

Displaying changes in bivariate relationships over age and time

Tim Riffe*¹ and José Manuel Aburto ^{†2}

¹Max Planck Institute for Demographic Research

²Epidemiology, Biostatistics and Biodemography, Department of Public Health,
University of Southern Denmark

May 23, 2018

Abstract

Background Lexis surfaces are an established visualization technique to show how a given value changes over age and time. Vector fields are a two-dimensional representation of two variables: direction and speed (or force).

Objective We aim to increase the information density of patterns shown on the Lexis surface by placing a vector field on the Lexis surface.

Results We show Lexis fields using different combinations of visual encodings, such as color, contour layering, and angle, length, curvature, and thickness of field elements. These instruments enable information layering that is not otherwise possible on standard Lexis surfaces.

Conclusions Lexis fields extend the analytic power of the Lexis surface, and these can be rendered to display information at higher densities than commonly made Lexis surfaces.

Introduction

Lexis surfaces are a graphical tool used to display data on the Lexis coordinate plane, a Cartesian plane that is also a simplex relationship between age, period, and cohort. Surfaces are often displayed as heat maps, contour maps, perspective plots, or variants of these things. Various kinds of quantities, such as raw magnitudes, differences, ratios, intensities, proportions, derivatives, and even compositions (Schöley and Willekens 2017)) can be displayed on Lexis surfaces in order to put age, period, cohort, or other patterns in relief.

Maps in general combine layers of categorical, continuous, and symbolic information on a common spatial projection. Lexis surfaces in contrast almost exclusively display one visual layer at a time. Even the composite surfaces of Schöley and Willekens (2017), which display layered information are rendered as a single visual layer. Small multiples of Lexis surfaces on the other hand constitute a de-layering, as these are spatially disjoint. We propose to enrich Lexis surfaces by adding a visual layer of quantitative information coded symbolically as a vector field, and we liken to cartographic information layering.

Vector fields are a graphical form generally used to display variation in speed, direction, or force over a plane. Point estimates of these quantities on the plane are often represented with segments or arrows, where length may be proportional to a function of magnitude (force, speed), and angle indicates direction, potentially disambiguated with an arrowhead. We propose a fusion of Lexis surfaces and vector fields, *Lexis fields*, as a tool to display variation in relationships between variables over age and time. A Lexis field may either be rendered atop on a Lexis surface, representing two map layers — a true Lexis map — or as a single-layer alone visualization.

We give an overview of constructing a Lexis field, and a single application. Our example explores the relationship between remaining life expectancy and the standard deviation of remaining lifespan over age and time based on all available populations in the Human Mortality Database from 1950 onward. Other potential applications are discussed.

*riffe@demogr.mpg.de

†jmaburto@health.sdu.dk

1 Lexis field construction

Constructing a Lexis field involves several degrees of designer freedom, which can however be codified into four basic steps, which we summarize in Figure 1. It only makes sense to plot a Lexis field if data contain a relationship that can be summarized with a line, a simple curve, or similar, and that varies by age and/or time. The steps to do so are outlined in the following steps, referenced to regions of Figure 1.

- A** Determine the Lexis reference grid size for data selection. For example, a five-year grid implies 5×5 Lexis cells. Data may be selected from multiple populations on the same reference grid. Presumably two variables are required.
- B** Abstract a model from the data, such as a linear, parabolic, or similarly simple fit. We consider the case of a bivariate linear model that produces a result of the form $y = a + b * x$.
- C** Translate the model fit to the characteristics of a line segment, or field *pointer*. Treat each grid cell on the Lexis diagram as a plot area, by default with equal year units in x and y , for example as implied by a 5×5 cell. In our implementation, the pointer always passes through the centroid of the Lexis cell. The pointer angle or slope may be taken as-is from a linear model, or exaggerated by the same multiplier for the entire field to increase definition. Within the Lexis cell a circle tangent to the four cell borders standardizes the length of the pointer, where the radius of the standard circle can be adjusted by defining an inner margin width (pad) to the Lexis cell. In the simplest case, all pointers may be of the same length, irrespective of slope, as determined by this reference circle. Otherwise, length may be proportional to some other data or model characteristic, such as the observed range of x , the goodness of model fit, or similar. Likewise, other segments characteristics, such as color, or width, may also map to data characteristics.
- D** Render the segment in the corresponding Lexis grid cell, repeating all steps for each cell in the diagram. Variations over the Lexis plane in pointer aesthetics

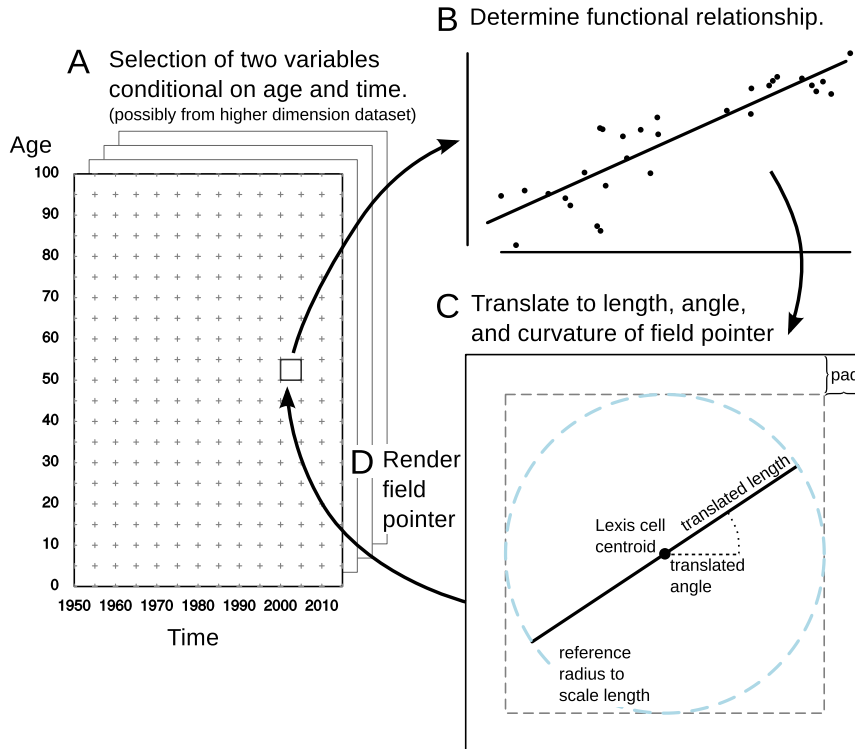


Figure 1: A diagram depicting the translation of functional relationships in data conditioned on age and time to visual encoding on a Lexis field. **A**: Condition data selection on age and time. **B**: model the functional relationship in data subset. **C**: Translate the model to field elements, ‘pointers’, using angle, and possibly also length, curvature, thickness, etc to encode model qualities. **D**: Populate the Lexis plane with field pointers to create a Lexis field.

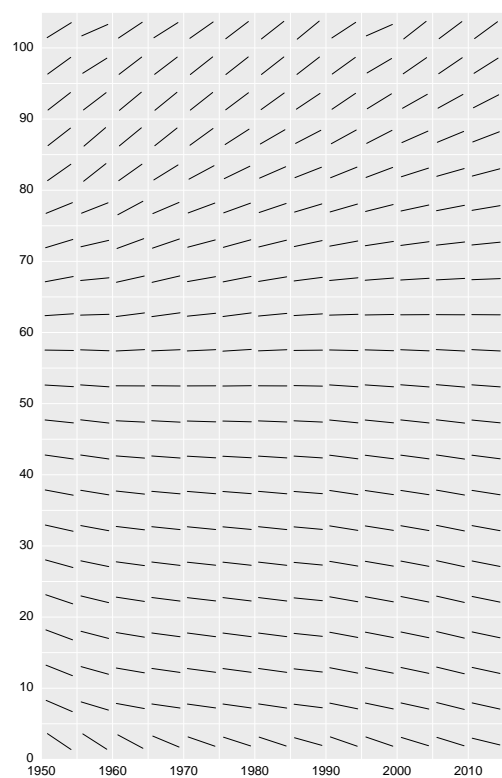
2 Application

We offer a selection of variants of Lexis field design for our example of the relationship between the mean and standard deviation of remaining lifespan. The first of these, Figure 2a, is a bare-bones Lexis field that serves to illustrate the underlying concept. This display differs from a standard vector field in two notable ways. 1) Segments are drawn rather than arrows, because each slope can be interpreted as a subplot, where the convention is to always read a relationship from left to right. 2) The left and right extremes are fixed for each segment according to the year-width rendered, in this case 5-year groupings. This means that relationships farther from zero also result in longer line segments, but the length of the line segment is not strictly proportional to the magnitude of the slope coefficient or any other quantity. For each subplot, the x range represents the same hypothetical range of variation in remaining life expectancy or 2 years. The y-axis is arbitrarily, but identically scaled for each subplot. Other criteria for slope length would also be possible and will be considered.

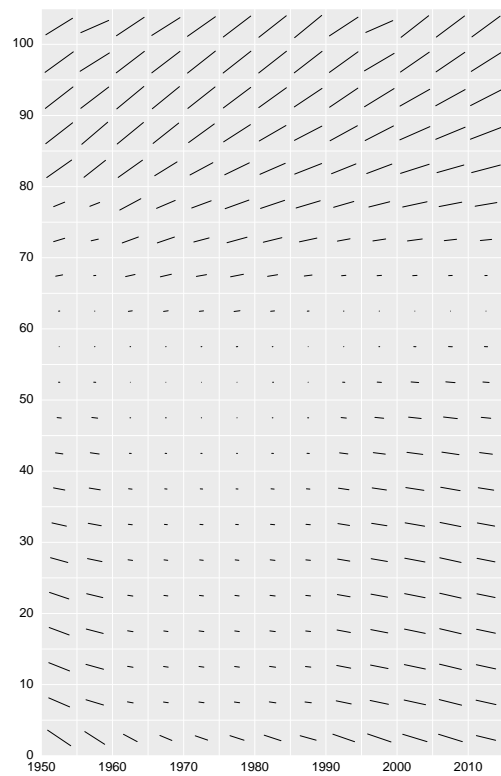
Figure 2b is identical to the previous, but adds a red contour line at the turning point in slopes. In this case there is only one sign change in slope coefficients over the age pattern. It is easier to locate and detect change in this (novel) threshold with the line markup than from the bare Lexis field. Indeed one could add further isoclines at slope break points, as in Figure 2c or ??, and insodoing we begin a transition into the more standard graphical form of contour maps. Figure 2c blends the Lexis field with a heatmap, where hue indicates both direction and magnitude of slope coefficients. In this case, slope coefficients are double-coded, and this display is likely the most legible, but it comes at a cost of clutter. Figure ?? removes the slope segments, but is otherwise identical. We wish to explore variants of the first three of these Lexis field proposals, and may possibly display more example relationships between other variables in a full manuscript, data permitting.

3 Discussion

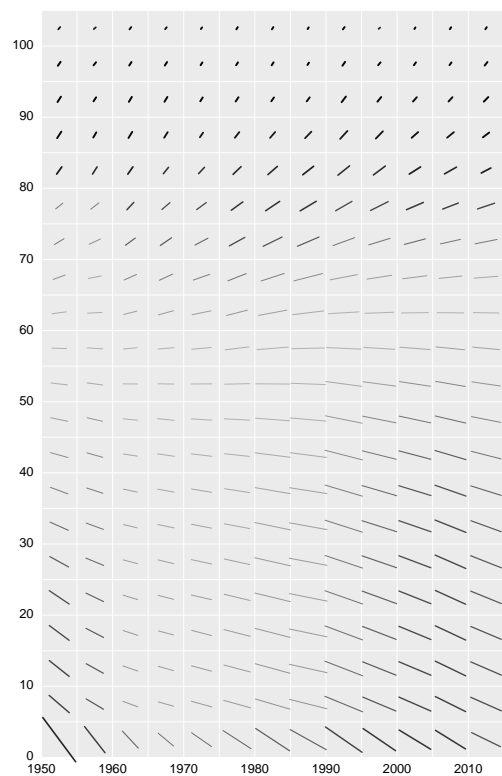
Although patterns in data may be much more complex than can be expressed with linear models, these simple model fits can be thought of as regular samples from the complex space implied by data, such that the pattern revealed on the Lexis field is still intelligible.



(a) A simple field.



(b) A simple field, with the zero-slope isocline drawn (red).



(c) Slopes doubly coded with field segments and discrete color, further slope isoclines drawn.

Figure 2: Four versions of Lexis fields displaying the linear relationship between the standard deviation and mean of remaining lifespan, males (HMD)

References

- Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on 17 April, 2017).
- Jonas Schöley and Frans Willekens. Visualizing compositional data on the lexis surface. *Demographic Research*, 36:627–658, 2017.
- James W Vaupel, Bradley A Gambill, and Anatoli I Yashin. Thousands of data at a glance: shaded contour maps of demographic surfaces. 1987.