Appendix: Constructing Prior Distributions Conditional on Constraints

Jason Hilton, Erengul Dodd, Jon Forster, Peter W. F. Smith
Centre for Population Change,
University of Southampton,
Southampton,
United Kingdom,
SO17 1BJ
J.D.Hilton@soton.ac.uk
January 26, 2018

Supplemental Material to “Projecting UK Mortality using Bayesian Generalised Additive Models”

R Code

Code to run all of the analysis in the paper is provided as a github repository at [https://github.com/jasonhilton/mortality_bgam](https://github.com/jasonhilton/mortality_bgam).

Period Parameters

This section describes how prior distributions for the period constraints are constructed. Some details are repeated from the body of the text for completeness. Period innovations are normally distributed so that:

\[ \kappa_t = \kappa_{t-1} + \epsilon_t \]
\[ \epsilon_t \sim \text{Normal}(0, \sigma^2_\epsilon) \]

Defining \( T \) as the number of periods in the dataset, the 2 by \( T \) matrix \( C \) describes the two constraints on the vector of \( \kappa \) parameters, that they must sum to zero and show no linear growth.

\[ C = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & 1 & 1 \\ 0 & 1 & 2 & 3 & \ldots & T - 2 & T - 1 \end{bmatrix} \]

When the constraints hold,

\[ C\kappa = 0 \]
\[ CS\epsilon = 0, \]

where \( S \) is the cumulative sum matrix. A distribution for period innovations conditional on the constraints is obtained by first transforming \( \epsilon \) into a new set of parameters \( \eta \), where the first two elements of \( \eta \) are zero when the constraints hold, and the remaining elements are identical to the equivalents in \( \epsilon \). The matrix \( Z \) used for this transformation is a \( T \) by \( T \) identity matrix with the first two rows replaced by the matrix \( CS \).

\[ \eta = \begin{bmatrix} \eta^1 \\ \eta^* \end{bmatrix} = Z\epsilon \]
\[ \eta \sim \text{MVN}(0, ZZ^T\sigma^2_\epsilon). \]

By conditioning on the first two values of \( \eta \) (denoted \( \eta^1 \)) equalling zero, we can find the distribution of the last \( t - 2 \) elements of \( \eta \) using the standard conditional relationship for multivariate normal variables. Conditioning on these elements of \( \eta \) being equal to zero is equivalent to conditioning on the constraints holding. We can therefore calculate the values of the first two values of \( \epsilon \) by multiplying \( \eta \) by the inverse of the \( Z \) matrix.
\[ \Sigma = ZZ^T \sigma^2 \]

\[ \eta^*|(\eta^\dagger = 0) \sim MVN(0, \Sigma_{ss} - \Sigma_{st} \Sigma_{tt}^{-1} \Sigma_{ts}) \]

\[ \epsilon = Z^{-1} \begin{bmatrix} 0 \\ \eta^* \end{bmatrix}, \]

where subscripts on the covariance matrices indicate partitions so that \( \Sigma_{st} \) is the sub-matrix of \( \Sigma \) with rows corresponding to \( \eta^* \) and columns to \( \eta^\dagger \).

**Correlated period parameters for males and females**

When both sexes are modelled together, the period innovations for each sex retain separate variance parameters, but are assumed to be joint multivariate normal, with the correlation determined by parameter \( \rho \). The joint distribution for the period innovations \( \epsilon \) is therefore as follows:

\[
\begin{bmatrix} \epsilon_{mf} \\ \epsilon_{km} \end{bmatrix} \sim \text{Multivariate Normal } (0, P)
\]

\[
P = \begin{bmatrix} I_T \sigma_{sf}^2 & I_T \sigma_{km}^2 \\ I_T \sigma_{km}^2 & I_T \sigma_{sf}^2 \end{bmatrix}
\]

\[ \rho \sim \text{Beta}(1, 1), \]

where \( I_T \) is the identity matrix of dimension \( T \), the number of periods, and \( \epsilon_{mf} \) and \( \epsilon_{km} \) refer to period innovations for males and females respectively. The implied joint multivariate distribution can then be conditioned on the constraints holding true for both sexes in the same way as before:

\[
\eta = \begin{bmatrix} \eta^\dagger \\ \eta^* \\ \eta^m \end{bmatrix} = X \epsilon
\]

\[
X = \begin{bmatrix} Z & 0 \\ 0 & Z \end{bmatrix}
\]

\[
\Xi = X P X^T
\]

\[ \eta \sim MVN(0, \Xi) \]

\[ \eta^*|(\eta^\dagger = 0) \sim N(0, \Xi_{ss} - \Xi_{st} \Xi_{tt}^{-1} \Xi_{ts}). \]

**Cohort Parameters**

Cohort effects are modelled as p-splines, with innovations normally distributed, so that

\[
s_\gamma(t - x) = \beta^\gamma B(t - x)
\]

\[
\beta^\gamma_i = \beta^\gamma_{i-1} + \epsilon^\gamma_i
\]

\[
\epsilon^\gamma_i \sim \text{Normal}(0, \sigma^2_\gamma)
\]

with \( B(.) \) giving the B-spline basis function, and \( i \) indexes the individual basis functions. As with the period effects, a matrix \( C \) is describes the constraints on the smooth function \( s_\gamma \), applying to the first, second, and final elements of the parameter vector.

\[
C = \begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 1 & 1 & 1 & \ldots & 1 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix}
\]

The constraints hold when

\[
Cs_\gamma = 0
\]

\[
CBS\epsilon^\gamma = 0.
\]
A new parameter vector $\eta^\gamma$ is obtained as for the period effects

$$\eta^\gamma = W\epsilon^\gamma$$

$$\eta^\gamma = \begin{bmatrix} \eta^\gamma_1 \\ \eta^\gamma_2 \\ \eta^\gamma_3 \end{bmatrix}$$

$$\eta^\gamma \sim \text{MVN}(0, WW^T\sigma_\gamma^2).$$

The matrix $W$ is an identity matrix with the first, second and final rows replaced by the rows of the matrix $CBS$, where $B$ is the matrix of basis functions $b(\cdot)$ evaluated for each cohort. A distribution for $\eta^\gamma_3$ given the constraints can now be constructed by conditioning on $\eta^\gamma_1 = \eta^\gamma_3 = 0$, in the same manner as for the period effects. New cohort basis function innovations can be drawn from the normal distribution with mean 0 and variance $\sigma_\gamma^2$. 