Section 9.7). However, these calculations are awkward and error-prone, because they involve all the eigenvalues and eigenvectors of the projection matrix. McCarthy, Townley & Hodgson (2008) introduced an alternative approach for calculating the second derivatives of eigenvalues (they call them 'second-order sensitivities') based on transfer functions, partially to avoid the calculation of all the eigenvectors. However, they consider only rank-one perturbations of a subset of the matrix elements, excluding fertilities, and their calculations are perhaps equally difficult.

Here, we reformulate the second derivative calculations using matrix calculus, providing easily computable results. We extend previous results by including not only second derivatives with respect to matrix elements, but also those with respect to any lower-level parameters that may affect the matrix elements, and by presenting the second derivatives of the continuous-time invasion exponent r and the net reproductive rate R_0.

The key to our approach is that the calculation of first derivatives using matrix calculus yields a particular expression, the differentiation of which leads directly to the second derivatives. Second derivatives are easily computed by this method in any matrix-oriented language, such as MATLAB or R. Although we consider only the second derivatives of population growth rates, our approach extends naturally to other scalar-dependent variables.

In the section 'A case study: *Calathea ovandensis*', we present an example of the calculation of second derivatives in a case study of the tropical herb *Calathea ovandensis*.

NOTATION

Matrices are denoted by upper-case boldface letters (e.g. $\mathbf{A}$) and vectors by lower-case boldface letters (e.g. $\mathbf{w}$); unless otherwise indicated, all vectors are column vectors. Transposes of matrices and vectors are indicated by the superscript T. The matrix $\mathbf{I}_n$ is the $n \times n$ identity matrix, the vector $\mathbf{e}$ is a vector of ones, and $\mathbf{e}_1$ is a vector with 1 as its first entry and zeros elsewhere. The matrix $\mathbf{K}_{m,n}$ is a $mn \times mn$ commutation matrix (vec-permutation matrix) (Magnus & Neudecker 1979, Henderson & Searle 1981), which can be calculated using the MATLAB function provided in Appendix S1-D. The expression diag($\mathbf{x}$) indicates the square matrix with $\mathbf{x}$ on the diagonal and zeros elsewhere.

The Kronecker product is denoted by $\mathbf{X} \otimes \mathbf{Y}$ and the Hadamard (element-by-element) product by $\mathbf{X} \circ \mathbf{Y}$. The vec operator

(e.g. vec$\mathbf{A}$) stacks the columns of a matrix into a single vector. For convenience, we will write (vec$\mathbf{A}$)$^\mathsf{T}$ as vec$^\mathsf{T}\mathbf{A}$. We will make frequent use of Roth's theorem (Roth 1934), which states that for any matrices $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$:

$$\mathrm{vec}(\mathbf{XYZ}) = (\mathbf{Z}^\mathsf{T} \otimes \mathbf{X})\mathrm{vec}\mathbf{Y} . \qquad \text{eqn 3}$$

Matrix calculus

MATRIX CALCULUS NOTATION

Matrix calculus is a system for manipulating vectors and matrices in multivariable calculus and simplifies partial derivative calculations by allowing the differentiation of scalar, vector or matrix functions with respect to scalar, vector or matrix arguments. While there are multiple matrix calculus notations, we will use the system of Magnus & Neudecker (1999). For a more detailed introduction to these methods in an ecological context, see Appendix 1 of Caswell (2007).

The first derivative of a $m \times 1$ vector $\mathbf{y}$ with respect to a $n \times 1$ vector $\mathbf{x}$ is defined to be the $m \times n$ Jacobian matrix

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{x}^\mathsf{T}} = \left(\frac{\mathrm{d}y_i}{\mathrm{d}x_j}\right); \qquad \text{eqn 4}$$

that is, a matrix whose (i,j) entry is the derivative of y_i with respect to x_j. We will also write this as an operator $D[\mathbf{y};\mathbf{x}]$; the first argument of D is the vector-valued function $\mathbf{y}$ to be differentiated, and the second argument is the vector-valued variable $\mathbf{x}$ with respect to which differentation is carried out. Thus,

$$D[\mathbf{y}; \mathbf{x}] = \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{x}^\mathsf{T}} . \qquad \text{eqn 5}$$

As in the scalar case, second derivatives are obtained by differentiating first derivatives. If we consider a scalar-valued function $y(\mathbf{x})$ of a vector-valued argument $\mathbf{x}$, the matrix of second derivatives (the Hessian matrix) is given by the operator

$$H[y; \mathbf{x}] = \left(\frac{\mathrm{d}^2 y}{\mathrm{d}x_i \mathrm{d}x_j}\right) \qquad \text{eqn 6}$$

$$= \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}^\mathsf{T}} \left[\left(\frac{\mathrm{d}y}{\mathrm{d}\mathbf{x}^\mathsf{T}}\right)^\mathsf{T}\right] . \qquad \text{eqn 7}$$

The matrix of second derivatives of a vector-valued function $\mathbf{y}(\mathbf{x})$, where $\mathbf{y}$ has dimensions $m \times 1$, is obtained by stacking the Hessian matrices for each of the elements of $\mathbf{y}$; that is,

$$H[\mathbf{y}; \mathbf{x}] = \begin{pmatrix} H[y_1; \mathbf{x}] \\ H[y_2; \mathbf{x}] \\ \vdots \\ H[y_m; \mathbf{x}] \end{pmatrix} \qquad \text{eqn 8}$$

$$= \frac{\mathrm{d}}{\mathrm{d}\mathbf{x}^\mathsf{T}} \mathrm{vec}\left[\left(\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{x}^\mathsf{T}}\right)^\mathsf{T}\right] . \qquad \text{eqn 9}$$

These first and second derivative definitions are written in terms of vector-valued functions and arguments. When matrices appear, they are transformed into vectors using the vec

operator, which stacks the columns of the matrix into a column vector. Thus, the first and second derivatives of λ with respect to the entries of the matrix **A** would be written, respectively, as $D[\lambda;\text{vec}\mathbf{A}]$ and $H[\lambda;\text{vec}\mathbf{A}]$.

THE IDENTIFICATION THEOREMS

Magnus & Neudecker (1985, 1999) showed how to obtain first and second derivatives from the differentials of functions. Their 'first identification theorem' showed that

$$d\mathbf{y} = \mathbf{Q}d\mathbf{x} \quad \Longrightarrow \quad D[\mathbf{y};\mathbf{x}] = \mathbf{Q}. \qquad \text{eqn 10}$$

That is, if an expression of the form d**y**=**Q**d**x** can be obtained, then the Jacobian matrix of first derivatives is given by **Q**.

The 'second identification theorem' does the same for the Hessian matrix of second derivatives, showing that

$$d^2 y = d\mathbf{x}^\mathsf{T}\mathbf{B}d\mathbf{x} \quad \Longrightarrow \quad H[y;\mathbf{x}] = \frac{1}{2}(\mathbf{B} + \mathbf{B}^\mathsf{T}). \qquad \text{eqn 11}$$

Thus, our goal will be to find expressions of the form $d^2 y = d\mathbf{x}^\mathsf{T}\mathbf{B}d\mathbf{x}$, where y is a measure of population growth rate and **x** represents either matrix entries or lower-level parameters; the matrix **B** will then provide the Hessian matrix using (11). The key to our approach is to begin with the expression (10) for the first differential, differentiate it to obtain the second differential and manipulate the results to obtain a matrix **B** in the form of (11).

Second derivatives of growth rates

We now apply the identification theorems to three measures of population growth rate, the discrete-time growth rate λ, the continuous-time growth rate $r = \log \lambda$ and the net reproductive rate R_0.

SECOND DERIVATIVES OF THE DISCRETE-TIME GROWTH RATE λ

Second derivatives of λ with respect to matrix entries: $H[\lambda;\text{vec}\mathbf{A}]$

We assume a population projection matrix **A** of dimension $n \times n$. The discrete-time growth rate λ is the dominant eigenvalue of **A**. To derive $H[\lambda;\text{vec}\mathbf{A}]$, we begin with an expression of the form (10) for the first differential of λ. As shown in Caswell (2010),

$$d\lambda = (\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T})d\text{vec}\mathbf{A} \qquad \text{eqn 12}$$

where **w** and **v** are the right and left eigenvectors of **A** corresponding to λ, scaled so that

$$\mathbf{v}^\mathsf{T}\mathbf{w} = 1 \qquad \text{eqn 13}$$

$$\mathbf{e}^\mathsf{T}\mathbf{w} = 1 \qquad \text{eqn 14}$$

where **e** is a $n \times 1$ vector of ones.

Differentiate (12) to obtain the second differential

$$d^2\lambda = d(\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T})d\text{vec}\mathbf{A} + (\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T})d^2\text{vec}\mathbf{A}. \qquad \text{eqn 15}$$

Because we are calculating second derivatives with respect to **A**, the second term will drop out because $d^2\text{vec}\mathbf{A} = 0$ (Magnus & Neudecker 1999). Apply the vec operator to obtain

$$d^2\lambda = d(\text{vec}^\mathsf{T}\mathbf{A})d\text{vec}(\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T}). \qquad \text{eqn 16}$$

The differential of $\text{vec}(\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T})$ is

$$d\text{vec}(\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T}) = (\mathbf{I}_n \otimes \mathbf{v})d\mathbf{w} + (\mathbf{w} \otimes \mathbf{I}_n)d\mathbf{v} \qquad \text{eqn 17}$$

(Magnus & Neudecker 1999). Substituting (17) into (16) gives

$$d^2\lambda = d(\text{vec}^\mathsf{T}\mathbf{A})[(\mathbf{I}_n \otimes \mathbf{v})d\mathbf{w} + (\mathbf{w} \otimes \mathbf{I}_n)d\mathbf{v}]. \qquad \text{eqn 18}$$

By the chain rule,

$$d\mathbf{w} = \frac{d\mathbf{w}}{d\text{vec}^\mathsf{T}\mathbf{A}}d\text{vec}\mathbf{A} \qquad \text{eqn 19}$$

$$d\mathbf{v} = \frac{d\mathbf{v}}{d\text{vec}^\mathsf{T}\mathbf{A}}d\text{vec}\mathbf{A}. \qquad \text{eqn 20}$$

The first derivatives of **w** and **v**, subject to (13) and (14), are given in Caswell (2008) and H. Caswell and Y. Vindenes (unpublished data), respectively, as:

$$\frac{d\mathbf{w}}{d\text{vec}^\mathsf{T}\mathbf{A}} = (\lambda\mathbf{I}_n - \mathbf{A} + \mathbf{we}^\mathsf{T}\mathbf{A})^{-1}[\mathbf{w}^\mathsf{T} \otimes (\mathbf{I}_n - \mathbf{we}^\mathsf{T})] \qquad \text{eqn 21}$$

$$\frac{d\mathbf{v}}{d\text{vec}^\mathsf{T}\mathbf{A}} = (\lambda\mathbf{I}_n - \mathbf{A}^\mathsf{T} + \lambda\mathbf{vw}^\mathsf{T})^{-1}$$
$$\left([(\mathbf{I}_n - \mathbf{vw}^\mathsf{T}) \otimes \mathbf{v}^\mathsf{T}] - \lambda(\mathbf{v} \otimes \mathbf{v}^\mathsf{T})\frac{d\mathbf{w}}{d\text{vec}^\mathsf{T}\mathbf{A}}\right). \qquad \text{eqn 22}$$

Substituting (19) and (20) into (18) gives

$$d^2\lambda = d\text{vec}^\mathsf{T}\mathbf{A}\left[(\mathbf{I}_n \otimes \mathbf{v})\frac{d\mathbf{w}}{d\text{vec}^\mathsf{T}\mathbf{A}} + (\mathbf{w} \otimes \mathbf{I}_n)\frac{d\mathbf{v}}{d\text{vec}^\mathsf{T}\mathbf{A}}\right]d\text{vec}\mathbf{A}. \qquad \text{eqn 23}$$

This is of the form

$$d^2\lambda = (d\text{vec}^\mathsf{T}\mathbf{A})\mathbf{B}(d\text{vec}\mathbf{A}) \qquad \text{eqn 24}$$

and hence

$$H[\lambda;\text{vec}\mathbf{A}] = \frac{1}{2}(\mathbf{B} + \mathbf{B}^\mathsf{T}) \qquad \text{eqn 25a}$$

where

$$B = (\mathbf{I}_n \otimes \mathbf{v})\frac{d\mathbf{w}}{d\text{vec}^\mathsf{T}\mathbf{A}} + (\mathbf{w} \otimes \mathbf{I}_n)\frac{d\mathbf{v}}{d\text{vec}^\mathsf{T}\mathbf{A}} \qquad \text{eqn 25b}$$

and the first derivatives of **w** and **v** are given by (21) and (22).

Second derivatives of λ with respect to lower-level parameters: $H[\lambda;\theta]$

Because many life-history traits and environmental factors affect multiple life cycle transitions, the entries of **A** are usually functions of lower-level parameters. The first derivatives with respect to lower-level parameters are calculated with the chain rule. To calculate the second derivatives of λ with respect to a

$s \times 1$ vector $\boldsymbol{\theta}$ of lower-level parameters, we must develop a chain rule for the Hessian.

To do so, we begin with the first differential of λ in (12) and differentiate to obtain the second differential (15). Because we are calculating second derivatives with respect to $\boldsymbol{\theta}$ rather than $\mathbf{A}$, $d^2 \text{vec} \mathbf{A}$ is no longer zero. By the chain rule,

$$d\text{vec}\mathbf{A} = \frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}} d\boldsymbol{\theta}. \qquad \text{eqn 26}$$

Differentiate (26) to obtain

$$d^2\text{vec}\mathbf{A} = d\left(\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}\right)d\boldsymbol{\theta} + \frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}d^2\boldsymbol{\theta}. \qquad \text{eqn 27}$$

Because $d^2\boldsymbol{\theta} = 0$, the second term drops out.

Substituting (26) and (27) into the expression for the second differential in (15) yields

$$d^2\lambda = d(\mathbf{w}^{\mathsf{T}} \otimes \mathbf{v}^{\mathsf{T}})\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}d\boldsymbol{\theta} + (\mathbf{w}^{\mathsf{T}} \otimes \mathbf{v}^{\mathsf{T}})d\left(\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}\right)d\boldsymbol{\theta}.$$
$$\text{eqn 28}$$

To simplify this expression, define

$$\mathbf{S} = \frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}} \qquad \text{eqn 29}$$

$$\mathbf{T} = \frac{d\lambda}{d\text{vec}^{\mathsf{T}}\mathbf{A}} = \mathbf{w}^{\mathsf{T}} \otimes \mathbf{v}^{\mathsf{T}} \qquad \text{eqn 30}$$

in terms of which (28) can be rewritten as

$$d^2\lambda = [(d\mathbf{T})\mathbf{S} + \mathbf{T}(d\mathbf{S})]d\boldsymbol{\theta}. \qquad \text{eqn 31}$$

Then apply the vec operator and Roth's theorem (3) to obtain

$$d^2\lambda = d\boldsymbol{\theta}^{\mathsf{T}}\text{vec}[\mathbf{T}(d\mathbf{S}) + (d\mathbf{T})\mathbf{S}] \qquad \text{eqn 32}$$

$$= d\boldsymbol{\theta}^{\mathsf{T}}[(\mathbf{I}_s \otimes \mathbf{T})d\text{vec}\mathbf{S} + \mathbf{S}^{\mathsf{T}}d\text{vec}\mathbf{T}] \qquad \text{eqn 33}$$

$$= d\boldsymbol{\theta}^{\mathsf{T}}\left[(\mathbf{I}_s \otimes \mathbf{T})\frac{d\text{vec}\mathbf{S}}{d\boldsymbol{\theta}^{\mathsf{T}}} + \mathbf{S}^{\mathsf{T}}\frac{d\text{vec}\mathbf{T}}{d\boldsymbol{\theta}^{\mathsf{T}}}\right]d\boldsymbol{\theta} \qquad \text{eqn 34}$$

where, as shown by (A-10) and (A-13) in Appendix S1-A,

$$(\mathbf{I}_s \otimes \mathbf{T})\frac{d\text{vec}\mathbf{S}}{d\boldsymbol{\theta}^{\mathsf{T}}} = (\mathbf{T} \otimes \mathbf{I}_s)H[\text{vec}\mathbf{A}; \boldsymbol{\theta}] \qquad \text{eqn 35}$$

$$\frac{d\text{vec}\mathbf{T}}{d\boldsymbol{\theta}^{\mathsf{T}}} = H[\lambda; \text{vec}\mathbf{A}]\mathbf{S}. \qquad \text{eqn 36}$$

The expression (34) is of the form

$$d^2\lambda = d\boldsymbol{\theta}^{\mathsf{T}}\mathbf{B}d\boldsymbol{\theta} \qquad \text{eqn 37}$$

and hence by the second identification theorem (11),

$$H[\lambda; \boldsymbol{\theta}] = \frac{1}{2}(\mathbf{B} + \mathbf{B}^{\mathsf{T}}) \qquad \text{eqn 38a}$$

where

$$\mathbf{B} = (\mathbf{T} \otimes \mathbf{I}_s)H[\text{vec}\mathbf{A}; \boldsymbol{\theta}] + (\mathbf{S}^{\mathsf{T}})H[\lambda; \text{vec}\mathbf{A}]\mathbf{S}$$
$$= (\mathbf{w}^{\mathsf{T}} \otimes \mathbf{v}^{\mathsf{T}} \otimes \mathbf{I}_s)H[\text{vec}\mathbf{A}; \boldsymbol{\theta}] + \left(\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}\right)^{\mathsf{T}}H[\lambda; \text{vec}\mathbf{A}]\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}.$$
$$\text{eqn 38b}$$

The first and second derivatives of $\mathbf{A}$ with respect to $\boldsymbol{\theta}$, which appear in $\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}$ and $H[\text{vec}\mathbf{A}; \boldsymbol{\theta}]$, respectively, can be evaluated by hand or by using a symbolic math program. This result is in agreement with the Hessian chain rule derived in a different way by Magnus & Neudecker (1999, p. 125).

These results can be used to parameterize constraints or covariation among traits. As a simple example, suppose that survival and fertility are constrained to covary as $F_i = cP_i$, and one wants the total second derivative including this constraint. This is obtained by defining a lower-level parameter $\boldsymbol{\theta}$, setting $F_i = \theta$ and $P_i = c\theta$ and calculating $H[\lambda; \boldsymbol{\theta}]$.

SECOND DERIVATIVES OF THE INVASION EXPONENT r: $H[r; \text{VEC}\mathbf{A}]$ AND $H[r; \theta]$

The population growth rate in continuous time is the invasion exponent $r = \log \lambda$. By the definition of the Hessian in (7), the Hessian of r with respect to $\mathbf{A}$ is

$$H[r; \text{vec}\mathbf{A}] = \frac{d}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\text{vec}\left[\left(\frac{d\log\lambda}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\right)^{\mathsf{T}}\right]. \qquad \text{eqn 39}$$

We insert the first derivative of $\log \lambda$,

$$H[r; \text{vec}\mathbf{A}] = \frac{d}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\left[\frac{1}{\lambda}\text{vec}\left(\frac{d\lambda}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\right)^{\mathsf{T}}\right] \qquad \text{eqn 40}$$

and then apply the product rule to obtain

$$H[r; \text{vec}\mathbf{A}] = \left(\frac{d\lambda}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\right)^{\mathsf{T}}\frac{d}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\frac{1}{\lambda}$$
$$+ \frac{1}{\lambda}\frac{d}{d\text{vec}^{\mathsf{T}}\mathbf{A}}\text{vec}\left[\left(\frac{d\lambda}{\partial\text{vec}^{\mathsf{T}}\mathbf{A}}\right)^{\mathsf{T}}\right] \qquad \text{eqn 41}$$

which simplifies to

$$H[r; \text{vec}\mathbf{A}] = -\frac{1}{\lambda^2}\left(\frac{d\lambda}{\partial\text{vec}^{\mathsf{T}}\mathbf{A}}\right)^{\mathsf{T}}\frac{d\lambda}{d\text{vec}^{\mathsf{T}}\mathbf{A}} + \frac{1}{\lambda}H[\lambda; \text{vec}\mathbf{A}]$$
$$\text{eqn 42}$$

$$= -\frac{1}{\lambda^2}(\mathbf{w}\mathbf{w}^{\mathsf{T}} \otimes \mathbf{v}\mathbf{v}^{\mathsf{T}}) + \frac{1}{\lambda}H[\lambda; \text{vec}\mathbf{A}] \qquad \text{eqn 43}$$

where $H[\lambda; \text{vec}\mathbf{A}]$ is given by (25).

Replacing vec$\mathbf{A}$ in (42) with a parameter vector $\boldsymbol{\theta}$ gives the Hessian

$$H[r; \boldsymbol{\theta}] = -\frac{1}{\lambda^2}\left(\frac{d\lambda}{d\boldsymbol{\theta}^{\mathsf{T}}}\right)^{\mathsf{T}}\frac{d\lambda}{d\boldsymbol{\theta}^{\mathsf{T}}} + \frac{1}{\lambda}H[\lambda; \boldsymbol{\theta}] \qquad \text{eqn 44}$$

$$= -\frac{1}{\lambda^2}\left(\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}}\right)^{\mathsf{T}}(\mathbf{w}\mathbf{w}^{\mathsf{T}} \otimes \mathbf{v}\mathbf{v}^{\mathsf{T}})\frac{d\text{vec}\mathbf{A}}{d\boldsymbol{\theta}^{\mathsf{T}}} + \frac{1}{\lambda}H[\lambda; \boldsymbol{\theta}].$$
$$\text{eqn 45}$$

The derivatives $\frac{d\text{vec}\mathbf{A}}{\partial\boldsymbol{\theta}^{\mathsf{T}}}$ can be calculated by hand or with a symbolic math program, and $H[\lambda; \boldsymbol{\theta}]$ can be obtained from (38).

SECOND DERIVATIVES OF THE NET REPRODUCTIVE RATE R_0

The net reproductive rate R_0 measures the population growth rate per generation and is used as an alternative fitness measure to r under some special conditions (Pásztor, Meszéna & Kidsi

1996, Brommer 2000). If **A** is decomposed into transition and fertility matrices, **A** = **U**+**F**, then R_0 is the dominant eigenvalue of the next generation matrix **R** = **FN** (Cushing & Zhou 1994), where **N** is the fundamental matrix:

$$\mathbf{N} = (\mathbf{I}_n - \mathbf{U})^{-1} \qquad \text{eqn 46}$$

The (i,j) entry of **N** gives the expected number of visits to stage i for an individual starting in stage j. The (i,j) entry of **R** gives the expected lifetime production of stage i offspring by an individual starting in stage j.

Because R_0 is an eigenvalue, our results for $H[\lambda;\text{vec}\mathbf{A}]$ and $H[\lambda;\theta]$ can be applied to find its second derivatives, but with **R** taking the place of matrix **A**. The resulting expressions are more complicated than the corresponding expressions for λ, because parameters can affect R_0 through **U**, **F** or both. In the important special case where only a single type of offspring is produced (suppose it is numbered as stage 1), then **R** is an upper triangular matrix and R_0 is its (1,1) entry; in this case, eigenvalue calculations are not necessary.

R with respect to **U**. The first derivatives are obtained by differentiating **R** = **FN**, applying the vec operator and noting that $\text{dvec}\mathbf{N} = (\mathbf{N}^\mathsf{T} \otimes \mathbf{N})\text{dvec}\mathbf{U}$ (Caswell 2006, 2009), to obtain

$$\frac{\text{dvec}\mathbf{R}}{\text{dvec}^\mathsf{T}\mathbf{U}} = (\mathbf{N}^\mathsf{T} \otimes \mathbf{R}). \qquad \text{eqn 48}$$

The second derivatives of **R** are obtained from the definition of the Hessian matrix (9):

$$H[\text{vec}\mathbf{R}; \text{vec}\mathbf{U}] = \frac{\text{d}}{\text{dvec}^\mathsf{T}\mathbf{U}} \text{vec}\left[\left(\frac{\text{dvec}\mathbf{R}}{\text{dvec}^\mathsf{T}\mathbf{U}}\right)^\mathsf{T}\right] \qquad \text{eqn 49}$$

$$= \frac{\text{d}}{\text{dvec}^\mathsf{T}\mathbf{U}} \text{vec}[\mathbf{N} \otimes \mathbf{R}^\mathsf{T}]. \qquad \text{eqn 50}$$

The derivative of vec $(\mathbf{N}\otimes\mathbf{R}^\mathsf{T})$ is given by a result of Magnus & Neudecker (1985, Theorem 11; 1999, p. 209); for a $m\times n$ matrix **X** and a $p\times q$ matrix **Y**,

$$\text{dvec}(\mathbf{X} \otimes \mathbf{Y}) = (\mathbf{I}_n \otimes \mathbf{K}_{qm} \otimes \mathbf{I}_p)[(\mathbf{I}_{mn} \otimes \text{vec}\mathbf{Y})\text{dvec}\mathbf{X} \\ + (\text{vec}\mathbf{X} \otimes \mathbf{I}_{pq})\text{dvec}\mathbf{Y}] \qquad \text{eqn 51}$$

Thus, (50) can be rewritten as

$$H[\text{vec}\mathbf{R}; \text{vec}\mathbf{U}] = (\mathbf{I}_n \otimes \mathbf{K}_{n,n} \otimes \mathbf{I}_n)\left[(\mathbf{I}_{n^2} \otimes \text{vec}[\mathbf{R}^\mathsf{T}])\frac{\text{dvec}\mathbf{N}}{\text{dvec}^\mathsf{T}\mathbf{U}} + (\text{vec}\mathbf{N} \otimes \mathbf{I}_{n^2})\frac{\text{dvec}\mathbf{R}^\mathsf{T}}{\text{dvec}^\mathsf{T}\mathbf{U}}\right] \qquad \text{eqn 52}$$

$$= (\mathbf{I}_n \otimes \mathbf{K}_{n,n} \otimes \mathbf{I}_n)[(\mathbf{I}_{n^2} \otimes \text{vec}[\mathbf{R}^\mathsf{T}])(\mathbf{N}^\mathsf{T} \otimes \mathbf{N}) + (\text{vec}\mathbf{N} \otimes \mathbf{I}_{n^2})\mathbf{K}_{n,n}(\mathbf{N}^\mathsf{T} \otimes \mathbf{R})]. \qquad \text{eqn 53}$$

We defer the fully general calculation of $H[R_0;\theta]$ to Appendix S1-C and show results here for two useful special cases: the second derivatives with respect to the entries of the transition matrix **U** and with respect to the entries of the fertility matrix **F**. We consider both single and multiple types of offspring.

If we apply (38) to the case of R_0, replacing vec**A** with vec**R**, we obtain

$$H[R_0; \theta] = \frac{1}{2}(\mathbf{B} + \mathbf{B}^\mathsf{T}) \qquad \text{eqn 47a}$$

where

$$\mathbf{B} = (\mathbf{w}_R^\mathsf{T} \otimes \mathbf{v}_R^\mathsf{T} \otimes \mathbf{I}_s)H[\text{vec}\mathbf{R}; \theta] \\ + \left(\frac{\text{dvec}\mathbf{R}}{\text{d}\theta^\mathsf{T}}\right)^\mathsf{T} H[R_0; \text{vec}\mathbf{R}]\frac{\text{dvec}\mathbf{R}}{\text{d}\theta^\mathsf{T}} \qquad \text{eqn 47b}$$

and $\mathbf{w}_R$ and $\mathbf{v}_R$ are the right and left eigenvectors of **R**.

To evaluate (47), we must calculate the second derivatives of R_0 with respect to **R**, and the first and second derivatives of **R** with respect to θ. For the former, the Hessian $H[R_0;\text{vec}\mathbf{R}]$ is given by (25), using the dominant eigenvalues and eigenvectors of **R** rather than those of **A**. For the latter, we will consider the derivatives of **R** with respect to **U** and **F** in turn. The derivatives of **R** with respect to general parameters θ are shown in Appendix S1-B.

Second derivatives of R_0 to the transition matrix: $H[R_0; \text{vec}\mathbf{U}]$

The second derivatives of R_0 with respect to the entries of the transition matrix **U** require the first and second derivatives of

As a result,

$$H[R_0; \text{vec}\mathbf{U}] = \frac{1}{2}(\mathbf{B} + \mathbf{B}^\mathsf{T}) \qquad \text{eqn 54a}$$

where

$$\mathbf{B} = (\mathbf{w}_R^\mathsf{T} \otimes \mathbf{v}_R^\mathsf{T} \otimes \mathbf{I}_{n^2})H[\text{vec}\mathbf{R}; \text{vec}\mathbf{U}] \\ + \left(\frac{\text{dvec}\mathbf{R}}{\text{vec}^\mathsf{T}\mathbf{U}}\right)^\mathsf{T} H[R_0; \text{vec}\mathbf{R}]\frac{\text{dvec}\mathbf{R}}{\text{vec}^\mathsf{T}\mathbf{U}} \qquad \text{eqn 54b}$$

where $\frac{\text{dvec}\mathbf{R}}{\text{dvec}^\mathsf{T}\mathbf{U}}$ is given by (48) and $H[\text{vec}\mathbf{R}; \text{vec}\mathbf{U}]$ is given by (53).

Second derivatives of R_0 to the fertility matrix: $H[R_0; \text{vec}\mathbf{F}]$

Now consider the second derivatives of R_0 with respect to the entries of the fertility matrix **F**. Differentiating **R** = **FN** with respect to **F** yields the first derivatives

$$\frac{\text{dvec}\mathbf{R}}{\text{dvec}^\mathsf{T}\mathbf{F}} = (\mathbf{N}^\mathsf{T} \otimes \mathbf{I}_n). \qquad \text{eqn 55}$$

The second derivatives of **R** are given by the Hessian matrix

$$H[\text{vec}\mathbf{R}; \text{vec}\mathbf{F}] = \frac{\text{d}}{\text{dvec}^\mathsf{T}\mathbf{F}} \text{vec}[\mathbf{N} \otimes \mathbf{I}_n]. \qquad \text{eqn 56}$$

However, because **N** depends only on **U**, and not on **F**, this is a zero matrix.

Substituting (55) and (56) into (47) gives

$$H[R_0; \text{vec}\mathbf{F}] = \frac{1}{2}(\mathbf{B} + \mathbf{B}^\mathsf{T}) \qquad \text{eqn 57a}$$

where

$$\mathbf{B} = (\mathbf{N} \otimes \mathbf{I}_n) H[R_0; \text{vec}\mathbf{R}](\mathbf{N}^\mathsf{T} \otimes \mathbf{I}_n). \qquad \text{eqn 57b}$$

Single type of offspring

In the common case where there is only one type of offspring (Appendix S1-B), $H[R_0; \boldsymbol{\theta}]$ simplifies to

$$H[R_0; \boldsymbol{\theta}] = (\mathbf{e}_1^\mathsf{T} \otimes \mathbf{e}_1^\mathsf{T} \otimes \mathbf{I}_s) H[\text{vec}\mathbf{R}; \boldsymbol{\theta}] \qquad \text{eqn 58}$$

where $\mathbf{e}_1$ is the $n \times 1$ vector with 1 as its first entry and zeros elsewhere.

A case study: *Calathea ovandensis*

Calathea ovandensis is a neotropical perennial herb that inhabits forest understories. Horvitz & Schemske (1995) developed a stage-structured matrix model for *C. ovandensis* that contains eight stages distinguished by size and reproductive ability: seeds, nonreproductive stages (seedlings, juveniles, pre-reproductive), and reproductive stages (small, medium, large and extra large). Plants may grow larger, remain in the same size class, or shrink at each time step; larger adults are typically more fecund.

Horvitz and Schemske summarized four years of population dynamics from four plots of *C. ovandensis* with a series of 8×8 projection matrices. The average of these matrices, weighted by the observed stage abundances and transition frequencies, is given in Table 8 of Horvitz & Schemske (1995) as

ple matrix entries. To analyse these using (38), write the survival probabilities in a vector $\boldsymbol{\sigma}$, which is given by the column sums of $\mathbf{U}$, so that

$$\mathbf{U} = \mathbf{G}\text{diag}(\boldsymbol{\sigma}) \qquad \text{eqn 60}$$

where $\mathbf{G}$ describes stage transitions conditional on survival (Caswell 2011). The Hessian of λ with respect to $\boldsymbol{\sigma}$ is given by (38), with the parameter vector $\boldsymbol{\theta}$ given by $\boldsymbol{\sigma}$. Calculating this Hessian requires the first and second derivatives of $\mathbf{A}$ with respect to $\boldsymbol{\sigma}$. The first derivatives, assuming that $\mathbf{F}$ does not depend on $\boldsymbol{\sigma}$ (i.e. prebreeding census), are

$$\frac{\text{dvec}\mathbf{A}}{\text{d}\boldsymbol{\sigma}^\mathsf{T}} = (\mathbf{I}_n \otimes \mathbf{G}) \text{ diag } (\text{vec}\mathbf{I}_n)(\mathbf{e} \otimes \mathbf{I}_n) \qquad \text{eqn 61}$$

(see Caswell and Salguero-Gómez 2013, Appendix A).

The second derivatives of $\mathbf{A}$ are given by $H[\text{vec}\mathbf{A}; \boldsymbol{\sigma}]$, the derivative of (61) with respect to $\boldsymbol{\sigma}$. However, none of the terms in (61) depend on $\boldsymbol{\sigma}$, so $H[\text{vec}\mathbf{A}; \boldsymbol{\sigma}]$ is a zero matrix. Thus, the matrix $\mathbf{B}$ in (38) reduces to

$$\mathbf{B} = \left(\frac{\text{dvec}\mathbf{A}}{\text{d}\boldsymbol{\sigma}^\mathsf{T}}\right)^\mathsf{T} H[\lambda; \text{vec}\mathbf{A}] \frac{\text{dvec}\mathbf{A}}{\text{d}\boldsymbol{\sigma}^\mathsf{T}} \qquad \text{eqn 62}$$

where $\frac{\text{dvec}\mathbf{A}}{\text{d}\boldsymbol{\sigma}^\mathsf{T}}$ is given by (61) and $H[\lambda; \text{vec}\mathbf{A}]$ is given by (25).

The resulting Hessian matrix with respect to the lower-level survival probabilities, $H[\lambda; \boldsymbol{\sigma}]$, is shown in Fig. 2. These second derivatives are generally of smaller magnitude than those of $H[\lambda; \text{vec}\mathbf{A}]$ (Fig. 1). The largest second derivatives in $H[\lambda; \boldsymbol{\sigma}]$ appear in rows 1 and 2 (equivalently, columns 1 and 2). Figure

$$\mathbf{A} = \begin{pmatrix} 0.4983 & 0 & 0.5935 & 7.139 & 14.2715 & 24.6953 & 34.9027 & 40.5437 \\ 0.0973 & 0.0110 & 0.0191 & 0 & 0 & 0 & 0 & 0 \\ 0.0041 & 0.0442 & 0.3378 & 0.0698 & 0.0251 & 0.0065 & 0.0085 & 0 \\ 0 & 0.0014 & 0.1355 & 0.4286 & 0.1736 & 0.0968 & 0.0427 & 0.0435 \\ 0 & 0 & 0.0363 & 0.3841 & 0.6025 & 0.4258 & 0.2991 & 0.2174 \\ 0 & 0 & 0.0019 & 0.0254 & 0.113 & 0.2387 & 0.1709 & 0.2826 \\ 0 & 0 & 0 & 0.0095 & 0.0272 & 0.1548 & 0.3248 & 0.1957 \\ 0 & 0 & 0 & 0.0032 & 0.0063 & 0.0452 & 0.1282 & 0.2391 \end{pmatrix}. \qquad \text{eqn 59}$$

The dominant eigenvalue of this matrix is 0·9923, indicating a near-steady state population.

To obtain the second derivatives of λ to the entries of $\mathbf{A}$, we calculated the Hessian $H[\lambda; \text{vec}\mathbf{A}]$ using (25). It is a symmetric 64×64 matrix (Fig. 1). In this example, and in others with large projection matrices, $H[\lambda; \text{vec}\mathbf{A}]$ contains many entries and may be difficult to interpret, even when entries that are fixed at 0 are omitted. Most of the second derivatives here are small in magnitude (Fig. 1b) with the exception of a few entries, including the highly negative $\partial^2\lambda/\partial a_{3,1}^2 = -217.53$ and $\partial^2\lambda/\partial a_{3,1}\partial a_{4,2} = -75.64$, where $a_{3,1}$ is the transition probability from seed to juvenile and $a_{4,2}$ is the transition probability from seedling to pre-reproductive.

Using (38), we calculated the Hessian $H[\lambda; \boldsymbol{\theta}]$ for a set of lower-level parameters $\boldsymbol{\theta}$. For example, the stage-specific survival probabilities are lower-level parameters that affect multi-

3 highlights the mixed second derivatives $\partial^2\lambda/\partial\sigma_1\partial\sigma_i$ and $\partial^2\lambda/\partial\sigma_2\partial\sigma_i$, along with the pure second derivatives $\partial^2\lambda/\partial\sigma_i^2$.

C. ovandensis has several large second derivatives involving σ_1 and σ_2 (the first two rows or columns of Fig. 2, which are shown separately in Fig. 3b,c). As discussed in the section 'Second-order sensitivity analysis and growth rate estimation,' this indicates that the sensitivity of λ to stage 1 (seed) and stage 2 (seedling) survival will be especially responsive to changes in later survival. Similarly, the sensitivity of λ to later survival is especially responsive to changes in seed and seedling survival.

When interpreted in terms of selection gradients, recall from the section, 'Characterizing nonlinear selection processes' that selection on a single trait is concave, linear or convex if $\partial^2\lambda/\partial\theta^2$ is negative, zero or positive. Selection on two traits is negatively or positively correlational if $\partial^2\lambda/\partial\theta_1\partial\theta_2$ is negative or positive. *C. ovandensis* is experiencing nearly linear selection on survival
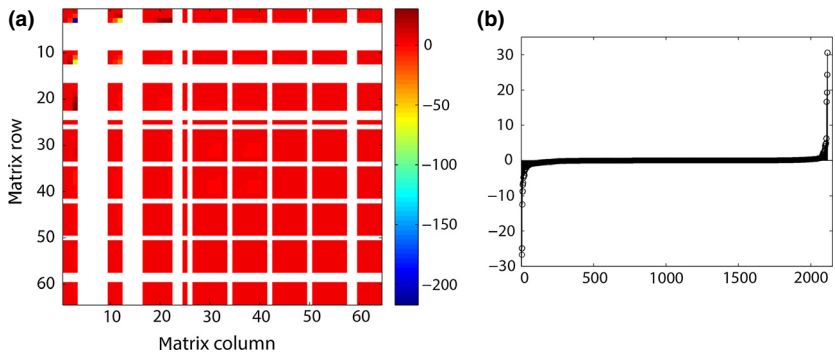
Fig. 1. (a) The Hessian matrix $H[\lambda; \text{vec}\mathbf{A}]$, giving the second derivatives of λ with respect to the entries of the projection matrix $\mathbf{A}$, for *C. ovandensis*. Entries corresponding to fixed zeros (unobserved transitions) in the matrix (59) are omitted. (b) The entries of the Hessian matrix in 1a, sorted in ascending order. The derivatives $\partial^2\lambda/\partial a_{3,1}^2 = -217{\cdot}53$ and $\partial^2\lambda/\partial a_{3,1}\partial a_{4,2} = -75{\cdot}64$ are omitted due to their magnitude.
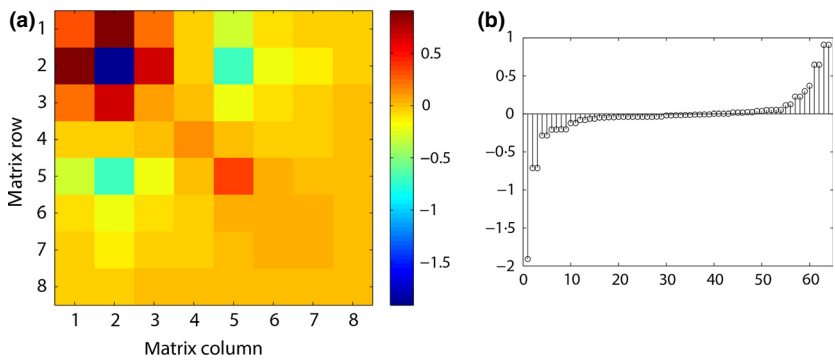
Fig. 2. (a) The Hessian $H[\lambda;\boldsymbol{\sigma}]$, giving the second derivatives of λ with respect to stage-specific survival probabilities σ_i, for *C. ovandensis*. (b) The Hessian entries in 2a, sorted in ascending order.
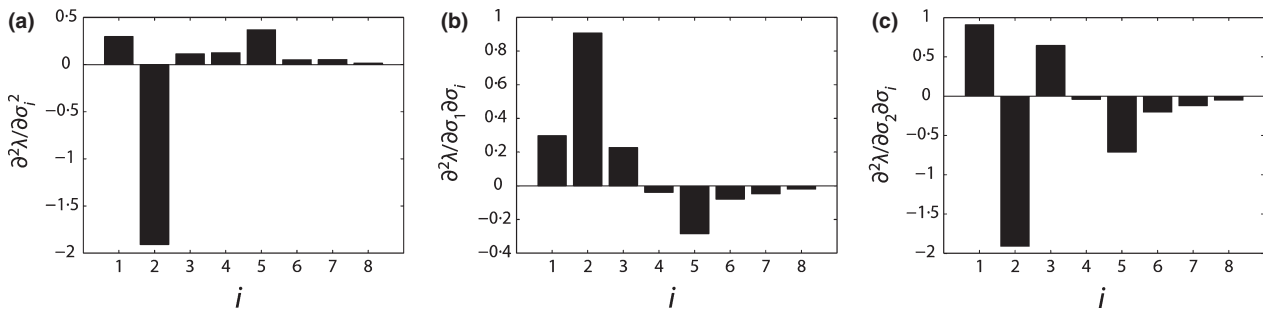

Fig. 3. Three sets of second derivatives from $H[\lambda; \boldsymbol{\sigma}]$ (Figure 2). (a) The pure second derivatives $\partial^2\lambda/\partial\sigma_i^2$. (b) The mixed second derivatives $\partial^2\lambda/\partial\sigma_1\partial\sigma_i$. (c) The mixed second derivatives $\partial^2\lambda/\partial\sigma_2\partial\sigma_i$.

in stage 8 ($\partial^2\lambda/\partial\sigma_8 \approx 0$), concave selection on survival in stage 2, and convex selection on survival in stages 1, 3, 4 and 5. There is negative correlational selection between survival in stage 1 (seeds) or 2 (seedlings) and survival in stages 4-8 (adults), and positive correlational selection between seed or seedling survival and survival in stages 1–3 (pre-adults). This indicates that seed and seedling survival are being selected to decrease their correlation with adult survival, but to increase their correlation with pre-adult survival.

Because the Hessian matrices include second derivatives with respect to all possible pairs of characters (matrix entries or lower-level parameters), they contain a great deal of information, and there are no established standards for displaying the results. We have shown several possibilities that may be useful: colour plots, plots that remove matrix entries that are of no interest because they are structural zeros, and plots displaying the range of magnitudes of the second derivatives. Others

will no doubt be developed. The MATLAB code used to generate the analysis is included in the Supporting Information.

Sensitivity analysis of stochastic growth rates

An application in which second derivatives are not the objective, but in which the Hessian matrix plays a role, is the sensitivity of Tuljapurkar's small-noise approximation to the stochastic growth rate $\log\lambda_s$ ('Section Sensitivity of the stochastic growth rate'). Tuljapurkar's approximation can be written in terms of the Jacobian matrix $\mathbf{D}$ of first derivatives of the dominant eigenvalue of the mean projection matrix, $\mathbf{D} = D[\lambda; \text{vec}\bar{\mathbf{A}}]$. Assuming that environments are uncorrelated in time,

$$\log\lambda_s \approx \log\lambda - \frac{\mathbf{DCD}^\mathsf{T}}{2\lambda^2} \qquad \text{eqn 63}$$

Table 2. An overview of the formulae for the second derivatives of population growth rates (λ, r, R_0) with respect to matrix entries (**A**, **U**, **F**), or to lower-level parameters ($\boldsymbol{\theta}$, $\boldsymbol{\sigma}$). The equation number for the corresponding Hessian matrix is given in the third column; auxiliary equations for terms in the Hessian expressions are given in the fourth column. The MATLAB functions used to calculate each Hessian, as provided in the supplemental material, are listed in the last column

Growth rate	Variables	Hessian equation	Auxiliary equations	MATLAB script
λ	**A**	(25)	(21), (22)	Hlambda_A.m
	$\boldsymbol{\theta}$	(38)	(25)	Hlambda_theta.m
	$\boldsymbol{\sigma}$	(62)	(25), (61)	Hlambda_sigma.m
r	**A**	(43)	(25)	Hr_A.m
	$\boldsymbol{\theta}$	(45)	(38)	Hr_theta.m
R_0	**U**	(54)	(25), (48), (53)	HR0_U.m
	F	(57)	(25)	HR0_F.m
	$\boldsymbol{\theta}$	(47)	(25), (B-2), (B-7) (or B-12/B-13), (B-8), (B-9), (B-10), (B-11)	HR0_theta.m
		(58) if one offspring type	(B-7) (or B-12/B-13), (B-8), (B-9), (B-10), (B-11)	HR0_theta_1.m

where **C** is the covariance matrix of the entries of the population projection matrix

$$\mathbf{C} = E\left[(\mathrm{vec}\mathbf{A}_t - \mathrm{vec}\bar{\mathbf{A}})(\mathrm{vec}\mathbf{A}_t - \mathrm{vec}\bar{\mathbf{A}})^{\mathsf{T}}\right]. \qquad \text{eqn 64}$$

The sensitivity of the stochastic growth rate can be obtained by differentiating (63) with respect to the entries of $\bar{\mathbf{A}}$. The sensitivity of the stochastic growth rate to $\bar{\mathbf{A}}$, leaving the variances and covariances fixed, depends on the second derivatives of λ as

$$\frac{\mathrm{d}\log\lambda_s}{\mathrm{d}\mathrm{vec}^{\mathsf{T}}\bar{\mathbf{A}}} = \frac{\mathbf{D}}{\lambda}\left(1 - \frac{\mathbf{CH}}{\lambda} + \frac{\mathbf{DCD}^{\mathsf{T}}}{\lambda^2}\right) \qquad \text{eqn 65}$$

where $\mathbf{H} = H[\lambda; \mathrm{vec}\bar{\mathbf{A}}]$ is the Hessian matrix of second derivatives. A derivation of (65) is provided in Appendix S1-C. Much more powerful and general approaches to sensitivity analysis of the stochastic growth rate are available in recent developments of the Monte Carlo method (e.g. Caswell, 2005, 2010, Tuljapurkar, Horvitz & Pascarella 2003, Haridas & Tuljapurkar 2005, Horvitz, Tuljapurkar & Pascarella 2005). This approximate result may, however, be useful in situations where the stochastic environment is defined directly in terms of the covariance matrix **C** of the vital rates.

Discussion

Although the first derivatives of population growth rates are commonly used in ecology and demography, tools for calculating the second derivatives are not nearly as well-established, even though second derivatives also have a variety of potential applications. To this end, we have derived new, more easily computable formulae for the second derivatives of three population growth rate measures – the discrete-time growth rate λ, the continuous-time growth rate r, and the per-generation growth rate R_0 – both with respect to projection matrix entries and to lower-level parameters. Table 2 provides an overview of the results, with directions to the equations defining the Hessian matrix, containing all second-order partial derivatives, for each type of growth rate and each type of independent variable.

The matrix calculus approach is comprehensive, and even though the formulae may appear complicated, they are easy to

apply with any matrix-oriented software. Other methods for finding second derivatives are either more limited or require more difficult and error-prone calculations. Cohen (1978), for instance, derives the second pure derivatives of λ with respect to the diagonal elements of the projection matrix ($\partial^2\lambda/\partial a_{ii}^2$) only. The approaches of Deutsch & Neumann (1984) and Kirkland & Neumann (1994) rely on the calculation of group inverses, while those of Caswell (1996) require all the eigenvalues and eigenvectors of the projection matrix. McCarthy *et al.*'s method (2008) uses transfer functions rather than eigenvectors and is more complicated when handling lower-level parameters.

Population growth rate, no matter how it is measured, is important in many ecological and evolutionary problems. It is hoped that the methods presented here will contribute to a deeper understanding of the response of growth rates to changes in parameters.

Acknowledgements

Data accessibility

MATLAB scripts: uploaded as online supporting information

References

Brodie, E.D., Moore, A.J. & Janzen, F.J. (1995) Visualizing and quantifying natural selection. *Trends in Ecology and Evolution*, **10**, 313–318.

Brommer, J.E. (2000) The evolution of fitness in life-history theory. *Biological Reviews*, **75**, 377–404.

Carslake, D., Townley, S. & Hodgson, D.J. (2008) Nonlinearity in eigenvalue-perturbation curves of simulated population projection matrices. *Theoretical Population Biology*, **73**, 498–505.

Caswell, H. (1996) Second derivatives of population growth rate: calculation and applications. *Ecology*, **77**, 870–879.

Caswell, H. (2001) *Matrix Population Models.* Sinauer Associates, Inc., Sunderland, MA.

Caswell, H. (2005) Sensitivity analysis of the stochastic growth rate: three extensions. *Australian and New Zealand Journal of Statistics*, **47**, 75–85.

Caswell, H. (2007) Sensitivity analysis of transient population dynamics. *Ecology Letters*, **10**, 1–15.

Caswell, H. (2008) Perturbation analysis of nonlinear matrix population models. *Demographic Research*, **18**, 59–116.

Caswell, H. (2009) Stage, age and individual stochasticity in demography. *Oikos*, **118**, 1763–1782.

Caswell, H. (2010) Life table response experiment analysis of the stochastic growth rate. *Journal of Ecology*, **98**, 324-333.

Caswell, H. (2011) Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theoretical Ecology*, **5**, 403–417.

Caswell, H. & Salgueró-Gomez, R. (2013) Age, stage and senescence in plants. *Journal of Ecology*, **101**, 585–595.

Cohen, J.E. (1978) Derivatives of the spectral radius as a function of non-negative matrix elements. *Mathematical Proceedings of the Cambridge Philosophical Society*, **83**, 183–190.

Cushing, J.M. & Zhou, Y. (1994) The net reproductive value and stability in matrix population models. *Natural Resources Modeling*, **8**, 297–333.

Deutsch, E. & Neumann, M. (1984) Derivatives of the Perron root of an essentially nonnegative matrix and the group inverse of an M-matrix. *Journal of Mathematical Analysis and Applications*, **102**, 1–29.

Diekmann, O. (2004) A beginner's guide to adaptive dynamics. *Banach Center Publications*, **63**, 47–86.

Doebeli, M. (2011) *Adaptive Diversification (MPB-48). Monographs in Population Biology*. Princeton University Press, Princeton.

Geritz, S.A.H., Kidsi, É., Meszéna, G. & Metz, J.A.J. (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, **12**, 35–57.

Haridas, C. & Tuljapurkar, S. (2005) Elasticities in variable environments: properties and implications. *American Naturalist*, **166**, 481–495.

Henderson, H.V. & Searle, S.R. (1981) The vec-permutation, the vec operator and Kronecker products: a review. *Linear Multilinear Algebra*, **9**, 271–288.

Horvitz, C.C. & Schemske, D.W. (1995) Spatiotemporal variation in demographic transitions of a tropical understory herb: projection matrix analysis. *Ecological Monographs*, **65**, 155–192.

Horvitz, C.C., Tuljapurkar, S. & Pascarella, J.B. (2005) Plant-animal interactions in random environments: habitat-stage elasticity, seed predators, and hurricanes. *Ecology*, **86**, 3312–3322.

Kirkland, S.J. & Neumann, M. (1994) Convexity and concavity of the Perron root and vector of Leslie matrices with applications to a population model. *SIAM Journal of Matrix Analysis and Applications*, **15**, 1092–1107.

Lande, R. (1982) A quantitative genetic theory of life history evolution. *Ecology*, **63**, 607–615.

Lande, R. & Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution*, **37**, 1210–1226.

Magnus, J.R. & Neudecker, H. (1979) The commutation matrix: some properties and applications. *Annals of Statistics*, **7**, 381–394.

Magnus, J.R. & Neudecker, H. (1985) Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. *Journal of Mathematical Psychology*, **29**, 474–492.

Magnus, J.R. & Neudecker, H. (1999) *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2nd edn. John Wiley and Sons Ltd, New York.

McCarthy, D., Townley, S. & Hodgson, D. (2008) On second order sensitivity for stage-based population projection matrix models. *Theoretical Population Biology*, **74**, 68–73.

Pásztor, L., Meszéna, G. & Kidsi, É. (1996) R_0 or r: a matter of taste? *Journal of Evolutionary Biology*, **9**, 511–518.

Phillips, P.C. & Arnold, S.J. (1989) Visualizing multivariate selection. *Evolution*, **43**, 1209–1222.

Roth, W.E. (1934) On direct product matrices. *Bulletin of the American Mathematical Society*, **40**, 461–468.

Stearns, S.C. (1992) *The Evolution of Life Histories*. Oxford University Press, Oxford, UK.

Tuljapurkar, S.D. (1982) Population dynamics in variable environments. III. Evolutionary dynamics of r-selection. *Theoretical Population Biology*, **21**, 141–165.

Tuljapurkar, S., Horvitz, C. & Pascarella, J.B. (2003) The many growth rates and elasticities of populations in random environments. *American Naturalist*, **162**, 489–502.

Waxman, D. & Gavrilets, S. (2005) 20 questions on adaptive dynamics. *Journal of Evolutionary Biology*, **18**, 1139–1154.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Supplemental files S1–17. MATLAB scripts for calculating second derivatives (as summarized in Table 2).

Appendix S1. A) First derivatives of **S** and **T**. B) Second derivatives of R0 to lower-level parameters: H[R$_0$;θ]. C) Sensitivity of the stochastic growth rate. D) Commutation matrix code.