demogR: A Package for the Construction and Analysis of Age-structured Demographic Models in R

James Holland Jones

Stanford University

Abstract

The analysis of matrix population models has become a fundamental tool in ecology, conservation biology, and life history theory. In this paper, I present **demogR**, a package for analyzing age-structured population models in R. The package includes tools for the construction and analysis of matrix population models. In addition to the standard analyses commonly used in evolutionary demography and conservation biology, **demogR** contains a variety of tools from classical demography. This includes the construction of period life tables, and the generation of model mortality and fertility schedules for human populations. The tools in **demogR** are generally applicable to age-structured populations but are particularly useful for analyzing problems in human ecology. I illustrate some of the capabilities of the package by doing an evolutionary demographic analysis of several human populations.

Keywords: demography, human ecology, life history theory, matrix population models, model life tables, R.

1. Introduction

demogR is a package for doing demographic analyses of age-structured populations in R (R Development Core Team 2007). By default, it works with the data format typically used by human demographers. However, all of the routines for the analysis of matrix population models will work for any age-structured life cycle. In addition to the functions particularly of interest to ecologists, there are a variety of routines for classical formal demography of human populations.

In this paper, I will focus primarily on describing the functions associated with matrix population models. Matrix population models have proved especially important in two complemen-

tary areas of ecology: (1) conservation biology and (2) evolutionary demography, including human evolutionary ecology (Hill and Hurtado 1996; Kaplan, Hill, Lancaster, and Hurtado 2000)

1.1. Conservation biology

The judicious use of matrix population models has become an indispensable tool for conservation biology (Lande 1988; Doak, Kareiva, and Klepetka 1994; Beissinger and Westphal 1998; Benton and Grant 1999). The rate of increase of a population is a critical parameter of interest in conservation biology, since a robust growth rate – and low variance in this growth rate – is the best insurance against extinction. The asymptotic growth rate of an age-structured population is the dominant eigenvalue of the projection matrix. This dominant eigenvalue, λ[1] is guaranteed to be unique, strictly greater than all other eigenvalues, and real if standard conditions for irreducibility and aperiodicity are met (Caswell 2001). Sensitivity and elasticity analyses of λ with respect to perturbations of vital events provide important insights into the relative importance of different vital rates for population viability and the potential impact of measurement error on the estimation of a population's rate of increase (Caswell 2001). The assumptions of the perturbation analysis that lie behind demographic elasticity analysis limit the uncritical use of elasticities in conservation biology. For example, real perturbations in a conservation context can be large and affect multiple vital rates simultaneously (Mills, Doak, and Wisdom 1999). In fact, a number of the routines in **demogR** are specifically designed for just these problems, including those for calculating the second derivatives of λ with respect to perturbations, and the sensitivities of eigenvalue elasticities.

1.2. Evolutionary demography

The asymptotic growth rate, λ, is also the fitness measure in a structured population with overlapping generations. Intense interest has focused on understanding how λ responds to perturbations in the elements of the life cycle (Caswell 1978; Benton and Grant 1996; Pfister 1998; de Kroon, van Groenendael, and Ehrlen 2000). Lande (1982) showed that the expected change in the mean life history phenotype (or any other phenotype) $\Delta \mathbf{z}$ is given by

$$\lambda \Delta \mathbf{z} = \mathbf{G} \nabla \lambda, \tag{1}$$

where $\nabla \lambda$ is the fitness gradient (which is simply a vector of the sensitivities of λ) and $\mathbf{G}$. is the additive genetic covariance matrix. The $\mathbf{G}$ matrix describes how a multivariate phenotype can respond to selection. Fisher (1958) showed that the rate of evolution is proportional to the variance in fitness. The $\mathbf{G}$ matrix generalizes this idea to multiple fitness components. These components all may have additive genetic variance, but they are also likely to covary with each other. In the simplest case where there are not (non-zero) off-diagonal elements and the diagonal elements of $\mathbf{G}$ are greater than zero, the life cycle traits will move in the direction of the gradient $\nabla \lambda$ which is the vector of sensitivities. When the covariances (i.e., off-diagonal elements) are non-zero, the response to selection will be more complex. The methods encompassed by matrix population models are therefore fundamental for understanding evolutionary change in life histories. There is currently some excellent work being done linking evolutionary quantitative genetics studies with ecology (e.g., Coulson, Benton,

---

[1]Wherever no confusion arises by doing so, I will simply write λ rather than λ_1 for the dominant eigenvalue.

Lundberg, Dall, and Kendall 2006; Pelletier, Clutton-Brock, Pemberton, Tuljapurkar, and Coulson 2007).

In addition, Hamilton's perturbation analysis of the discrete-time characteristic equation for the human life history provided the canonical theory of senescence for the past 40 years (Hamilton 1966; Baudisch 2005; Steinsaltz, Evans, and Wachter 2005). The force of selection against deleterious mutations with age-specific effects declines with age and this decline is measured by the sensitivities of λ. The demographic methods used by **demogR** are therefore of central importance to the evolution of senescence and life histories in general and such analyses are increasingly being carried out in an ecological context on wild populations (Charmantier and Garant 2005; Nussey, Clutton-Brock, Albon, Pemberton, and Kruuk 2005; Charmantier, Perrins, McCleery, and Sheldon 2006; Nussey, Kruuk, Donald, Fowlie, and Clutton-Brock 2006)

1.3. Human evolutionary ecology

The observation that the asymptotic rate of increase of the population projection matrix is the proper fitness measure for an age-structured population places demography at the center of studies of human evolutionary ecology. The human life cycle is highly age-structured and there is little value to studies of human adaptation that fail to account for the variable force of selection on traits with respect to their age expression. Work in human evolutionary ecology has tended to focus on the demography of small-scale populations such as hunter-gatherers or horticulturalists (e.g., Howell 1979; Early and Peters 1990; Hill and Hurtado 1996). Understanding the evolution of human life histories, characterized as they are by long lifespan, modest fertility, extensive biparental care, overlapping periods of juvenile dependency, and long post-reproductive lifespan, remains a major challenge in human evolutionary ecology. An important area where human evolutionary ecology and conservation biology intersect is the analysis of subsistence and resource management in small populations (e.g. Alvard 1998; Smith and Wishnie 2000; Bird, Bird, and Parker 2005). Humans often play an instrumental role in regulating (and disturbing) communities and ecosystem processes and understanding the role that changes in the size and composition of human populations play is essential.

1.4. The need for a package for demographic analysis

The methods of analysis for matrix population models are widely used and a number of excellent texts are available to describe their use. Caswell (2001) provides MATLAB code fragments and pseudo-code for many of the analyses described in his comprehensive book, while Morris and Doak (2002) provide extensive MATLAB code to support their text. However, these functions are written for fairly expensive proprietary software, limiting their availability to all users. This is particularly problematic for scientists in lesser developed countries. Having a single package that collects many of the most commonly used analyses for age-structured population models, written for freely available software and distributed under the GPL solves this problem. **demogR** also allows population biologists and demographers to enjoy the many other advantages of the R programming environment including a more natural environment for handling complex data, advanced tools for statistical inference and superior graphics.

2. Using demogR

2.1. Life table methods

I will first discuss the construction of life tables from the census counts and death tabulations. Throughout the rest of the paper, I will use the vital event data from three countries: Venezuela (1965), Madagascar (1966), USA (1967). These are the three populations used by Goodman, Keyfitz, and Pullum (1974) in their classic paper on the demographic theory of kinship and they illustrate differing combinations of extremely high and low fertility and mortality. The dataset `goodman` is included in the package and consists of raw numbers used to construct the life tables and age-specific fertility schedules analyzed in the Goodman *et al.* (1974) paper. This dataset is a 19×10 data frame with columns representing age followed by the mid-year population size ($_nK_x$), number of deaths ($_nD_x$) and births (B_x) for Venezuela, Madagascar, and the USA for ages $x \in 0, 1, 5, 10, \ldots 80, 85$. Following standard demographic notation, x is the age index and n indicates the width of the age interval. Thus, $_nK_x$ is the midyear population for individuals age x to $x + n$ years.

Construct a period life table with the routine `life.table`, which uses the Greville equation to convert observed death rates ($_nM_x = {_nD_x}/{_nK_x}$) into interval probabilities $_nq_x$. A period life table is a model for a hypothetical population. Through it, we apply period measures (i.e., cross-sectional measures taken at a specific time) and create a hypothetical cohort whose mortality experience we summarize. In effect, the period life table asks questions like: if current mortality conditions pertained to a particular population, what would the average age of death of that population be? If we had an actual cohort, then we would observe (cohort) age-specific death rates $_nm_x = {_nd_x}/{_nL_x}$, the observed number of deaths divided by the number of person-years in an age class. For the synthetic cohort of the period life table, we substitute the period death rates $_nM_x$, taking the mid-interval population size in age class x as our estimate for the number of person-years lived in the interval ($_nL_x$). As Greville showed, the only information needed to move between $_nM_x$ and $_nq_x$ is the average person-years lived by individuals dying in the interval $_na_x$ (Greville 1943). Given an estimate of the central death rate $_nM_x$, the conversion to $_nq_x$ is given by:

$$_nq_x = \frac{n \cdot {_nm_x}}{1 + {_nm_x}(n - {_na_x})}.$$

There are a variety of methods available for specifying the $_na_x$ schedule (Keyfitz 1977; Preston, Heuveline, and Guillot 2001). Probably the most common approach uses regression on the observed $_nM_x$ values. In terms of mortality, the most rapidly changing portion of the human life cycle are the ages under five, and especially from birth to the first birthday. This general pattern of high and declining early mortality probably pertains to most mammals (Charnov 1993). For ages above five, we can typically assume that deaths are randomly distributed across the 5-year age interval, so that the average number of person-years lived by individuals dying at ages $x > 5$ is $_na_x = 2.5$. Thus, the regression approach typically focuses on the first two $_na_x$ values (i.e., for age classes 0-1 and 1-4). The most common regression-based approach to graduating the $_na_x$ schedule is attributable to Keyfitz and Flieger (1990) and this is the default value for `life.table`.

In using regression-based graduation for the $_na_x$ schedules, human demographers capitalize on the remarkably robust nature of the human life cycle. Despite the great diversity of eco-

nomic, social, and demographic environments that people experience, there are many general features of the human mortality experience. For many ecological applications, the detailed demographic information that allows human demographers to specify an $_na_x$ schedule by regression or borrowing from a similar population may be absent. It is common in ecological applications to simply construct a cohort life table (Deevey 1947; Caughley 1966). The assumptions underlying a cohort life table are not unreasonable if we assume that a population has lived in a relatively constant (average) environment. In the presence of non-stationary demographic change (e.g., in the case of demographic transitions in recent human populations), the cohort and period life tables will give very different pictures of the mortality experience of the population. Thus, for most applications to human ecology, the options `type="cd"` or `type="kf"` should be used, while for most non-human ecological applications, `type="cohort"` is probably the best option.

First, load the data and then calculate a period life tables using the default Keyfitz-Flieger graduation for the three populations. For the purposes of this paper, I will also limit the number of significant digits that are printed.

```
R> library("demogR")
R> options(digits = 3, scipen = 1)
R> data("goodman")
R> names(goodman)

 [1] "age"     "ven.nKx" "ven.nDx" "ven.bx"  "mad.nKx" "mad.nDx" "mad.bx"
 [8] "usa.nKx" "usa.nDx" "usa.bx"


R> ven <- with(goodman, life.table(x = age, nKx = ven.nKx, nDx = ven.nDx))
R> mad <- with(goodman, life.table(x = age, nKx = mad.nKx, nDx = mad.nDx))
R> usa <- with(goodman, life.table(x = age, nKx = usa.nKx, nDx = usa.nDx))
R> ven


    x   nax    nMx    nqx     lx    ndx    nLx     Tx     ex
1   0 0.148 0.0461 0.0443 1.000 0.0443 0.962 67.70  67.70
2   1 1.500 0.0057 0.0226 0.956 0.0216 3.769 66.74  69.84
3   5 2.500 0.0012 0.0058 0.934 0.0055 4.657 62.97  67.41
4  10 2.500 0.0007 0.0036 0.929 0.0034 4.635 58.31  62.80
5  15 2.500 0.0011 0.0054 0.925 0.0050 4.614 53.68  58.01
6  20 2.500 0.0013 0.0065 0.920 0.0060 4.586 49.07  53.32
7  25 2.500 0.0017 0.0086 0.914 0.0079 4.552 44.48  48.65
8  30 2.500 0.0022 0.0109 0.906 0.0098 4.507 39.93  44.05
9  35 2.500 0.0033 0.0163 0.897 0.0146 4.446 35.42  39.51
10 40 2.500 0.0042 0.0209 0.882 0.0184 4.364 30.97  35.12
11 45 2.500 0.0055 0.0273 0.864 0.0236 4.259 26.61  30.81
12 50 2.500 0.0083 0.0406 0.840 0.0341 4.115 22.35  26.61
13 55 2.500 0.0111 0.0542 0.806 0.0436 3.920 18.24  22.63
14 60 2.500 0.0231 0.1090 0.762 0.0831 3.603 14.32  18.78
15 65 2.500 0.0214 0.1015 0.679 0.0689 3.224 10.71  15.77
16 70 2.500 0.0338 0.1558 0.610 0.0951 2.814  7.49  12.27
```

```
17 75 2.500 0.0485 0.2164 0.515 0.1115 2.297  4.68  9.08
18 80 2.500 0.1137 0.4426 0.404 0.1787 1.572  2.38  5.89
19 85 3.587 0.2788 1.0000 0.225 0.2250 0.807  0.81  3.59
```

The life table contains nine columns: age (x), $_na_x$, the period central death rate $(_nM_x)$, the interval mortality probability $(_nq_x)$, survivorship (l_x), the proportion of deaths in the interval $(_nd_x)$, person-years lived in the interval $(_nL_x)$, person-years of remaining life in the cohort (T_x) and life expectancy (e_x). In most human demography – particularly for national-level data – the radix of the life table is typically set to $l_0 = 100,000$, leading the the interpretation of l_x as the number of survivors in the artificial cohort at age x and $_nd_x$ as the number of deaths in the interval. Following ecological and evolutionary conventions, the default of `life.table` is set for a radix of $l_0 = 1$, making l_x the probability of surviving to exact age x and $_nd_x$ the proportion of deaths between x and $x + n$ $(= {}_nm_xl_x)$.

`life.table` can also use the Coale-Demeny approach to graduating the $_na_x$ schedule, which can be more useful in anthropological applications (i.e., those of small populations, frequently experiencing high mortality) that are more likely to be used in ecological or evolutionary analyses. The Coale-Demeny approach uses different values in the regression depending on the severity of early mortality (Coale, Demeny, and Vaughn 1983). In practice, the two different methods produce very similar life tables for the three national-level populations.

To demonstrate the usage of the cohort life table, I use Caughley's survival data on Himalayan thar (*Hemitargus jemlahicus*). These data are included in the dataframe `thar`.

```
R> data("thar")
R> thar.lt <- with(thar, life.table(x = age, nDx = deaths, nKx = count,
+     type = "cohort", iwidth = 1, width12 = c(1, 1)))
R> thar.lt

    x nax    nMx    nqx     lx    ndx    nLx   Tx   ex
1   0 0.5 0.5317 0.5317 1.0000 0.5317 0.7341 3.49 3.49
2   1 0.5 0.0208 0.0208 0.4683 0.0098 0.4634 2.75 5.87
3   2 0.5 0.0532 0.0532 0.4585 0.0244 0.4463 2.29 4.99
4   3 0.5 0.1124 0.1124 0.4341 0.0488 0.4098 1.84 4.24
5   4 0.5 0.1392 0.1392 0.3854 0.0537 0.3585 1.43 3.72
6   5 0.5 0.1912 0.1912 0.3317 0.0634 0.3000 1.07 3.24
7   6 0.5 0.2182 0.2182 0.2683 0.0585 0.2390 0.77 2.88
8   7 0.5 0.2558 0.2558 0.2098 0.0537 0.1829 0.53 2.55
9   8 0.5 0.3125 0.3125 0.1561 0.0488 0.1317 0.35 2.25
10  9 0.5 0.3182 0.3182 0.1073 0.0341 0.0902 0.22 2.05
11 10 0.5 0.3333 0.3333 0.0732 0.0244 0.0610 0.13 1.77
12 11 0.5 0.4000 0.4000 0.0488 0.0195 0.0390 0.07 1.40
13 12 1.0 1.0000 1.0000 0.0293 0.0293 0.0293 0.03 1.00
```

2.2. Projection matrix construction

Having constructed the life tables, we are in a position to construct and analyze demographic projection matrices for the age-structured population. The routine `leslie.matrix` takes

as arguments a measure of age-specific survivorship (either l_x or $_nL_x$) and the age-specific fertility rates, m_x. `life.table` assumes that you want to construct a life table from data tabulated in a manner similar to `goodman`. That is, there are age-specific tabulations of mid-year populations size (or its equivalent), deaths, and births for ages $x = 0, 1, 5, 10, \ldots$

The inputs for the Leslie matrix are a measure of survivorship (either l_x or $_nL_x$) and age-specific fertility m_x. The default value is to use $_nL_x$, with corresponding argument `L=TRUE`. For national-level demographic data frequently employed in human demography, it is conventional to use person-years ($_nL_x$) to construct a projection matrix. This is because projections for national (or other large) populations are typically done with fairly coarse age intervals (e.g., 5 or 10 years). Since $_nL_x = \int_x^{x+n} l(x)dx$, $_nL_x$ will give better estimates of survival for such large age-classes. It is assumed that both input vectors will be of the same length and that they will both contain a value for one year-olds which will need to be removed from the m_x and l_x vectors when `L=FALSE` and will need to be combined with $_4L_1$ when `L=TRUE`.

Births are assumed to be continuously distributed throughout the year. That is, `leslie.matrix` assumes a birth-flow population. Following standard demographic conventions[2] (Keyfitz 1977; Preston *et al.* 2001), the Leslie matrix entry for the ith element of the first row is given by

$$F_i = {}_nL_0 \, \frac{m_i + P_i m_{i+1}}{2}. \tag{2}$$

When the option `L=FALSE` is used, $l(0.5) \cdot n$ is substituted for $_nL_x$ in equation 2, where $l(0.5) = \sqrt{l(2)}$ (Caswell 2001).

```
R> ven.mx <- with(goodman, ven.bx/ven.nKx)
R> mad.mx <- with(goodman, mad.bx/mad.nKx)
R> A <- leslie.matrix(lx = ven$nLx, mx = ven.mx)
R> B <- leslie.matrix(lx = mad$nLx, mx = mad.mx)
R> A
```

```
        [,1]    [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]   [,9]  [,10]     [,11]
 [1,] 0.000 0.00218 0.148 0.504 0.730 0.648 0.489 0.290 0.0957 0.0213 0.000937
 [2,] 0.984 0.00000 0.000 0.000 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [3,] 0.000 0.99525 0.000 0.000 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [4,] 0.000 0.00000 0.995 0.000 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [5,] 0.000 0.00000 0.000 0.994 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [6,] 0.000 0.00000 0.000 0.000 0.992 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [7,] 0.000 0.00000 0.000 0.000 0.000 0.990 0.000 0.000 0.0000 0.0000 0.000000
 [8,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.986 0.000 0.0000 0.0000 0.000000
 [9,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.000 0.981 0.0000 0.0000 0.000000
[10,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.000 0.000 0.9759 0.0000 0.000000
[11,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.000 0.000 0.0000 0.9661 0.000000
attr(,"class")
[1] "leslie.matrix"
```

---

[2]Actually most human demographers would also divide by $n \cdot l_0$, where l_0 is the radix of the life table. However, since I follow biological conventions by setting $l_0 = 1$, I ignore this.

If data are not in the format expected by `leslie.matrix`, the utility function `odiag` can facilitate constructing a Leslie matrix. `odiag` takes two arguments: `x` and `d`. For vector argument `x`, `odiag` produces a square matrix with `x` along the off-diagonal given by the integer `d`. For matrix argument `x`, `odiag` extracts as a vector the `d`th off-diagonal. To place values along the sub-diagonal of a matrix, use `d=-1`.

```
R> Px <- c(0.77, 0.95, 0.98, 0.97)
R> Fx <- c(0, 0, 1, 1.2, 1)
R> L <- odiag(Px, -1)
R> L[1, ] <- Fx
R> L


     [,1] [,2] [,3] [,4] [,5]
[1,] 0.00 0.00 1.00 1.20    1
[2,] 0.77 0.00 0.00 0.00    0
[3,] 0.00 0.95 0.00 0.00    0
[4,] 0.00 0.00 0.98 0.00    0
[5,] 0.00 0.00 0.00 0.97    0
```

2.3. Projection

Caswell (2001) notes that projection is the simplest form of analysis. Having constructed the Leslie matrix, and given a starting population vector, we can now project it forward some number of years and observe its behavior. Here we construct the Leslie matrix for the USA in 1967 and project it forward 100 years using the routine `project.leslie`, which takes as arguments the Leslie matrix, a starting population vector and a time horizon.

```
R> usa <- with(goodman, life.table(x = age, nKx = usa.nKx, nDx = usa.nDx))
R> usa.mx <- goodman$usa.bx/goodman$usa.nKx
R> C <- leslie.matrix(lx = usa$nLx, mx = usa.mx)
R> C


        [,1]    [,2]  [,3]  [,4]  [,5]  [,6]  [,7]   [,8]   [,9]    [,10]
 [1,] 0.000 0.00105 0.082 0.288 0.378 0.265 0.141 0.0586 0.0134 0.000333
 [2,] 0.997 0.00000 0.000 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [3,] 0.000 0.99836 0.000 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [4,] 0.000 0.00000 0.998 0.000 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [5,] 0.000 0.00000 0.000 0.997 0.000 0.000 0.000 0.0000 0.0000 0.000000
 [6,] 0.000 0.00000 0.000 0.000 0.996 0.000 0.000 0.0000 0.0000 0.000000
 [7,] 0.000 0.00000 0.000 0.000 0.000 0.995 0.000 0.0000 0.0000 0.000000
 [8,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.992 0.0000 0.0000 0.000000
 [9,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.000 0.9886 0.0000 0.000000
[10,] 0.000 0.00000 0.000 0.000 0.000 0.000 0.000 0.0000 0.9829 0.000000
attr(,"class")
[1] "leslie.matrix"
```

We now project the Leslie matrix with `project.leslie` using the census population of 1967 as the starting population size and plot the trajectory of the 10 age classes over the 100 year projection.

```
R> no <- goodman$usa.nKx[3:11]
R> no <- c(sum(goodman$usa.nKx[1:2]), no)
R> tmax <- 20
R> N <- project.leslie(A = C, no = no, tmax = tmax)

R> cols <- rgb(0, (10:1)/10, (1:10)/10)
R> plot(5 * (0:20), N[1, ]/100000, type = "l", xlab = "Years",
+      ylab = "Population Size (x100,000)",
+      ylim = c(16, 175), col = cols[1])
R> for (i in 2:10) lines(5 * (0:20), N[i, ]/100000, col = cols[i])
```

The call to `project.leslie` take `tmax=20` as an argument. However, since the width of the age classes in the Leslie matrix is five years, the projection actually spans 100 years, as shown
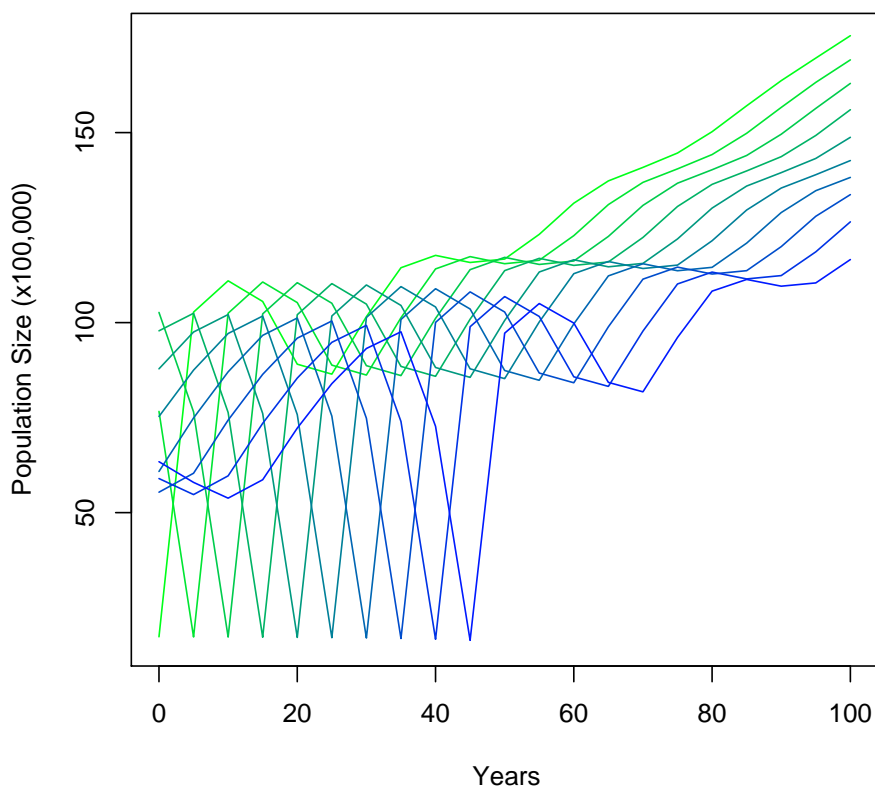

Figure 1: Convergence of the population to the stable age distribution.

in figure 1. The plot shows the very interesting transient dynamics which result from the fact that the baby boom had ended five years prior to the census and the peak of fertility occurred with the births of the cohort who were 10-14 years of age at the time of the census. The very low mortality of the United States meant that these unevenly sized cohorts passed through their life cycles essentially intact. The passage of the first cohort of newborns out of the range of ages projected by the Leslie matrix can be seen with the abrupt smoothing of the trajectories at $t = 50$. The actual population growth of the United States since 1965 looks quite different, of course. As a simple projection of a closed population, there is no account taken of immigration or changing vital rates.

When doing a projection, we sometimes do not care about the size of the different age classes and simply want the total population size. The option `pop.sum` has the default value of `pop.sum=FALSE`. Changing this to `pop.sum=TRUE` sums the population projection.

```
R> N.tot <- project.leslie(A = C, no = no, tmax = tmax, pop.sum = TRUE)
R> N.tot
```

```
 [1]  75717000  78261068  81865093  86923160  92733449  98000255 102053724
 [8] 105490420 108682801 111984424 116519384 121796188 126452846 130801781
[15] 135380074 140525030 146059631 151660059 157269825 163052060 169155329
```

2.4. Eigenvalue analysis

The most common forms of analysis of the projection matrix are contained in the function `eigen.analysis`, which takes the projection matrix as its sole argument. The return value of `eigen.analysis` is a list containing: (1) `lambda1`: the dominant eigenvalue λ of the projection matrix, (2) `rho`: the damping ratio $\rho = \lambda_1/|\lambda_2|$, which is a measure of the rate of convergence to the stable age distribution, (3) `sensitivities`: a matrix of eigenvalue sensitivities, (4) `elasticities`: a matrix of eigenvalue elasticities, (5) `stable.age`: the stable age distribution $\mathbf{u}$ normalized so that $\|\mathbf{u}\| = 1$, and (6) `repro.value`: reproductive value $\mathbf{v}$ normalized such that $v_1 = 1$.

```
R> Aea <- eigen.analysis(A)
R> Bea <- eigen.analysis(B)
R> Cea <- eigen.analysis(C)
R> Aea
```

```
$lambda1
[1] 1.21
```

```
$rho
[1] 3.01
```

```
$sensitivities
         [,1]  [,2]  [,3]  [,4]   [,5]   [,6]   [,7]    [,8]     [,9]    [,10]
 [1,] 0.000 0.154 0.127 0.104 0.0861 0.0708 0.0581 0.04749 0.038628 0.0312414
 [2,] 0.231 0.000 0.000 0.000 0.0000 0.0000 0.0000 0.00000 0.000000 0.0000000
```

```
 [3,] 0.000 0.228 0.000 0.000 0.0000 0.0000 0.0000 0.00000 0.000000 0.0000000
 [4,] 0.000 0.000 0.209 0.000 0.0000 0.0000 0.0000 0.00000 0.000000 0.0000000
 [5,] 0.000 0.000 0.000 0.156 0.0000 0.0000 0.0000 0.00000 0.000000 0.0000000
 [6,] 0.000 0.000 0.000 0.000 0.0932 0.0000 0.0000 0.00000 0.000000 0.0000000
 [7,] 0.000 0.000 0.000 0.000 0.0000 0.0471 0.0000 0.00000 0.000000 0.0000000
 [8,] 0.000 0.000 0.000 0.000 0.0000 0.0000 0.0184 0.00000 0.000000 0.0000000
 [9,] 0.000 0.000 0.000 0.000 0.0000 0.0000 0.0000 0.00447 0.000000 0.0000000
[10,] 0.000 0.000 0.000 0.000 0.0000 0.0000 0.0000 0.00000 0.000707 0.0000000
[11,] 0.000 0.000 0.000 0.000 0.0000 0.0000 0.0000 0.00000 0.000000 0.0000243
      [,11]
 [1,] 0.025
 [2,] 0.000
 [3,] 0.000
 [4,] 0.000
 [5,] 0.000
 [6,] 0.000
 [7,] 0.000
 [8,] 0.000
 [9,] 0.000
[10,] 0.000
[11,] 0.000
attr(,"class")
[1] "leslie.matrix"

$elasticities
        [,1]     [,2]    [,3]    [,4]    [,5]    [,6]    [,7]    [,8]     [,9]
 [1,] 0.000 0.000277 0.0156 0.0436 0.0521 0.0380 0.0236 0.01143 0.003064
 [2,] 0.188 0.000000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.000000
 [3,] 0.000 0.187927 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.000000
 [4,] 0.000 0.000000 0.1724 0.0000 0.0000 0.0000 0.0000 0.00000 0.000000
 [5,] 0.000 0.000000 0.0000 0.1287 0.0000 0.0000 0.0000 0.00000 0.000000
 [6,] 0.000 0.000000 0.0000 0.0000 0.0767 0.0000 0.0000 0.00000 0.000000
 [7,] 0.000 0.000000 0.0000 0.0000 0.0000 0.0386 0.0000 0.00000 0.000000
 [8,] 0.000 0.000000 0.0000 0.0000 0.0000 0.0000 0.0151 0.00000 0.000000
 [9,] 0.000 0.000000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00364 0.000000
[10,] 0.000 0.000000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.000572
[11,] 0.000 0.000000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.000000
          [,10]      [,11]
 [1,] 0.0005524 0.0000194
 [2,] 0.0000000 0.0000000
 [3,] 0.0000000 0.0000000
 [4,] 0.0000000 0.0000000
 [5,] 0.0000000 0.0000000
 [6,] 0.0000000 0.0000000
 [7,] 0.0000000 0.0000000
 [8,] 0.0000000 0.0000000
 [9,] 0.0000000 0.0000000
```

```
[10,] 0.0000000 0.0000000
[11,] 0.0000194 0.0000000
attr(,"class")
[1] "leslie.matrix"

$stable.age
 [1] 0.2023 0.1651 0.1361 0.1123 0.0925 0.0761 0.0625 0.0511 0.0415 0.0336
[11] 0.0269

$repro.value
 [1] 1.000000 1.225854 1.484045 1.649851 1.495850 1.082902 0.664755 0.317260
 [9] 0.094129 0.018304 0.000776
```

We see that the population of Venezuela in 1965 was increasing very rapidly: $r = \log(\lambda)/n =$ 0.0382, or nearly 4% annual increase. This is what happens when a population with a life expectancy at birth of nearly 70 also has a total fertility rate of of 6.4. Using the projection matrix $\mathbf{A}$, we can calculate both the net reproduction number R_0 and the generation time T. Following Caswell (2001), we can decompose $\mathbf{A}$ into two components, $\mathbf{A} = \mathbf{T} + \mathbf{F}$, where $\mathbf{T}$ is a matrix of transitions (simply the subdiagonal of an age-classified matrix) and $\mathbf{F}$ contains the fertilities. Using this decomposition, Caswell (2001) shows how to convert the projection matrix into a matrix of Markov transition probabilities for the states of the life cycle. Define the fundamental matrix of the Markov chain as $\mathbf{N} = (\mathbf{I} - \mathbf{T})^{-1}$, where $\mathbf{I}$ is an identity matrix of the same rank as $\mathbf{A}$. The fundamental matrix $\mathbf{N}$ contains the expected number of visits to a transient state i prior to absorption (i.e., death) conditional on starting state j and is an object of interest in its own right in addition to its utility for calculating other quantities. Caswell shows that R_0 is the dominant eigenvalue of the matrix $\mathbf{R} = \mathbf{FN}$. For the age-classified model, this is simply the r_{11}th entry of $\mathbf{R}$. The generation time is the amount of time it takes for a population growing at period rate λ to increase by a factor R_0: $\lambda^T = R_0$, and can be found easily by solving this equation for T.

```
R> calc.ro(A)


[1] 2.82


R> gen.time(A)


[1] 27.6
```

The average woman in Venezuela in 1965 replaced herself with almost three daughters and took approximately 27.6 years to do so.

The sensitivities of the growth rate, $\partial\lambda/\partial a_{ij}$, provide important demographic and evolutionary information. One possibility for calculating sensitivities is implicit differentiation of the characteristic equation (Hamilton 1966; Demetrius 1969). Caswell (1978) derived the sensitivities for a demographic projection matrix using a perturbation analysis of the matrix equation $\mathbf{Au} = \lambda\mathbf{u}$. The calculation of sensitivities in this way only requires suitably scaled dominant left and right eigenvectors of the projection matrix. Caswell shows that

$\partial \lambda / \partial a_{ij} = v_i u_j$, where $\langle \mathbf{v}, \mathbf{u} \rangle = 1$, $\mathbf{v}$ is the dominant left eigenvector and $\mathbf{u}$ is the dominant right eigenvector. Elasticities are simply proportional sensitivities $e_{ij} = (a_{ij}/\lambda)s_{ij}$ and have the convenient property that $\sum_{i,j} e_{ij} = 1$. We can compare the both the sensitivities and elasticities of the three populations in a convenient graphical manner using the plotting function `plot.leslie.matrix`. `plot.leslie.matrix` takes a projection matrix as its primary argument. It then plots the subdiagonal and first row of a square matrix on common axes. The figure uses the generic function `plot` to plot the fitness sensitivities (a-c) and elasticities (d-f) from the `leslie.matrix` objects.

```
R> par(mfrow = c(2, 3))
R> plot(Aea$sensitivities)
R> title("(a)")
R> plot(Bea$sensitivities)
R> title("(b)")
R> plot(Cea$sensitivities)
R> title("(c)")
R> plot(Aea$elasticities)
R> title("(d)")
R> plot(Bea$elasticities)
R> title("(e)")
R> plot(Cea$elasticities)
R> title("(f)")
```

The results are plotted in figure 2. Overall, sensitivities to pre-reproductive survival are highest, but there is a great deal of diversity in these three populations in the magnitude of the sensitivities to fertility. When the sensitivities are scaled to be proportional, we find that the general pattern is much more uniform across the three populations despite the tremendous diversity in vital rates.

As Stearns (1992) noted, sensitivities are "situational." As partial derivatives of the growth rate, they are conditional on all other values of the projection matrix remaining constant. However, there are at least three important contexts where the vital rates are likely to change. First in an evolutionary context, when the structure of $\mathbf{G}$, the additive genetic covariance matrix for the life cycle, permits it, the life cycle transitions will respond to selection and change. Second in a conservation context, we may undertake specific policies to improve the survival and/or fertility of specific age classes (Crouse, Crowder, and Caswell 1987; Doak *et al.* 1994; Morris, Tuljapurkar, Haridas, Menges, Horvitz, and Pfister 2006). A successful management intervention will change vital rates. Third in the case of culture change, vital rates may change in response to exogenous drivers such as economic development or endogenous drivers such as ideational change (Borgerhoff Mulder 1998).

Given this interest, we can ask the question how will a small change to some matrix element a_{kl} affect the other sensitivities? Sensitivities are partial derivatives. To investigate the change in sensitivities, we need to calculate eigenvalue second derivatives. (Caswell 1996) derived a formula for the second derivative of λ:

$$\frac{\partial^2 \lambda^{(1)}}{\partial a_{ij} \partial a_{kl}} = s_{il}^{(1)} \sum_{m \neq 1} \frac{s_{kj}^{(m)}}{\lambda^{(1)} - \lambda^{(m)}} + s_{kj}^{(1)} \sum_{m \neq 1} \frac{s_{il}^{(m)}}{\lambda^{(1)} - \lambda^{(m)}}. \tag{3}$$
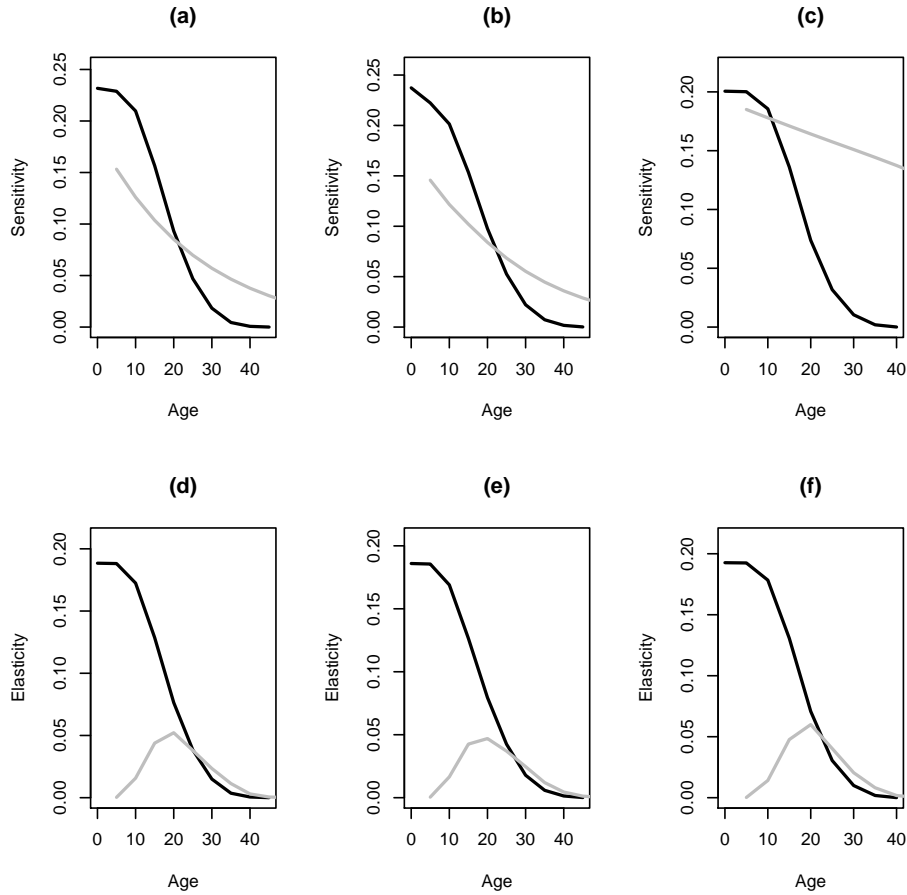
Figure 2: Sensitivities (a-c) and elasticities (d-f) for Venezuela (a,d), Madagascar (b, e), and the USA (c, f) in 1966. Sensitivities/elasticities with respect to survival in black and fertility in grey.

where the superscript (i) indicates which eigenvalue/eigenvector (assuming they have been sorted in descending order of the real parts of the eigenvalues).

Calculate the second derivatives of λ to a perturbation of the klth element of **A** using `secder`, which takes as arguments the projection matrix **A** followed by the row index k and column index l of the perturbation.

```
R> secder(A, k = 2, l = 1)

        [,1]    [,2]    [,3]    [,4]      [,5]      [,6]      [,7]      [,8]
[1,]   0.000  0.1380  0.0896  0.0540  0.028006  0.00950  -0.00331  -0.01179
[2,]  -0.173  0.0000  0.0000  0.0000  0.000000  0.00000   0.00000   0.00000
[3,]   0.000  0.0603  0.0000  0.0000  0.000000  0.00000   0.00000   0.00000
[4,]   0.000  0.0000  0.0469  0.0000  0.000000  0.00000   0.00000   0.00000
[5,]   0.000  0.0000  0.0000  0.0196  0.000000  0.00000   0.00000   0.00000
[6,]   0.000  0.0000  0.0000  0.0000 -0.000955  0.00000   0.00000   0.00000
```

```
 [7,]  0.000 0.0000 0.0000 0.0000  0.000000 -0.00718  0.00000  0.00000
 [8,]  0.000 0.0000 0.0000 0.0000  0.000000  0.00000 -0.00556  0.00000
 [9,]  0.000 0.0000 0.0000 0.0000  0.000000  0.00000  0.00000 -0.00210
[10,]  0.000 0.0000 0.0000 0.0000  0.000000  0.00000  0.00000  0.00000
[11,]  0.000 0.0000 0.0000 0.0000  0.000000  0.00000  0.00000  0.00000
            [,9]       [,10]     [,11]
 [1,] -0.016973 -0.0197005 -0.0206
 [2,]  0.000000  0.0000000  0.0000
 [3,]  0.000000  0.0000000  0.0000
 [4,]  0.000000  0.0000000  0.0000
 [5,]  0.000000  0.0000000  0.0000
 [6,]  0.000000  0.0000000  0.0000
 [7,]  0.000000  0.0000000  0.0000
 [8,]  0.000000  0.0000000  0.0000
 [9,]  0.000000  0.0000000  0.0000
[10,] -0.000450  0.0000000  0.0000
[11,]  0.000000 -0.0000199  0.0000
```

We can see that selection of the survival of 0-4 year-olds is highly concave, which is a multivariate generalization of stabilizing selection (Caswell 1996). Since the second derivative of λ with respect to P_1 is negative (a standard definition of a concave function), selection acting on P_1 will reduce variance in it (see also Arnold and Wade (1984) for more discussion of concavity vs. convex selection). In contrast, selection on fertility prior to the mean age of childbearing T is convex, which is the multivariate generalization of directional selection. Selection on early fertility will increase variance in fertility. The convexity of selection on survival of ages after 0-4 is more complex. There is a sign change in age 20-24, which happens to be the age-class with peak fertility.

We may also want to ask the question for elasticities. We use the eigenvalue second derivatives to calculate the sensitivities of the elasticities according to the relationship:

$$\frac{\partial e_{ij}}{\partial a_{kl}} = \frac{a_{ij}}{\lambda} \frac{\partial^2}{\partial a_{ij} \partial a_{kl}} - \frac{a_{ij}}{\lambda^2} \frac{\partial \lambda}{\partial a_{ij}} \frac{\partial \lambda}{\partial a_{kl}} + \frac{\delta_{ik}\delta_{jl}}{\lambda} \frac{\partial \lambda}{\partial a_{ij}} \tag{4}$$

To do this, we use the function `elassens`, which takes the same arguments as `secder`.

```
R> elassens(A, 2, 1)
```

```
        [,1]   [,2]   [,3]    [,4]    [,5]    [,6]    [,7]    [,8]    [,9]
 [1,] 0.000 0.0002 0.0080  0.0142  0.0070 -0.0022 -0.0058 -0.0050 -0.0019
 [2,] 0.014 0.0000 0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
 [3,] 0.000 0.0138 0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
 [4,] 0.000 0.0000 0.0057  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
 [5,] 0.000 0.0000 0.0000 -0.0085  0.0000  0.0000  0.0000  0.0000  0.0000
 [6,] 0.000 0.0000 0.0000  0.0000 -0.0154  0.0000  0.0000  0.0000  0.0000
 [7,] 0.000 0.0000 0.0000  0.0000  0.0000 -0.0133  0.0000  0.0000  0.0000
 [8,] 0.000 0.0000 0.0000  0.0000  0.0000  0.0000 -0.0074  0.0000  0.0000
 [9,] 0.000 0.0000 0.0000  0.0000  0.0000  0.0000  0.0000 -0.0024  0.0000
```

```
[10,] 0.000 0.0000 0.0000  0.0000  0.0000  0.0000  0.0000  0.0000 -0.0005
[11,] 0.000 0.0000 0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000
         [,10] [,11]
 [1,] -0.0005     0
 [2,]  0.0000     0
 [3,]  0.0000     0
 [4,]  0.0000     0
 [5,]  0.0000     0
 [6,]  0.0000     0
 [7,]  0.0000     0
 [8,]  0.0000     0
 [9,]  0.0000     0
[10,]  0.0000     0
[11,]  0.0000     0
attr(,"class")
[1] "leslie.matrix"
```

The function `fullsecder` computes all the second derivative of the dominant eigenvalue and returns them in a single matrix showing only the second derivatives with respect to non-zero matrix elements. The rows of this matrix, which will have a maximum size of $2k - 1 \times 2k - 1$ (where k is the rank of the Leslie matrix) are sorted such that all the fertilities are at the top. The matrix dimensions are labeled to facilitate easy access of the desired value of $\partial^2 \lambda / \partial a_{ij} \partial a_{kl}$. This application reveals a clear advantage to R over MATLAB, as such dimension labeling is not possible in MATLAB. This may seem, at first glance, like a small point but it greatly simplifies the interpretation of a very complex analysis. We can use this sorting to conveniently extract second derivatives for plotting. For instance, suppose that we want to know $\partial^2 \lambda / \partial P_i \partial P_j$ for Madagascar and plot the 10 lines.

```
R> age <- seq(0, 45, by = 5)
R> SD <- fullsecder(B)
R> SD
```

```
            12        13        14        15        16        17        18
12     0.055355  0.030603  0.012098 -0.001639 -0.011132 -0.017217 -0.020758
13     0.030603  0.012033 -0.001661 -0.011412 -0.017767 -0.021474 -0.023263
14     0.012098 -0.001661 -0.011625 -0.018314 -0.022280 -0.024195 -0.024681
15    -0.001639 -0.011412 -0.018314 -0.022546 -0.024645 -0.025201 -0.024734
16    -0.011132 -0.017767 -0.022280 -0.024645 -0.025367 -0.024958 -0.023845
17    -0.017217 -0.021474 -0.024195 -0.025201 -0.024958 -0.023903 -0.022389
18    -0.020758 -0.023263 -0.024681 -0.024734 -0.023845 -0.022389 -0.020655
19    -0.022443 -0.023683 -0.024176 -0.023585 -0.022291 -0.020614 -0.018789
110   -0.022776 -0.023125 -0.022980 -0.021978 -0.020458 -0.018693 -0.016873
111   -0.022164 -0.021907 -0.021342 -0.020103 -0.018490 -0.016730 -0.014979
21     0.147615  0.099651  0.063099  0.034432  0.013206 -0.001757 -0.011822
32    -0.016161  0.093492  0.059222  0.032342  0.012437 -0.001595 -0.011037
43    -0.020565 -0.039485  0.059147  0.033843  0.014978  0.001565 -0.007564
54    -0.026497 -0.039459 -0.048193  0.042872  0.025475  0.012690  0.003614
```

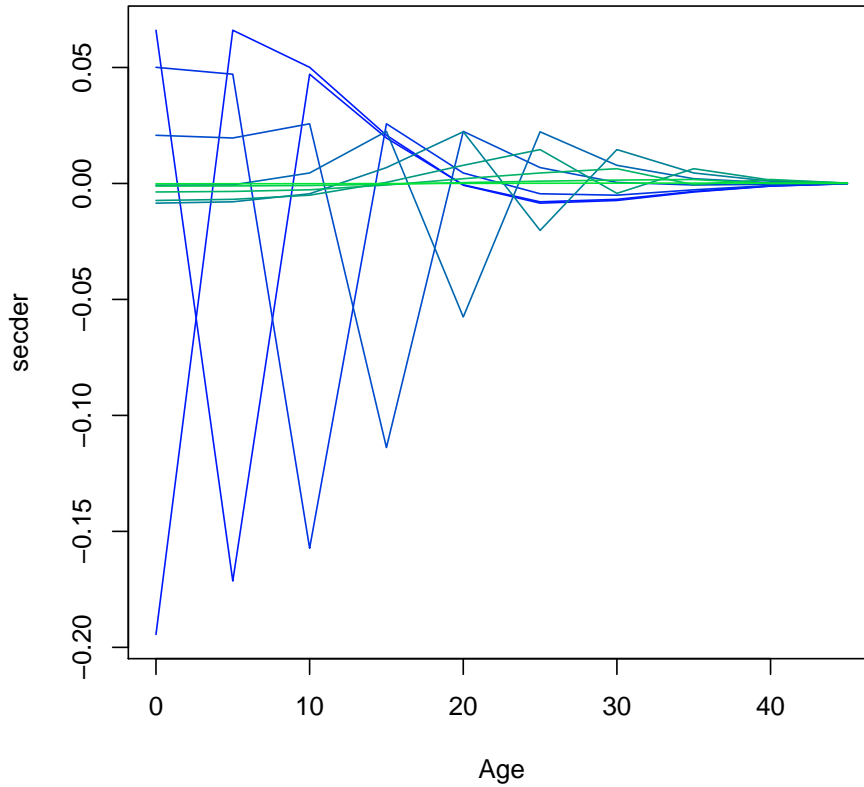Figure 3: Eigenvalue second derivatives with respect to survival P_j.

```
65   -0.025794 -0.032835 -0.037383 -0.039183  0.041099  0.028508  0.019034
76   -0.019221 -0.022289 -0.024134 -0.024513 -0.023859  0.043813  0.033603
87   -0.010647 -0.011592 -0.012078 -0.011955 -0.011421 -0.010651  0.044894
98   -0.004396 -0.004585 -0.004643 -0.004504 -0.004238 -0.003906 -0.003551
109  -0.001247 -0.001262 -0.001251 -0.001195 -0.001110 -0.001014 -0.000914
1110 -0.000147 -0.000145 -0.000141 -0.000133 -0.000122 -0.000111 -0.000099
             19         110        111         21         32         43          54
12   -0.0224429 -0.0227763 -0.02216   0.147615 -0.0161613 -0.020565 -0.0264967
13   -0.0236831 -0.0231249 -0.02191   0.099651  0.0934918 -0.039485 -0.0394591
14   -0.0241758 -0.0229797 -0.02134   0.063099  0.0592219  0.059147 -0.0481935
15   -0.0235847 -0.0219779 -0.02010   0.034432  0.0323419  0.033843  0.0428717
16   -0.0222907 -0.0204585 -0.01849   0.013206  0.0124374  0.014978  0.0254745
17   -0.0206140 -0.0186931 -0.01673  -0.001757 -0.0015953  0.001565  0.0126898
18   -0.0187895 -0.0168729 -0.01498  -0.011822 -0.0110372 -0.007564  0.0036143
19   -0.0169262 -0.0150771 -0.01329  -0.018204 -0.0170255 -0.013458 -0.0026217
110  -0.0150771 -0.0133389 -0.01169  -0.021836 -0.0204351 -0.016931 -0.0067081
111  -0.0132939 -0.0116931 -0.01020  -0.023454 -0.0219573 -0.018624 -0.0091835
```

```
21    -0.0182039 -0.0218355 -0.02345 -0.189792  0.0639678  0.048999  0.0209741
32    -0.0170255 -0.0204351 -0.02196  0.063968 -0.1672992  0.046083  0.0197817
43    -0.0134581 -0.0169314 -0.01862  0.048999  0.0460828 -0.154172  0.0254660
54    -0.0026217 -0.0067081 -0.00918  0.020974  0.0197817  0.025466 -0.1133197
65     0.0120088  0.0068747  0.00320 -0.000181 -0.0000881  0.004731  0.0219633
76     0.0255899  0.0193122  0.01443 -0.008181 -0.0076223 -0.004310  0.0066772
87     0.0359734  0.0287187  0.02284 -0.007321 -0.0068421 -0.005104  0.0003398
98     0.0417230  0.0338747  0.02741 -0.003763 -0.0035204 -0.002829 -0.0007518
109   -0.0008162  0.0361064  0.02937 -0.001210 -0.0011322 -0.000941 -0.0003847
1110  -0.0000879 -0.0000773  0.03004 -0.000155 -0.0001451 -0.000123 -0.0000607
               65         76         87         98        109       1110
12    -0.0257937 -0.0192207 -0.010647 -0.004396 -0.0012467 -1.47e-04
13    -0.0328350 -0.0222891 -0.011592 -0.004585 -0.0012618 -1.45e-04
14    -0.0373832 -0.0241336 -0.012078 -0.004643 -0.0012511 -1.41e-04
15    -0.0391834 -0.0245131 -0.011955 -0.004504 -0.0011945 -1.33e-04
16     0.0410993 -0.0238594 -0.011421 -0.004238 -0.0011105 -1.22e-04
17     0.0285079  0.0438130 -0.010651 -0.003906 -0.0010136 -1.11e-04
18     0.0190337  0.0336026  0.044894 -0.003551 -0.0009141 -9.90e-05
19     0.0120088  0.0255899  0.035973  0.041723 -0.0008162 -8.79e-05
110    0.0068747  0.0193122  0.028719  0.033875  0.0361064 -7.73e-05
111    0.0032045  0.0144251  0.022837  0.027409  0.0293686  3.00e-02
21    -0.0001805 -0.0081807 -0.007321 -0.003763 -0.0012097 -1.55e-04
32    -0.0000881 -0.0076223 -0.006842 -0.003520 -0.0011322 -1.45e-04
43     0.0047311 -0.0043102 -0.005104 -0.002829 -0.0009409 -1.23e-04
54     0.0219633  0.0066772  0.000340 -0.000752 -0.0003847 -6.07e-05
65    -0.0586080  0.0219958  0.007776  0.002039  0.0003533  2.12e-05
76     0.0219958 -0.0211701  0.014745  0.004620  0.0010283  9.53e-05
87     0.0077756  0.0147450 -0.004615  0.006587  0.0015383  1.51e-04
98     0.0020392  0.0046199  0.006587 -0.000601  0.0018176  1.81e-04
109    0.0003533  0.0010283  0.001538  0.001818 -0.0000390  1.94e-04
1110   0.0000212  0.0000953  0.000151  0.000181  0.0001941 -4.46e-07
```

```r
R> cols <- rgb(0, (1:10)/10, (10:1)/10)
R> plot(age, SD[11, 11:20], type = "l", col = cols[1], xlab = "Age",
+      ylab = "secder")
R> for (i in 12:20) lines(age, SD[i, 11:20], col = cols[i - 11])
```

The plot (figure 3) reveals that the value of $\partial^2 \lambda / \partial P_i^2$ is negative for all ages, indicating concave selection, though for ages greater than 25, it is very nearly zero. $\partial^2 \lambda / \partial P_i \partial P_j > 0$ for $j = i - 5$, indicating convex selection. This indicates that selection on survival at any age class will reduce variance in it. A detailed discussion of these results for human life cycles is beyond the scope of the present paper, but is taken up in a forthcoming paper (Jones, in prep).

It is a general feature of the human life cycle that the largest fraction of total selection occurs on pre-reproductive survival (Jones 2005). However, elasticities of any of the life cycle transitions are conditional upon the other values of the life cycle. Survival is important for fitness only to the extent that it leads eventually to reproduction. van Groenendael, De

Kroon, Kalisz, and Tuljapurkar (1994) presented a novel decomposition of the life-cycle using elasticities. They decompose the life cycle into its constituent loops (illustrated in figure 4). Using the properties of elasticities, they show that the total elasticity of the life cycle can be usefully decomposed into its loop elasticities. The algorithm presented by van Groenendael *et al.* (1994) is to identify all the loops of the life cycle, find the unique arc of each loop and multiply its elasticity, known as the loop's "characteristic elasticity," by the number of arcs in the loop. Wardle (1998) noted that identifying all the loops of a complex life cycle is not always a trivial task and presents a graph theoretic algorithm for enumerating all life cycle graphs. Fortunately, in the age classified life cycle, the unique arc of any loop will be the fertility arc of the loop. Thus the characteristic elasticity of any loop in an age-classified life cycle is simply the elasticity of the fertility arc in each loop.

With the function `loop.elas`, we can calculate and plot the loop elasticities of an age-classified projection matrix. `loop.elas` takes as arguments the projection matrix and the two optional arguments `draw.plot` (default=TRUE) and `peryear` (default=5).
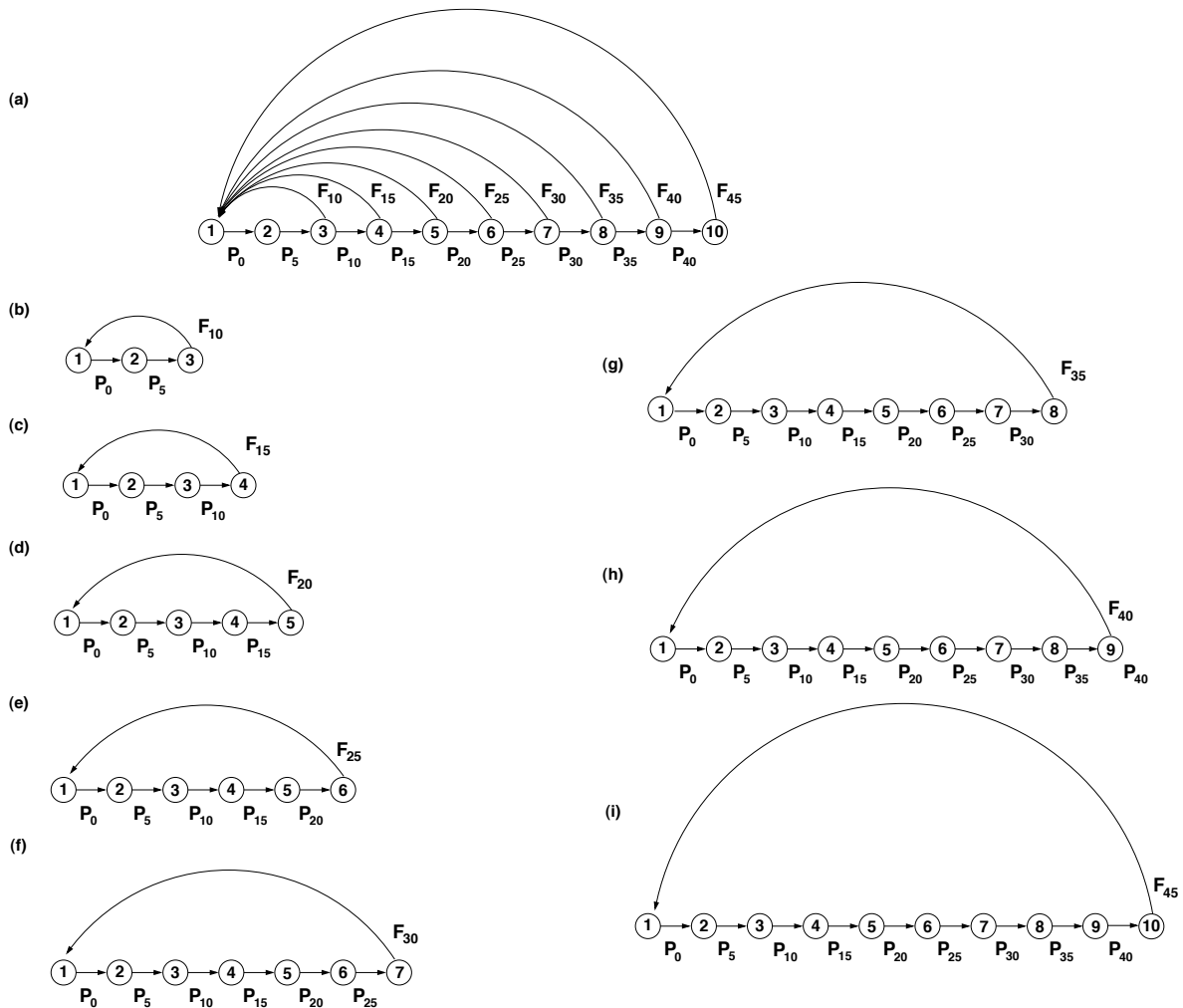

Figure 4: (a) Human life cycle with eight reproductive classes (ages 10-45). (b-i) eight unique loops of the life-cycle diagram in (a).
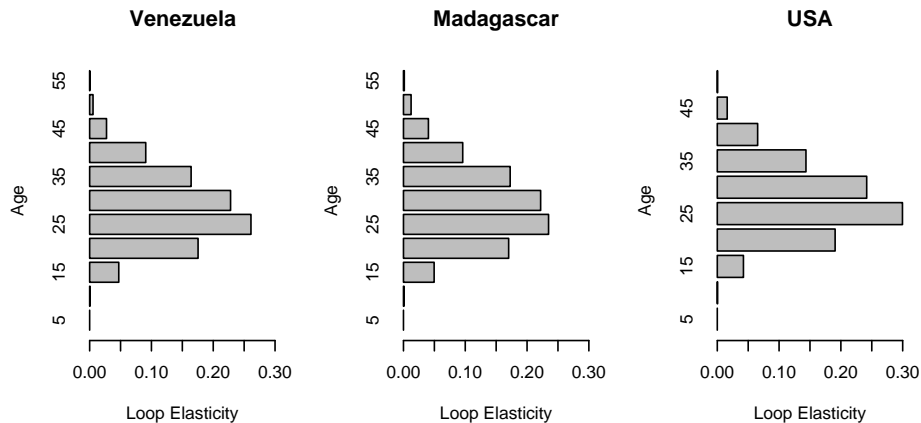
Figure 5: Loop elasticities for the three populations.

To facilitate comparison between the three populations, we suppress the plot call in `loop.elas` and plot the three populations' loop elasticities side-by-side with common scales.

```
R> ven.le <- loop.elas(A, draw.plot = FALSE)
R> mad.le <- loop.elas(B, draw.plot = FALSE)
R> usa.le <- loop.elas(C, draw.plot = FALSE)

R> par(mfrow = c(2, 3))
R> barplot(ven.le, names.arg = 5 * 1:11, horiz = TRUE, beside = TRUE,
+     xlim = c(0, 0.3), xlab = "Loop Elasticity", ylab = "Age")
R> barplot(mad.le, names.arg = 5 * 1:11, horiz = TRUE, beside = TRUE,
+     xlim = c(0, 0.3), xlab = "Loop Elasticity", ylab = "Age")
R> barplot(usa.le, names.arg = 5 * 1:10, horiz = TRUE, beside = TRUE,
+     xlim = c(0, 0.3), xlab = "Loop Elasticity", ylab = "Age")
```

Figure 5 presents the results. All three populations have a peak loop elasticity coinciding with the mean age of childbearing. The extent of this peak seems to be related to how spread out reproduction is around this mean. A full analysis of this pattern is again beyond the scope of this paper but is taken up in another forthcoming paper (Jones, in prep b).

3. Model life tables

Many parts of the world still lack vital event registration. This is particularly true for populations in the developing world that are of special interest to anthropologists and human behavioral ecologists. In addition to the lack of vital event registration, many populations of anthropological interest are quite small and modeling age-patterns of vital events (gathered, for example through genealogical surveys or archival reconstructions) can be challenging because the small number of events means that sampling variance can swamp out the demographic signal. Model life tables have been used to address both of these problems. The functions `cdmltw`, `cdmlts`, `cdmltn`, and `cdmlte` provide the Coale-Demeny families of model

life tables (West, South, North, and East respectively). These four functions take one optional argument, `sex`, which defaults to `sex=F`. The output of the functions is a list containing the ages, age-class widths, and a series of 25×21 arrays of the various life table functions (25 levels of e_{10} by 21 age classes): survivorship l_x, interval mortality probability $_nq_x$, person-years lived by those dying in the interval $_na_x$, deaths in the interval $_nd_x$, person-years lived in the interval $_nL_x$, central death rate $_nm_x$, person-years of remaining life T_x, and life expectancy e_x.

The code underlying these functions uses the coefficients provided in Coale *et al.* (1983) to calculate the 25 model life tables per family. Each of these model tables is indexed by e_{10}, life expectancy at age 10. The Coale-Demeny regional model life tables are derived from a series of 192 life tables all based on registration data and picked by stringent data-quality standards. The North, East, and South regions are all characterized by the particular pattern of mortality that relates to the causes of death that pertained in the countries from which the model was constructed. For the most part, the regions in the Coale-Demeny system correspond to regions of Europe: North is based largely on life tables from Scandinavian countries; East comes form eastern European countries; South comes from Mediterranean countries. The last "regional" model life table is based on the remaining countries, including the only non-western life tables in the collection. The West regional life tables therefore do not display the systematic deviations from the overall mean of the 192 life tables that the three other regions do. Murray, Ferguson, Lopez, Guillot, Salomon, and Ahmad (2003) provide a recent excellent overview of model life table systems.

Coale and Demeny's method for constructing the regional model life tables involved regression the observed age-specific probabilities of mortality in an interval x to $x + n$ $(_nq_x)$ on the observed life expectancy at age 10 (e_{10}) for the 192 input tables. These regressions were done by ordinary least-squares on both $\log_{10}$ transformed and untransformed values of $_nq_x$. Plots of the predicted values of both these regressions against age will intersect twice, and in constructing the model life tables, Coale and Demeny used predicted $_nq_x$ values from the simple regression for to the left of the first intersection and from the logarithmic regression to the right of the second intersection. Between the two, they used the arithmetic mean of the two predicted values. The $_nq_x$ values are extended to ages 105 by assuming Gompertz mortality above age 80:

$$l_x = l_{80} \exp\{-(\mu(80)/k)(e^{k(x-80)} - 1)\}, \qquad x = 85, 90, \ldots, 105$$

where $k = \log(\mu(105) - \mu(77.5))/27.5$, $\mu(105) = 0.551 + 1.75(_5q_{75})$ for males, $\mu(105) = 0.613 + 1.75(_5q_{75})$ for females, $\mu(77.5) = l_{75}(_5q_{75}/_5L_{75})$, and $\mu(80) = \mu(77.5)\exp(2.5k)$. In all these equations $\mu(\cdot)$ is the mortality hazard for the specified age

The values of e_{10} were chosen by an iterative procedure in such as way as to yield even 2.5-year intervals of e_0 from 20 to 80 years. Male values of e_{10} for a given model lifetable are a linear function of the female e_{10}. Unfortunately, Coale and Demeny do not provide the coefficients for this relationship, but they can be recovered by copying the published e_0 values and regressing the values of e_{10} for males onto the values of e_{10} for females.

We can demonstrate the use of the Coale-Demeny model life tables with an example from the demography of hunter-gatherers. Howell (1979) pioneered the application of formal demographic methods to anthropological populations with her analysis of the Dobe !Kung of Botswana. Faced with a very small number of events and a population that did not keep track

of their ages, Howell dealt with some serious challenges in constructing age-specific schedules of mortality and reproduction. She determined relative ages of the extant population and then assigned ages to as many people as possible using commonly known unusual events in the recent history of the people to triangulate individuals births and deaths with calendar years. She then used the Coale-Demeny West family model life table to smooth her estimated age distribution, which allowed her to estimate the age-specific mortality rate of the total population.

In their demographic study of the Ache of eastern Paraguay, Hill and Hurtado (1996) developed field techniques to allow them to avoid relying on model life tables for smoothing the age distribution. We can compare the !Kung and Ache l_x curves to model model l_x curves generated from the West family of life tables.

```
R> cdw <- cdmltw()
R> names(cdw)

 [1] "age"    "width" "lx"     "nqx"    "nax"    "ndx"    "nLx"    "nmx"    "Tx"
[10] "ex"

R> plot(cdw$age, cdw$lx[5, ], type = "l", lty = 3, col = "red",
+     xlab = "Age", ylab = "Survivorship")
R> for (i in 6:14) lines(cdw$age, cdw$lx[i, ], lty = 3, col = "red")
R> ache.nKx <- c(292, 1045, 827, 966, 817, 664, 502, 376, 327, 277,
+     199, 180, 125, 83, 43, 13, 2)
R> ache.nDx <- c(34, 46, 17, 10, 6, 8, 1, 6, 1, 4, 5, 3, 3, 3, 2,
+     1, 1)
R> ache.age <- c(0, 1, seq(5, 75, by = 5))
R> ache.lt <- life.table(x = ache.age, nKx = ache.nKx, nDx = ache.nDx,
+     type = "cd")
R> kung.lx <- c(1, 0.79, 0.6715, 0.6045, 0.5925, 0.5925, 0.551,
+     0.54, 0.508, 0.4475, 0.4045, 0.3765, 0.335, 0.325, 0.2275,
+     0.1595, 0.099, 0.039, 0.02)
R> kung.age <- c(0, 1, seq(5, 85, by = 5))
R> lines(ache.lt$x, ache.lt$lx, lwd = 2)
R> lines(kung.age, kung.lx, lwd = 2, col = "grey")
R> legend(72, 1, c("Ache", "!Kung"), lwd = 2, lty = 1, col = c("black",
+     "grey"))
```

Both populations show the same general shape (figure 6). The Ache survivorship function, however, crosses several model l_x schedules in the early part of life. This could reflect under-reporting of early deaths among the Ache. However, it could also easily be a real pattern since all of the 192 life tables that went into the Coale-Demeny life tables were for state-level populations. Hunter-gatherers could conceivably show quite distinct patterns of mortality as they live at low densities and may avoid some forms of infectious disease mortality which might otherwise cause childhood deaths in a population with comparable life expectancy at birth. Hill and Hurtado (1996) present a compelling argument for why the Ache should deviate from the Coale-Demeny model tables. Their critique suggests possible future development of model
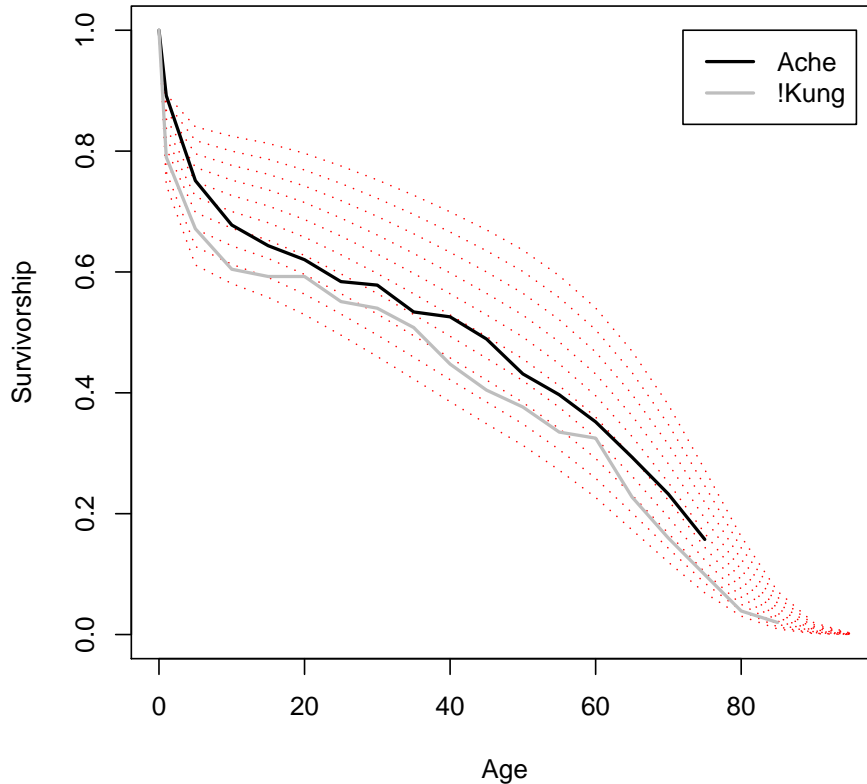
Figure 6: Survival curves for the !Kung (grey) and Ache (black) superimposed on Coale-Demeny West model life table survivorship schedules.

life tables that do not exclude the kind of catastrophic mortality frequently experienced by small-scale populations.

While this section is more specifically applicable to human demography, model life tables have been applied in nonhuman ecological contexts. For example, Dyke and colleagues have constructed a series of model life tables for nonhuman primates (Gage and Dyke 1988; Dyke, Gage, Ballou, Petto, Tardiff, and Williams 1993; Dyke, Gage, Alford, Swenson, and Williams-Blangero 1995). Caswell, Brault, Read, and Smith (1998) developed a novel application of the model life table concept to study incidental mortality in harbor porpoises.

4. Future work

Future development of **demogR** will focus on two primary areas: (1) incorporation of stochastic matrix models, and (2) curve-fitting to model age-schedules. For stochastic population models, I plan to write functions to estimate stochastic growth rate under independent and identically distributed (iid) and Markovian environments and calculate the sensitivities of

these stochastic growth rates to perturbations of both the vital rates and, in the case of Markovian environments, the environmental transition probabilities (Tuljapurkar and Haridas 2006). The curve-fitting for model schedules is more specifically related to human demography. Smoothing of empirical age distributions (particularly for small populations) using model life tables and observed growth rates is on particularly important application. A second will be to fit observed marriage transitions to a Coale-McNeil model nuptiality schedule (Coale and McNeil 1972). Another extension planned for the near future is a series of functions that provide other model life tables (e.g., UN, INDPETH) (United Nations 1981; Network 2002).

Acknowledgments

References

Alvard MS (1998). "Evolutionary Ecology and Resource Conservation." *Evolutionary Anthropology*, **7**(2), 62–74.

Arnold SJ, Wade MJ (1984). "On the Measurement of Natural Selection: Theory." *Evolution.*, **38**(4), 709–719.

Baudisch A (2005). "Hamilton's Indicators of the Force of Selection." *Proceedings of the National Academy of Sciences of the United States of America*, **102**(23), 8263–8268.

Beissinger SR, Westphal MI (1998). "On the Use of Demographic Models of Population Viability in Endangered Species Management." *Journal of Wildlife Management*, **62**(3), 8201–841.

Benton TG, Grant A (1996). "How to Keep Fit in the Real World: Elasticity Analyses and Selection Pressures on Life Histories In A Variable Environment." *American Naturalist*, **147**(1), 115–139.

Benton TG, Grant A (1999). "Elasticity Analysis as an Important Tool in Evolutionary And Population Ecology." *Trends in Ecology & Evolution*, **14**(12), 467–471.

Bird DW, Bird RB, Parker CH (2005). "Aboriginal Burning Regimes and Hunting Strategies in Australia's Western Desert." *Human Ecology*, **33**(4), 443–464.

Borgerhoff Mulder M (1998). "The Demographic Transition: Are We Any Closer to an Evolutionary Explanation?" *Trends in Ecology and Evolution*, **13**(7), 266–270.

Caswell H (1978). "A General Formula for the Sensitivity Of Population Growth Rate to Changes in Life History Parameters." *Theoretical Population Biology*, **14**, 215–230.

Caswell H (1996). "Second Derivatives of Population Growth Rate: Calculation and Applications." *Ecology*, **77**(3), 870–879.

Caswell H (2001). *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer, Sunderland, MA, 2nd edition.

Caswell H, Brault S, Read AJ, Smith TD (1998). "Harbor Porpoise and Fisheries: An Uncertainty Analysis of Incidental Mortality." *Ecological Applications*, **8**(4), 1226–1238.

Caughley G (1966). "Mortality Patterns in Mammals." *Ecology*, **47**, 906–918.

Charmantier A, Garant D (2005). "Environmental Quality and Evolutionary Potential: Lessons from Wild Populations." *Proceedings of the Royal Society B-Biological Sciences*, **272**(1571), 1415–1425.

Charmantier A, Perrins C, McCleery RH, Sheldon BC (2006). "Age-Dependent Genetic Variance in a Life-History Trait in the Mute Swan." *Proceedings of the Royal Society B-Biological Sciences*, **273**(1583), 225–232.

Charnov EL (1993). *Life History Invariants*. Oxford series in ecology and evolution. Oxford University Press, Oxford.

Coale A, Demeny P, Vaughn B (1983). *Regional Model Life Tables and Stable Populations*. Academic Press, New York, 2nd edition.

Coale AJ, McNeil D (1972). "The Distribution by Age of the Frequency of First Marriage in a Female Cohort." *Journal of the American Statistical Association*, **67**(340), 743–749.

Coulson T, Benton TG, Lundberg P, Dall SRX, Kendall BE (2006). "Putting Evolutionary Biology Back in the Ecological Theatre: A Demographic Framework Mapping Genes to Communities." *Evolutionary Ecology Research*, **8**(7), 1155–1171.

Crouse D, Crowder L, Caswell H (1987). "A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation." *Ecology.*, **68**, 1412–1423.

de Kroon H, van Groenendael J, Ehrlen J (2000). "Elasticities: A Review of Methods and Model Limitations." *Ecology*, **81**(3), 607–618.

Deevey E (1947). "Life Tables for Natural Populations of Animals." *Quarterly Review of Biology*, **22**(4), 283–314.

Demetrius L (1969). "The Sensitivity of Population Growth Rate to Perturbations in the Life Cycle Components." *Mathematical Biosciences*, **4**, 129–136.

Doak D, Kareiva P, Klepetka B (1994). "Modeling Population Viability for the Desert Tortoise in the Western Mojave Desert." *Ecological Applications*, **4**(3), 446–460.

Dyke B, Gage T, Ballou J, Petto A, Tardiff S, Williams L (1993). "Model Life Tables for the Smaller New World Monkeys." *American Journal of Primatology*, **29**, 268–285.

Dyke B, Gage TB, Alford PL, Swenson B, Williams-Blangero S (1995). "Model Life Table for Captive Chimpanzees." *American Journal of Primatology*, **37**, 25–37.

Early JD, Peters JF (1990). *The Population Dynamics of the Mucajai Yanomama*. Academic Press, San Diego.

Fisher RA (1958). *The Genetical Theory of Natural Selection*. Dover, New York, 2nd edition.

Gage T, Dyke B (1988). "Model Life Tables for the Larger Old World Monkeys." *American Journal of Primatology*, **16**, 305–320.

Goodman LA, Keyfitz N, Pullum TW (1974). "Family Formation and the Frequency of Various Kinship Relationships." *Theoretical Population Biology*, **5**, 1–27.

Greville T (1943). "Short Methods for Constructing Life Tables." *Record from the American Institute of Actuaries*, **32**, 29–42.

Hamilton W (1966). "The Moulding of Senescence by Natural Selection." *Journal of Theoretical Biology*, **12**, 12–45.

Hill K, Hurtado AM (1996). *Ache Life History*. Aldine de Gruyter, New York.

Howell N (1979). *The Demography of the Dobe !Kung*. Academic Press, New York.

Jones J (2005). "Fetal Programming: Adaptive Life-History Tactic or Making the Best of a Bad Start?" *American Journal of Human Biology*, **17**, 22–33.

Kaplan H, Hill K, Lancaster J, Hurtado AM (2000). "A Theory of Human Life History Evolution: Diet, Intelligence, and Longevity." *Evolutionary Anthropology*, **9**(4), 156–185.

Keyfitz N (1977). *Introduction to the Mathematics of Populations*. Addison-Wesley, Menlo Park, 2nd edition.

Keyfitz N, Flieger W (1990). *World Population Growth and Aging: Demographic Treends in the Late Twentieth Century*. University of Chicago Press, Chicago.

Lande R (1988). "Demographic Models of the Northern Spotted Owl (*Strix occidentalis caurina*)." *Oecologia*, **75**, 601–607.

Lande RA (1982). "A Quantitative Genetic Theory of Life History Evolution." *Ecology*, **63**, 607–615.

Mills LS, Doak DF, Wisdom MJ (1999). "Reliability of Conservation Actions Based on Elasticity Analysis of Matrix Models." *Conservation Biology*, **13**(4), 815–829.

Morris W, Doak D (2002). *Quantitative Conservation Biology*. Sinauer, Sunderland, MA.

Morris WF, Tuljapurkar S, Haridas CV, Menges ES, Horvitz CC, Pfister CA (2006). "Sensitivity of the Population Growth Rate to Demographic Variability within and between Phases of the Disturbance Cycle." *Ecology Letters*, **9**(12), 1331–1341.

Murray CJL, Ferguson BD, Lopez AD, Guillot M, Salomon JA, Ahmad O (2003). "Modified Logit Life Table System: Principles, Empirical Validation, and Application." *Population Studies*, **57**(2), 165–182.

Network I (2002). *Population, Health, and Survival at INDEPTH Sites*, volume 1 of *Population and Health in Developing Countries*. IDRC, Ottawa.

Nussey DH, Clutton-Brock TH, Albon SD, Pemberton J, Kruuk LEB (2005). "Constraints on Plastic Responses to Climate Variation in Red Deer." *Biology Letters*, **1**(4), 457–460.

Nussey DH, Kruuk LEB, Donald A, Fowlie M, Clutton-Brock TH (2006). "The Rate of Senescence in Maternal Performance Increases with Early-Life Fecundity in Red Deer." *Ecology Letters*, **9**(12), 1342–1350.

Pelletier F, Clutton-Brock T, Pemberton J, Tuljapurkar S, Coulson T (2007). "The Evolutionary Demography of Ecological Change: Linking Trait Variation and Population Growth." *Science*, **315**(5818), 1571–1574.

Pfister CA (1998). "Patterns of Variance in Stage-Structured Populations: Evolutionary Predictions and Ecological Implications." *Proceedings of the National Academy of Sciences, USA*, **95**(1), 213–218.

Preston S, Heuveline P, Guillot F (2001). *Demography: Measuring and Modeling Population Processes*. Blackwell, Oxford.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Smith EA, Wishnie M (2000). "Conservation and Subsistence in Small-Scale Societies." *Annual Review of Anthropology*, **29**, 493–524.

Stearns SC (1992). *The Evolution of Life Histories*. Oxford University Press, Oxford.

Steinsaltz D, Evans SN, Wachter KW (2005). "A Generalized Model of Mutation-Selection Balance with Applications to Aging." *Advances in Applied Mathematics*, **35**(1), 16–33.

Tuljapurkar S, Haridas CV (2006). "Temporal Autocorrelation and Stochastic Population Growth." *Ecology Letters*, **9**(3), 324–334.

United Nations (1981). "Model Life Tables for Developing Countries." *Technical Report XIII.7*, United Nations.

van Groenendael J, De Kroon H, Kalisz S, Tuljapurkar S (1994). "Loop Analysis: Evaluating Life History Pathways in Population Projection Matrices." *Ecology*, **75**(8), 2410–2415.

Wardle GM (1998). "A Graph Theory Approach to Demographic Loop Analysis." *Ecology*, **79**(7), 2539–2549.

Affiliation:

James Holland Jones
Department of Anthropological Sciences
Stanford University
450 Serra Mall
Building 360
Stanford, CA, 94505-2117, United States of America
Telephone: +1/650/723-4824
Fax: +1/650/725-9996
E-mail: jhj1@stanford.edu
URL: http://www.stanford.edu/~jhj1/