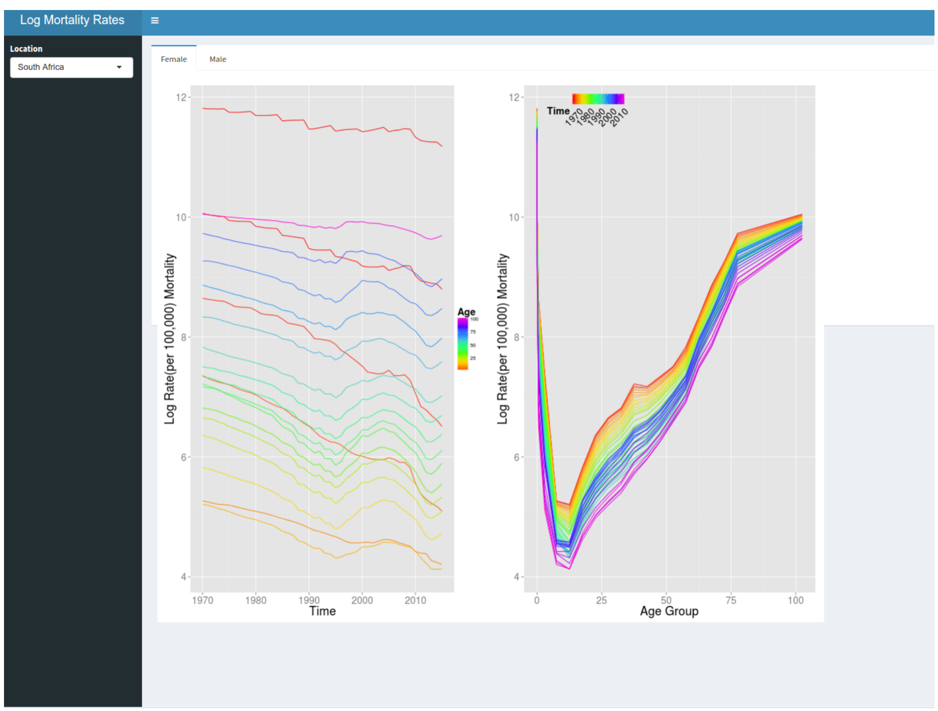
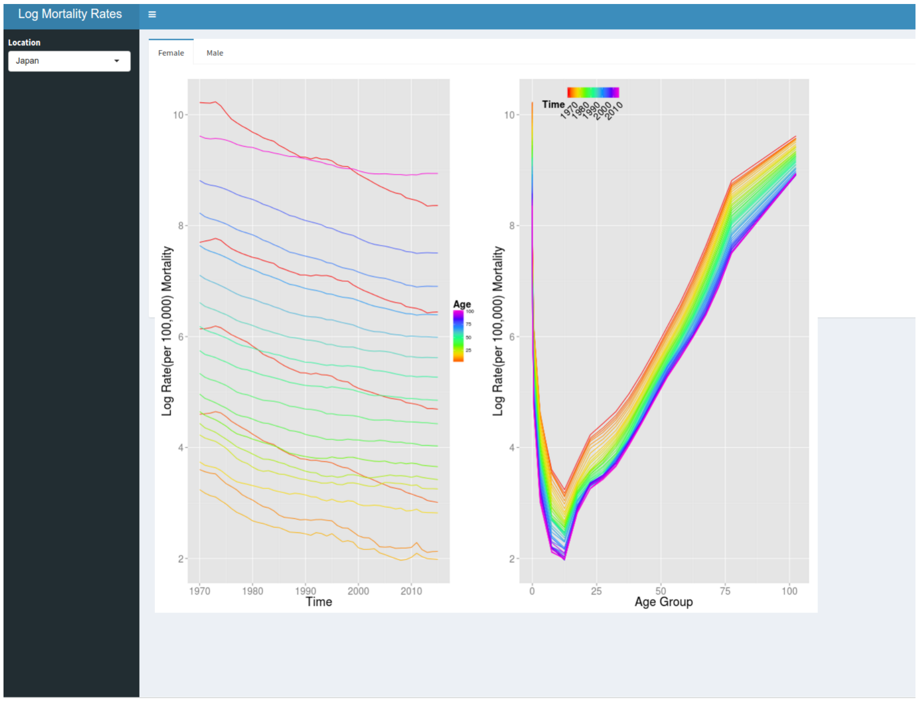
Measuring age specific mortality rates across space and time

Age specific mortality rates and its changing pattern over time have long been concern of demographers and urban planners. As life expectancy increases and the age structure of a society evolves social institutions need to plan accordingly in order to meet the needs of their populations.

In order to estimate life expectancy well, a good model for mortality needs to be used in order to estimate when people will die throughout their life trajectory. In 1825 Gompertz developed the law of human mortality, which William Makeham amended in 1860, which divides mortality into and independent and age dependent components. this model was used for sometime until Silder developed a competing hazards model which divided mortality further into infantile causes of death which decline over time and adult causes of death which increase over time.

This model held for some time and describes the trend of human mortality. Take for example both Japan's and South Africa's mortality rates. The two countries have very different levels of mortality but the pattern across age and time is similar.





In order to estimate mortality across multiple age groups, geographies and time it is important to capture the relatedness across these dimensions while still capturing the general shape of mortality. In order to do this I will use a combination of a modified Silder model which has more interpretable terms with random effects capturing systematic error across age, space and time.

The model skeleton

The underlying model by which we will estimate human mortality, referred to hereafter as the model skeleton, is the underlying shape that we believe human mortality follows at a global level. The form is as follows

$$inf\_rate_x = N0 + exp(x * lambda) + c$$

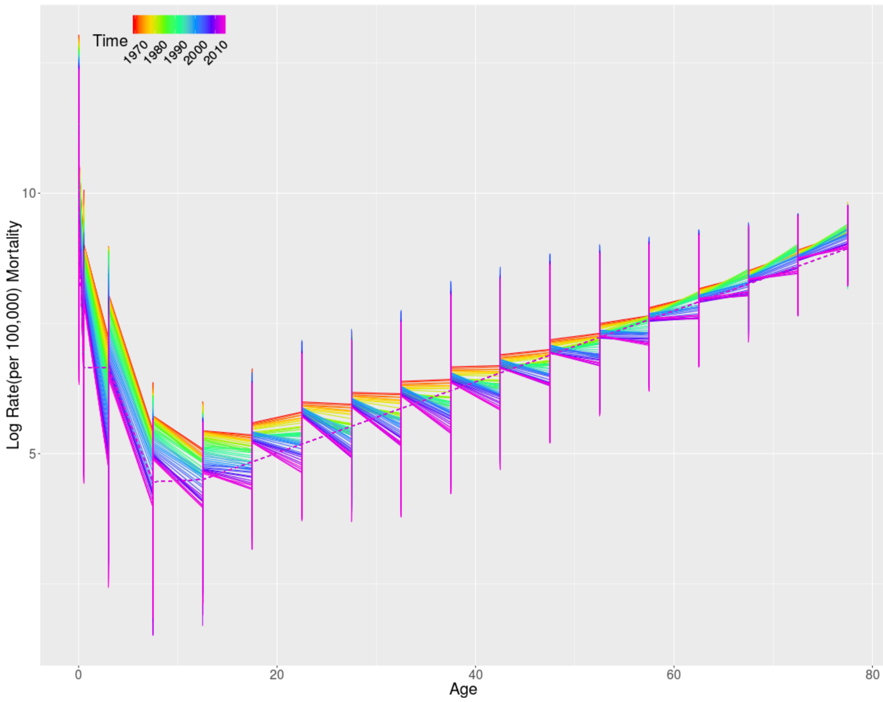
$$sns\_rate_x = m * x + b$$

$$reversal_x = 1/(1 + exp(3 * (eta - x)))$$

$$log(m_x)_{skeleton} = inf\_rate_x * (1 - reversal) + sns\_rate_x * reversal$$

This skeleton provides the basic underlying mortality that we observe across the past 20 years of observation across the globe. In order to simulate with reasonable values a simple model will be run with just these fixed terms so that a reasonable mean function can be used that is representative of observed mortality. In order to gain decent predictive validity however we will add in random effects which we believe capture the relatedness in the error of the model.

The skeleton fit can be seen in the dotted line below along side the data for all administrative locations.



Simulating structured random error terms

The first component is the geographical relatedness in our data. In order to simulate this structure we will begin wih an neighborhood matrix which defines relatedness of geographies based on development status rather than proximity. In this way We will capture the relatedness between countries such as Japan and the United States which have similar patterns of development over the past 50 years while not expecting a correlation between the US and Mexico, which by contrast have had very different patterns of development even though they are geographically close. These groups are defined by previous UN studies.

With these measures of relatedness we arrive at 21 non overlapping regions which house countries which are developmentally similar. From this a precision matrix can be defined using the following structure

$$Q^{geo}_{i,j} = \begin{cases} n_{\delta_i}, & \text{if } i = j \\ -1, & \text{if } i \sim j \\ 0, & \text{otherwise} \end{cases}$$

A precision matrix for time and age is used following an AR1 model with the following structure.

$$Q_{i,j} = \begin{cases} \frac{1+\rho^2}{\sigma^2}, & \text{if } i = j \\ \frac{-\rho}{\sigma^2}, & \text{if } i \sim j \\ 0, & \text{otherwise} \end{cases}$$

where similarity is defined as an adjacent time or age. An independent ρ and σ can be used here for age and time but the simulation used the same default value of .7 for ρ and 1 for σ.

This however only accounts for the similarity of time or age independent of geography. To get the effects of geography-time and geography-age the kroneker product of the two respective precision matrices is taken in order to account for the interactive effect

$$Q^{geo\_time} = Q^{geo} \otimes Q^{time}$$
$$Q^{geo\_age} = Q^{geo} \otimes Q^{age}$$

For each precision matrix the inverse is taken and draws from a multivariate normal are used in order to get correlated random effects for geo, geo_age, and geo_time.

$$\epsilon \sim MVN(0, Q^{-1})$$

Full Form

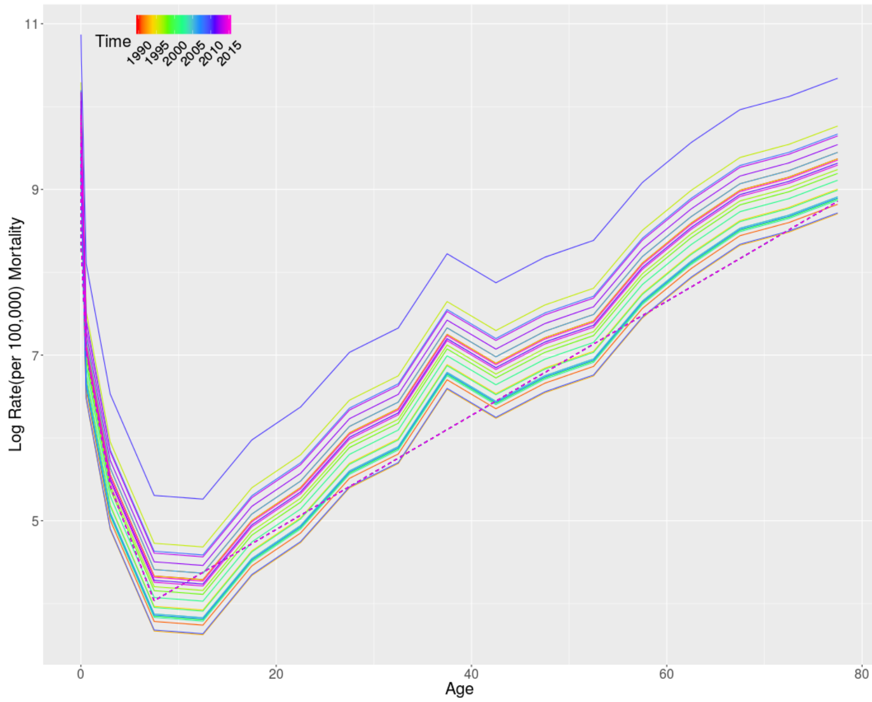
The final estimate takes into account the three sources of structured random effects with the skeleton.

$$log(m_{l,a,t}) = log(m_x)_{skeleton} + \epsilon geo + \epsilon geo\_time + \epsilon geo\_age$$

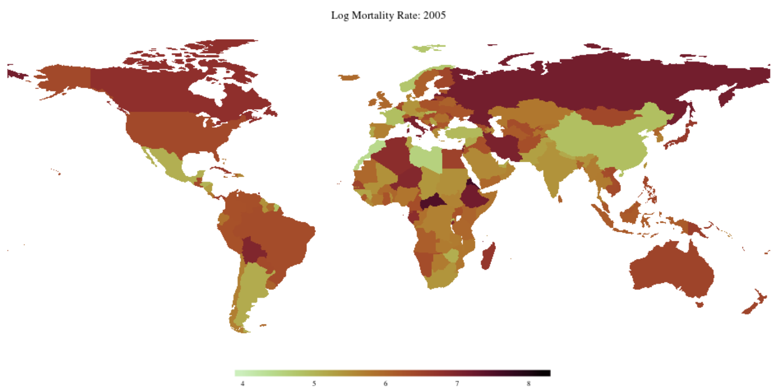
With this log rate estimate we can simulate mortality numbers with some observation error

$$log(m\_obs_{l,a,t}) \sim \mathcal{N}(log(m\_obs_{l,a,t}), \sigma_{obs})$$

Below is a simulated single country with correlated time and age components



Along with the development relatedness map


Log Mortality Rate: 2005
4
5
6
7
8