Extended abstract

Mortality and causes of death: matrix formulation and sensitivity analysis

Hal Caswell
Institute for Biodiversity and Ecosystem Dynamics
University of Amsterdam
Amsterdam, The Netherlands

Nadine Ouellette
Institut National d'Études Démographiques (INED)
Paris, France

Revised December 15, 2015

Contents

1 Introduction **1**

2 Matrix formulation of mortality by cause of death **2**
2.1 Notation . . . 2
2.2 An absorbing Markov chain . . . 3
2.3 Longevity . . . 4
2.4 Modifications of the operative risks . . . 4
2.5 Probability of eventual death due to each cause . . . 5
2.6 Sensitivity analysis . . . 5
2.6.1 Term ① . . . 5
2.6.2 Term ② . . . 5
2.6.3 Term ③ . . . 6
2.6.4 Term ④ . . . 6
2.6.5 Term ⑤ . . . 6
2.7 Parameterization of mortality hazards . . . 7

3 An example **8**

1 Introduction

If demography has an organizing principle, it is that the behavior of cohorts and populations results from differences among individuals. Perhaps the most fundamental difference is between the living and the dead. Much of demography is devoted to exploring differences among living individuals, due to age, sex, health, social or physiological status, etc. The dead often simply disappear from the accounting of the living, at rates that depend on the individual state.

Accounting for causes of death imposes a structure on the dead, as well as on the living. This structure reflects the fact that each cause has a different dependence on age (or other i-states, if present) and that an individual can die only once. The mortality risks due to the various causes are influenced by different factors, in different ways, and hence undergo different changes over time. In addition, they compete with each other, so that changes in the risks of one cause modify the likelihood of eventual death due to all other causes.

Analyses of mortality in terms of causes of death has a long history. Chiang 1961; 1968 formulated the basic stochastic theory and connected it to multiple decrement life tables, and there has been a continual stream of empirical analyses (e.g., Preston et al., 1972; Ouellette et al., 2014) and continuing analytical developments (e.g., Beltrán-Sánchez et al., 2008; Andersen et al., 2013; Andersen, 2013).

In this paper, we reformulate the cause of death problem; our goal is to present a complete (or reasonably so) analysis in terms of matrix operations. Our results are easily computable, and take advantage of the notational and computational advantages of matrix mathematics. They provide straightforward answers to the classical questions of competing causes of death, including the results of various hypothetical situations (certain causes operating alone, or the effect of removing certain causes)

For the purposes of this Extended Abstract, we give a telegraphic summary of the matrix formulation and sensitivity analysis.

We will go beyond current analyses in three ways. First, because we will formulate our model as a Markov chain, we will be able to calculate not only mean, or average, consequences of a set of mortality risks, but also the higher moments and statistics to measure variance and skewness. Finally, we will present a complete perturbation analysis of the model, so that the sensitivity and elasticity of any output to changes in any of the demographic parameters can be directly calculated.

Our approach is to formuate the life cycle as an absorbing Markov chain (Feichtinger, 1971; Caswell, 2001, 2006, 2009, 2011, 2013; Engelman et al., 2014; van Raalte and Caswell, 2013). Although we will focus here on age-classified models, the formulation is general enough to accomodate stage-classified and multi-state formulations of the life cycle, and permits analysis of time-varying vital rates.

2 Matrix formulation of mortality by cause of death

2.1 Notation

Matrices are generally denoted by uppercase boldface symbols (e.g., $\mathbf{U}$). Some block-structured matrices are denoted as, e.g., $\mathbb{U}$. Column vectors are boldface lower case symbols (e.g., $\boldsymbol{\eta}$); $\mathbf{x}^\mathsf{T}$ is the transpose of $\mathbf{x}$. The matrix $\mathbf{I}$ is the identity matrix; the vector $\mathbf{1}$ is a vector of ones, and $\mathbf{e}_i$ is the ith unit vector, with a 1 in entry i and zeros elsewhere. Where necessary, the order of matrices may be indicated by a subscript; e.g., $\mathbf{I}_\omega$ is the identity matrix of order ω. The matrix $\mathbf{E}_{ii}$ is the diagonal matrix with a 1 in the (i,i) entry and zeros elsewhere. The matrix $\mathcal{D}(\mathbf{x})$ is the diagonal matrix with $\mathbf{x}$ on the diagonal and zeros elsewhere. The matrix $\mathbf{X}_{\mathrm{dg}}$ is the matrix with the diagonal of $\mathbf{X}$ on the diagonal and zeros elsewhere. The Kronecker product is denoted by $\otimes$ and the Hadamard (element-by-element) product by $\circ$. As in MATLAB, row i and column j of a matrix $\mathbf{X}$ are denoted by $\mathbf{X}(i,:)$ and $\mathbf{X}(:,j)$, respectively. Logarithms are natural unless specified otherwise.

2.2 An absorbing Markov chain

We define age classes $1, \ldots, \omega$ and absorbing states (causes of death) $1, \ldots, a$ (see Figure 1). Living states are listed before absorbing states, and the transition matrix is

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U} & \mathbf{0} \\ \hline \mathbf{M} & \mathbf{I}_a \end{array} \right) \tag{1}$$

Transitions among transient (living) states are given by $\mathbf{U}$, which is of dimension $\omega \times \omega$. The mortality matrix $\mathbf{M}$ describing transitions from living to dead states is $a \times \omega$.

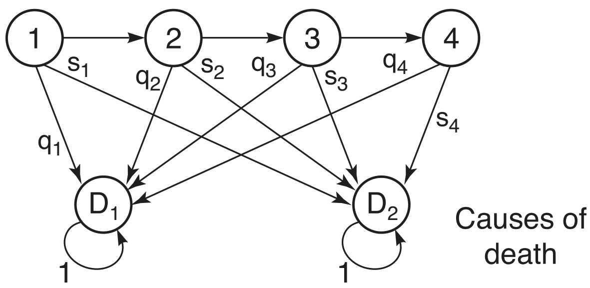

Figure 1: An age-classified life cycle with $\omega = 4$ age classes and $a = 2$ absorbing states based on two causes of death. The age-specific mortality probabilities for cause 1 and cause 2 are denoted by q_i and s_i, respectively.

Figure 1 shows a life cycle in which individuals move among a set of ω transient states until being absorbed into one of a absorbing states that represent different causes of death.

The matrices $\mathbf{U}$ and $\mathbf{M}$ are calculated from a matrix of age- and cause-specific hazards

$$\mathbf{H} = \left(\begin{array}{ccc} \mathbf{h}^{(1)} & \cdots & \mathbf{h}^{(a)} \end{array} \right) \tag{2}$$

in which column j gives the hazard due to cause j.

The total hazard is the sum of the cause-specific hazards

$$\tilde{\mathbf{h}} = \sum_{j=1}^{a} \mathbf{h}^{(j)} \tag{3}$$

$$= \mathbf{H}\mathbf{1}_a \tag{4}$$

The vector of probabilities of death for all age classes, under the total hazard $\tilde{\mathbf{h}}$, is

$$\tilde{\mathbf{q}} = \mathbf{1}_\omega - \exp\left(\tilde{\mathbf{h}}\right) \tag{5}$$

where the exponential function is applied elementwise. The vector of age-specific survival probabilities is

$$\tilde{\mathbf{p}} = \mathbf{1}_\omega - \tilde{\mathbf{q}} \tag{6}$$

In terms of these quantities, $\mathbf{U}$ contains the vector $\tilde{\mathbf{p}}$ on the subdiagonal and zeros elsewhere. Define $\mathbf{Z}$ as an indicator matrix with ones on the subdiagonal and zeros elsewhere. Then

$$\mathbf{U} = \mathbf{Z} \circ \left(\mathbf{1}_\omega \tilde{\mathbf{p}}^{\mathsf{T}}\right). \tag{7}$$

Assuming that the risks of death operate independently,

$$\mathbf{M} = \begin{pmatrix} \frac{h_{11}}{\tilde{h}_1}\tilde{q}_1 & \cdots & \frac{h_{\omega,1}}{\tilde{h}_\omega}\tilde{q}_\omega \\ \vdots & & \vdots \\ \frac{h_{1a}}{\tilde{h}_1}\tilde{q}_1 & \cdots & \frac{h_{\omega,a}}{\tilde{h}_1}\tilde{q}_\omega \end{pmatrix} \tag{8}$$

$$= \mathbf{H}^{\mathsf{T}} \mathcal{D}(\tilde{\mathbf{q}}) \mathcal{D}(\tilde{\mathbf{h}})^{-1} \tag{9}$$

2.3 Longevity

Let ν_{ij} be the number of visits to transient state i, prior to absorption, by an individual starting in transient state j. The expectations of the ν_{ij} are entries of the fundamental matrix $\mathbf{N} = \mathbf{N}_1 = \Big(E(\eta_{ij}) \Big)$:

$$\mathbf{N} = (\mathbf{I} - \mathbf{U})^{-1} \tag{10}$$

(e.g., Kemeny and Snell, 1960; Iosifescu, 1980)). Let $\mathbf{N}_k = \Big(E(\eta_{ij}^k) \Big)$ be a matrix containing the kth moments about the origin of the ν_{ij}. The first several of these matrices are (Iosifescu, 1980, Thm. 3.1)

$$\begin{aligned}
\mathbf{N}_1 &= (\mathbf{I} - \mathbf{U})^{-1} && (11)\\
\mathbf{N}_2 &= (2\mathbf{N}_{\rm dg} - \mathbf{I})\,\mathbf{N}_1 && (12)\\
\mathbf{N}_3 &= \big(6\mathbf{N}_{\rm dg}^2 - 6\mathbf{N}_{\rm dg} + \mathbf{I}\big)\,\mathbf{N}_1 && (13)\\
\mathbf{N}_4 &= \big(24\mathbf{N}_{\rm dg}^3 - 36\mathbf{N}_{\rm dg}^2 + 14\mathbf{N}_{\rm dg} - \mathbf{I}\big)\,\mathbf{N}_1. && (14)
\end{aligned}$$

Let η_j be the time to absorption starting in transient state j and let $\boldsymbol{\eta}_k = E\left(\ \eta_1^k, \cdots, \eta_\omega^k\ \right)^{\mathsf{T}}$. The first several of these moments are (Iosifescu, 1980, Thm. 3.2)

$$\begin{aligned}
\boldsymbol{\eta}_1^{\mathsf{T}} &= \mathbf{e}^{\mathsf{T}}\mathbf{N}_1 && (15)\\
\boldsymbol{\eta}_2^{\mathsf{T}} &= \boldsymbol{\eta}_1^{\mathsf{T}}\,(2\mathbf{N}_1 - \mathbf{I}) && (16)\\
\boldsymbol{\eta}_3^{\mathsf{T}} &= \boldsymbol{\eta}_1^{\mathsf{T}}\,\big(6\mathbf{N}_1^2 - 6\mathbf{N}_1 + \mathbf{I}\big) && (17)\\
\boldsymbol{\eta}_4^{\mathsf{T}} &= \boldsymbol{\eta}_1^{\mathsf{T}}\,\big(24\mathbf{N}_1^3 - 36\mathbf{N}_1^2 + 14\mathbf{N}_1 - \mathbf{I}\big)\,. && (18)
\end{aligned}$$

The variance, skewness, and other statistics are calculated from these moments.

2.4 Modifications of the operative risks

Modifications to the hazard matrix $\mathbf{H}$ can define the subsets of hazards operating. Create an indicator vector $\mathbf{r}$ (dimension $a \times 1$), where

$$r_j = \begin{cases} 1 & \text{cause } j \text{ operating} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

For example, to eliminate cause j, set $r_j = 0$ and all other entries of $\mathbf{r}$ to 1. To examine cause j in isolation, set $r_j = 1$ and all other entries of $\mathbf{r}$ to 0. Then, the modified hazard matrix can be written

$$\mathbf{H}_r = \mathbf{H}\mathcal{D}(\mathbf{r}) \tag{20}$$

In Chiang's (1961) terminology,[1] the *crude* probability of death from cause i when all causes are operating is obtained from setting $\mathbf{r} = \mathbf{1}_a$. The *net* probability of death from cause i when all other causes have been (hypothetically) removed is obtained by setting $\mathbf{r} = \mathbf{e}_i$. The *partial crude* probability, when one cause of death (say cause j) is eliminated but all others are operating is obtained by setting $\mathbf{r}$ to a vector with a zero in position j and ones elsewhere. Extensions to partial crude probabilities under arbitrary choices of causes operating and not operating are obvious.

[1]For a more complete discussion of the various terminologies used for these concepts, and the assumptions involved, see Andersen et al. (2013).

2.5 Probability of eventual death due to each cause

Let

$$b_{ij} = P\left[\text{death due to cause } i \,|\text{current age } j\right] \qquad i = 1, \ldots, a \quad j = 1, \ldots, \omega \tag{21}$$

Then

$$\mathbf{B} = \mathbf{M}\mathbf{N}_1 \tag{22}$$

The jth column of $\mathbf{B}$ gives the probability distribution of eventual cause of death for individuals in age class j. The ith row of $\mathbf{B}$ gives the probabilities that members of each age class eventually die of cause i.

2.6 Sensitivity analysis

Suppose that a vector $\boldsymbol{\theta}$ of parameters determines the age- and cause-specific hazards, so that $\mathbf{H} = \mathbf{H}[\boldsymbol{\theta}]$, and that some demographic output $\boldsymbol{\xi}$ (scalar- or vector-valued) has been calculated as a function of $\mathbf{U}$ and/or $\mathbf{M}$. Our goal is to compute $d\boldsymbol{\xi}/d\boldsymbol{\theta}^{\mathsf{T}}$.

A determined application of the chain rule gives

$$\frac{d\xi}{d\boldsymbol{\theta}^{\mathsf{T}}} = \underbrace{\frac{d\xi}{d\mathrm{vec}^{\mathsf{T}}\mathbf{U}} \frac{d\mathrm{vec}\,\mathbf{U}}{d\tilde{\mathbf{p}}^{\mathsf{T}}} \frac{d\tilde{\mathbf{p}}^{\mathsf{T}}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}} \frac{d\mathrm{vec}\,\mathbf{H}}{d\boldsymbol{\theta}^{\mathsf{T}}}}_{\text{effects through } \mathbf{U}} + \underbrace{\frac{d\xi}{d\mathrm{vec}^{\mathsf{T}}\mathbf{M}} \frac{d\mathrm{vec}\,\mathbf{M}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}} \frac{d\mathrm{vec}\,\mathbf{H}}{d\boldsymbol{\theta}^{\mathsf{T}}}}_{\text{effects through } \mathbf{M}} \tag{23}$$

The first term captures effects of $\boldsymbol{\theta}$ that operate through $\mathbf{U}$; the second term captures effects that operate through $\mathbf{M}$.

We consider each of the terms in (23) in turn.

$$\frac{d\xi}{d\boldsymbol{\theta}^{\mathsf{T}}} = \underbrace{\frac{d\xi}{d\mathrm{vec}^{\mathsf{T}}\mathbf{U}}}_{①} \underbrace{\frac{d\mathrm{vec}\,\mathbf{U}}{d\tilde{\mathbf{p}}^{\mathsf{T}}}}_{②} \underbrace{\frac{d\tilde{\mathbf{p}}^{\mathsf{T}}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}}}_{③} \underbrace{\frac{d\mathrm{vec}\,\mathbf{H}}{d\boldsymbol{\theta}^{\mathsf{T}}}}_{④} + \underbrace{\frac{d\xi}{d\mathrm{vec}^{\mathsf{T}}\mathbf{M}}}_{①} \underbrace{\frac{d\mathrm{vec}\,\mathbf{M}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}}}_{⑤} \underbrace{\frac{d\mathrm{vec}\,\mathbf{H}}{d\boldsymbol{\theta}^{\mathsf{T}}}}_{④} \tag{24}$$

2.6.1 Term ①

The derivatives

$$\frac{d\boldsymbol{\xi}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{U}} \quad \text{and} \quad \frac{d\boldsymbol{\xi}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{M}} \tag{25}$$

depend on $\boldsymbol{\xi}$ and how it is calculated from $\mathbf{U}$ and/or $\mathbf{M}$. For example, if $\boldsymbol{\xi} = \mathbf{N}$, then

$$\frac{d\mathrm{vec}\,\mathbf{N}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{U}} = \mathbf{N}^{\mathsf{T}} \otimes \mathbf{N} \qquad \frac{d\mathrm{vec}\,\mathbf{N}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{M}} = \mathbf{0}_{\omega^2 \times a\omega} \tag{26}$$

Caswell (2006); see Caswell (2009, 2013) for more examples.

2.6.2 Term ②

Differentiating $\mathbf{U}$ in (7) gives

$$d\mathbf{U} = \mathbf{Z} \circ \left(\mathbf{1}_\omega d\tilde{\mathbf{p}}^{\mathsf{T}}\right). \tag{27}$$

Applying the vec operator yields

$$d\mathrm{vec}\,\mathbf{U} = \mathcal{D}\left(\mathrm{vec}\,\mathbf{Z}\right)\left(\mathbf{I}_\omega \otimes \mathbf{1}_\omega\right) d\tilde{\mathbf{p}}. \tag{28}$$

2.6.3 Term ③

The derivative of $\tilde{\mathbf{p}}$ is obtained by differentiating (5) and (6):

$$\begin{aligned} d\tilde{\mathbf{p}} &= -d\tilde{\mathbf{q}} && (29) \\ &= -\mathcal{D}(\tilde{\mathbf{p}})\, d\tilde{\mathbf{h}} && (30) \end{aligned}$$

But we note that $\tilde{\mathbf{h}} = \mathbf{H}\mathbf{1}_a$, so that

$$d\tilde{\mathbf{h}} = \left(\mathbf{1}_a^{\mathsf{T}} \otimes \mathbf{I}_\omega\right) d\mathrm{vec}\,\mathbf{H} \tag{31}$$

and thus

$$d\tilde{\mathbf{p}} = -\mathcal{D}(\tilde{\mathbf{p}}) \left(\mathbf{1}_a^{\mathsf{T}} \otimes \mathbf{I}_\omega\right) d\mathrm{vec}\,\mathbf{H} \tag{32}$$

2.6.4 Term ④

The derivative of $\mathbf{H}$ with respect to $\boldsymbol{\theta}$ depends on the identity of $\boldsymbol{\theta}$ and how it determines the age- and cause-specific hazards. See Section 2.7 for some potentially interesting examples.

2.6.5 Term ⑤

To differentiate $\mathbf{M}$, we note from (8) that row i of $\mathbf{M}$ is $\mathbf{q}^{(i)\mathsf{T}}$. Using equation (9), $\mathbf{M}$ can be written as

$$\mathbf{M} = \mathbf{H}^{\mathsf{T}} \mathcal{D}(\tilde{\mathbf{q}})\, \mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} \tag{33}$$

Differentiating equation (33) gives

$$\begin{aligned} d\mathbf{M} &= \left(d\mathbf{H}^{\mathsf{T}}\right) \mathcal{D}(\tilde{\mathbf{q}})\, \mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} \\ &\quad + \mathbf{H}^{\mathsf{T}} \left(d\mathcal{D}(\tilde{\mathbf{q}})\right) \mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} \\ &\quad + \mathbf{H}^{\mathsf{T}} \mathcal{D}(\tilde{\mathbf{q}}) \left(d\mathcal{D}\left(\tilde{h}\right)^{-1}\right) && (34) \end{aligned}$$

Applying the vec operator to both sides yields

$$\begin{aligned} d\mathrm{vec}\,\mathbf{M} &= \left(\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} \mathcal{D}(\tilde{\mathbf{q}}) \otimes \mathbf{I}_a\right) d\mathrm{vec}\,\mathbf{H}^{\mathsf{T}} \\ &\quad + \left(\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} \otimes \mathbf{H}^{\mathsf{T}}\right) \mathrm{vec}\, d\mathcal{D}(\tilde{\mathbf{q}}) \\ &\quad + \left(\mathbf{I}_\omega \otimes \mathbf{H}^{\mathsf{T}} \mathcal{D}(\tilde{\mathbf{q}})\right) \mathrm{vec}\, d\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} && (35) \end{aligned}$$

The derivative of the transpose of $\mathbf{H}$ is

$$d\mathrm{vec}\,\mathbf{H}^{\mathsf{T}} = \mathbf{K}_{s,a} d\mathrm{vec}\,\mathbf{H} \tag{36}$$

where $\mathbf{K}_{s,a}$ is the vec-permutation matrix, or commutation matrix (Magnus and Neudecker, 1979). We can use results in Caswell (2006, 2009) to write

$$\mathrm{vec}\, d\mathcal{D}(\tilde{\mathbf{q}}) = \mathcal{D}(\mathrm{vec}\,\mathbf{I}_\omega) \left(\mathbf{I}_\omega \otimes \mathbf{1}_\omega\right) d\tilde{\mathbf{q}} \tag{37}$$

and

$$\mathrm{vec}\, d\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} = -\left[\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1} \otimes \mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1}\right] \mathcal{D}(\mathrm{vec}\,\mathbf{I}_\omega) \left(\mathbf{I}_\omega \otimes \mathbf{1}_\omega\right) d\tilde{\mathbf{h}} \tag{38}$$

Putting all the pieces together gives the derivative of $\mathbf{M}$, in three terms, as

$$\begin{aligned}\frac{d\mathrm{vec}\,\mathbf{M}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}} &= \left(\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1}\mathcal{D}\left(\tilde{\mathbf{q}}\right)\otimes\mathbf{I}_a\right)\mathbf{K}_{s,a} \\ &\quad+\left(\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1}\otimes\mathbf{H}^{\mathsf{T}}\right)\mathcal{D}\left(\mathrm{vec}\,\mathbf{I}_\omega\right)\left(\mathbf{I}_\omega\otimes\mathbf{1}_\omega\right)\frac{d\tilde{\mathbf{q}}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}} \\ &\quad-\left(\mathbf{I}_\omega\otimes\mathbf{H}^{\mathsf{T}}\mathcal{D}\left(\tilde{\mathbf{q}}\right)\right)\left[\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1}\otimes\mathcal{D}\left(\tilde{\mathbf{h}}\right)^{-1}\right]\mathcal{D}\left(\mathrm{vec}\,\mathbf{I}_\omega\right)\left(\mathbf{I}_\omega\otimes\mathbf{1}_\omega\right)\frac{d\tilde{\mathbf{h}}}{d\mathrm{vec}^{\mathsf{T}}\mathbf{H}}\end{aligned} \tag{39}$$

2.7 Parameterization of mortality hazards

The final step is to define an interesting parameter vector $\boldsymbol{\theta}$ and its influence on the hazard matrix $\mathbf{H}[\boldsymbol{\theta}]$. The possibilities are endless, and depend on the questions under investigation, but some interesting examples include the following:

1. Let θ_j be a proportional modification of the jth age-specific hazard (i.e., θ_j multiplies column j of $\mathbf{H}$). Then

$$\mathbf{H}[\boldsymbol{\theta}] = \mathbf{H}_0\mathcal{D}(\boldsymbol{\theta}) \tag{40}$$

where $\mathbf{H}_0$ is a baseline hazard matrix. In this case, differentiating $\mathbf{H}$ in (40) gives

$$\begin{aligned} d\mathbf{H} &= \mathbf{H}_0\mathcal{D}(d\boldsymbol{\theta}) && (41)\\ &= \mathbf{H}_0\left[\mathbf{I}_a\circ(d\boldsymbol{\theta}\mathbf{1}_a^{\mathsf{T}})\right] && (42)\end{aligned}$$

and thus

$$d\mathrm{vec}\,\mathbf{H} = \left(\mathbf{I}_a\otimes\mathbf{H}_0\right)\mathcal{D}\left(\mathrm{vec}\,\mathbf{I}_a\right)\left(\mathbf{1}_a\otimes\mathbf{I}_a\right)d\boldsymbol{\theta}. \tag{43}$$

2. Alternatively, θ_j could be an additive modification of the hazard due to cause j, in which case

$$\mathbf{H}[\boldsymbol{\theta}] = \mathbf{H}_0 + \mathbf{1}_\omega\boldsymbol{\theta}^{\mathsf{T}} \tag{44}$$

Differentiating (44) and applying the vec operator yields

$$d\mathrm{vec}\,\mathbf{H} = \left(\mathbf{I}_a\otimes\mathbf{1}_\omega\right)d\boldsymbol{\theta}. \tag{45}$$

3. Suppose that θ_i specifies proportional changes to the hazards from all causes of death, for an individual of age i, for $i=1,\ldots,s$. In this case, θ_i modifies row i of $\mathbf{H}$; i.e.,

$$\mathbf{H}[\boldsymbol{\theta}] = \mathcal{D}(\boldsymbol{\theta})\mathbf{H}_0, \tag{46}$$

and differentiating gives

$$d\mathrm{vec}\,\mathbf{H} = \left(\mathbf{H}_0\otimes\mathbf{I}_\omega\right)\mathcal{D}\left(\mathrm{vec}\,\mathbf{I}_\omega\right)\left(\mathbf{1}_\omega\otimes\mathbf{I}_\omega\right)d\boldsymbol{\theta}. \tag{47}$$

4. Finally, θ_i might make an additive change to the hazards from all causes of death for an individual of age i. In this case,

$$\mathbf{H}[\boldsymbol{\theta}] = \mathbf{H}_0 + \boldsymbol{\theta}\mathbf{1}_a^{\mathsf{T}}, \tag{48}$$

and

$$d\mathrm{vec}\,\mathbf{H} = \left(\mathbf{1}_a\otimes\mathbf{I}_\omega\right)d\boldsymbol{\theta} \tag{49}$$

This completes the information necessary to compute the derivatives of $\boldsymbol{\xi}$ to the parameters using (23).

3 An example

Figure 2 shows the sensitivity of remaining life expectancy to changes in mortality for each of seven causes of death, for U.S. males in 1979.[2] The perturbation shown is a proportional reduction in mortality across all ages. The smallest effects come from prostate and colorectal cancer; by far the largest impact would come from reductions in mortality due to heart disease. The sensitivity of life expectancy to reductions in lung cancer, cerebrovascular disease, and other cancers is intermediate.

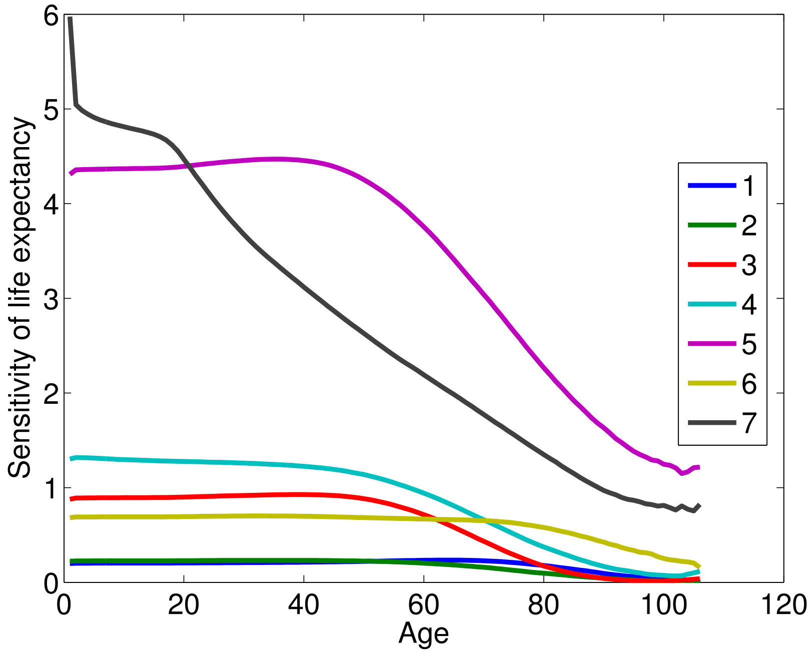

Figure 2: The sensitivity of life expectancy (elements of $\boldsymbol{\eta}_1$) to a proportional reduction (hence, the positive sign; a proportional increase in mortality would reduce life expectancy) in the entire age schedule of mortality due to each of seven causes (1 = prostate cancer, 2 = colorectal cancer, 3 = lung cancer, 4 = other cancers, 5 = heart disease, 6 = cerebrovascular disease, 7 = other). Data are for males in the U.S. in 1979.

[2]The age- and cause-specific male death rates (hazards) for this example were calculated from age- and cause-specific male death counts (from the National Center for Health Statistics; National Center for Health Statistics 2014) divided by the age-specific male population's exposure to the risk of death (from the Human Mortality Database; Human Mortality Database 2014).

References

Andersen, P. K. 2013. Decomposition of number of life years lost according to causes of death. Statistics in medicine **32**:5278–5285.

Andersen, P. K., V. Canudas-Romo, and N. Keiding. 2013. Cause-specific measures of life years lost. Demographic Research **29**:1127–1152.

Beltrán-Sánchez, H., S. H. Preston, and V. Canudas-Romo. 2008. An integrated approach to cause-of-death analysis: cause-deleted life tables and decompositions of life expectancy. Demographic Research **19**:1323–1350.

Caswell, H. 2001. Matrix Population Models. 2nd edition. Sinauer Associates, Sunderland, MA.

Caswell, H., 2006. Applications of Markov chains in demography. Pages 319–334 *in* MAM2006: Markov Anniversary Meeting. Boson Books, Raleigh, North Carolina, USA.

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

Caswell, H. 2011. Beyond R_0: demographic models for variability of lifetime reproductive output. PLoS ONE **6**:e20809.

Caswell, H. 2013. Sensitivity analysis of discrete Markov chains via matrix calculus. Linear Algebra and its Applications **438**:1727–1745.

Chiang, C. L., 1961. On the probability of death from specific causes in the presence of competing risks. Pages 169–180 *in* Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, volume 4. University of California Press, Berkeley, California.

Chiang, C. L. 1968. Introduction to Stochastic Processes in Biostatistics. Wiley, New York, New York.

Engelman, M., H. Caswell, and E. M. Agree. 2014. Why do lifespan variability trends for the young and old diverge? Demographic Research **30**:1367–1396.

Feichtinger, G. 1971. Stochastische Modelle demographischer Prozesse (Lecture Notes in Economics and Mathematical Systems). Springer.

Human Mortality Database, 2014. University of California, Berkley (USA), and Max Planck Institute for Demographic Research (Germany). URL http://www.mortality.org.

Iosifescu, M. 1980. Finite Markov Processes and Their Applications. Wiley, New York, New York.

Kemeny, J. G., and J. L. Snell. 1960. Finite Markov Chains. Van Nostrand, Princeton, New Jersey.

Magnus, J. R., and H. Neudecker. 1979. The commutation matrix: some properties and applications. The Annals of Statistics pages 381–394.

National Center for Health Statistics, 2014. Multiple Cause-of-death Mortality Data. URL http://www.nber.org/data/vital-statistics-mortality-data-multiple-cause-of-death.html.

Ouellette, N., M. Barbieri, and J. R. Wilmoth. 2014. Period-based mortality change: turning points in trends since 1950. Population and Development Review **40**:77–106.

Preston, S. H., R. Schoen, and N. Keyfitz. 1972. Causes of Death: Life Tables for National Populations. Seminar Press, New York, New York.

van Raalte, A. A., and H. Caswell. 2013. Perturbation analysis of indices of lifespan variability. Demography **50**:1615–1640.