Estimating Constrained Functions

Introduction
Positive functions
Fitting strictly . . .

Introduction
Positive functions
Fitting strictly . . .

1. Introduction

- Growth curves should be increasing functions.
- Smooths of magnitudes like precipitation must not be negative.
- Probability density functions have to be positive and integrate to one.
- Functions estimating probabilities have to have values in $[0, 1]$.

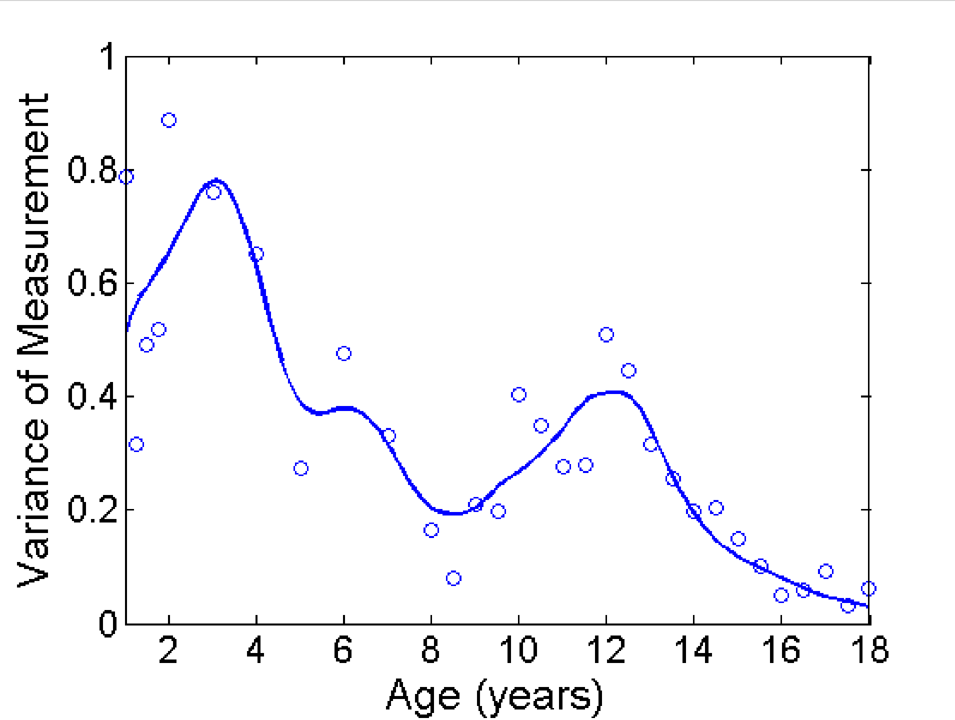
Strategy

- When curves are constrained in ways such as these, it is not helpful to express them as linear combinations of basis functions since it difficult to enforce the constraints.
- Instead, we express a constrained function as a transformation of an unconstrained function.
- Positive functions can be expressed as exponentials of unconstrained functions, for example.
- Expressing the constrained function in terms of a differential equation can often be helpful.

2. Positive functions

Introduction
Positive functions
Fitting strictly . . .
Home Page
Title Page
Page 4 of 18
Go Back
Full Screen
Close
Quit

- The scatter of the observed data about a best-fitting curve is likely to vary over the curve.
- Measurements of height of one year old children have a standard deviation of more than 7 mm, but for adults this is more like 3 mm.
- Consequently we need a curve indicating the variance and standard deviation of errors of measurement.
- But this variance curve must necessarily be positive.



The circles indicate squared residuals averaged across the 54 girls in the Berkeley growth study. The solid line is a positive smooth of these data.

Defining a positive smoothing spline

- We define as the exponential of an unconstrained function W:

$$x(t) = e^{W(t)}$$

- W is the logarithm of x.
- Since $W(t)$ can be positive or negative and is not in any other way constrained, we can expand W in terms of a set of basis functions,

$$W(t) = \sum_k c_k \phi_k(t) \ ,$$

- A Fourier series works for periodic data such as precipitation and B-spline serve for non-periodic data such as the mean squared residuals for the growth data.

Penalizing roughness

- The roughness of positive smoothing function x is determined by the roughness of its logarithm W
- For example, we may fit the data by minimizing

$$[\mathbf{y} - e^{W(\mathbf{t})}]'\mathbf{W}[\mathbf{y} - e^{W(\mathbf{t})}]^2 + \lambda \int [D^2 W(t)]^2 \, dt \; .$$

- We must resort to numerical optimization methods to minimize this or some other fitting criterion with respect to coefficients c_k.
- The growth variance curve was estimated using order 4 B-splines with knots at data locations and $\lambda = 0.0001$.

Positive functions defined by a differential equation

- The transformation $x(t) = e^{W(t)}$ can be expressed as

$$Dx(t) = w(t)x(t).$$

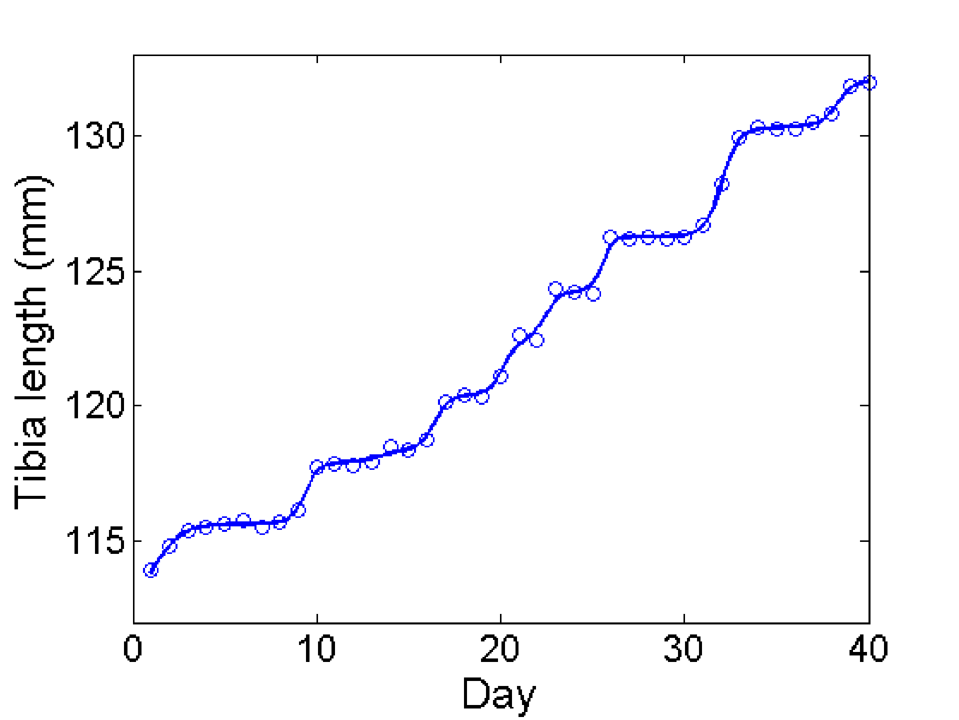
  where

$$w(t) = DW(t)$$

- Positive function x is the solution of a simple *differential equation*.
- This definition of x may seem pedantic and un-intuitive, but we will see that it is a powerful tool for defining other constrained functions.

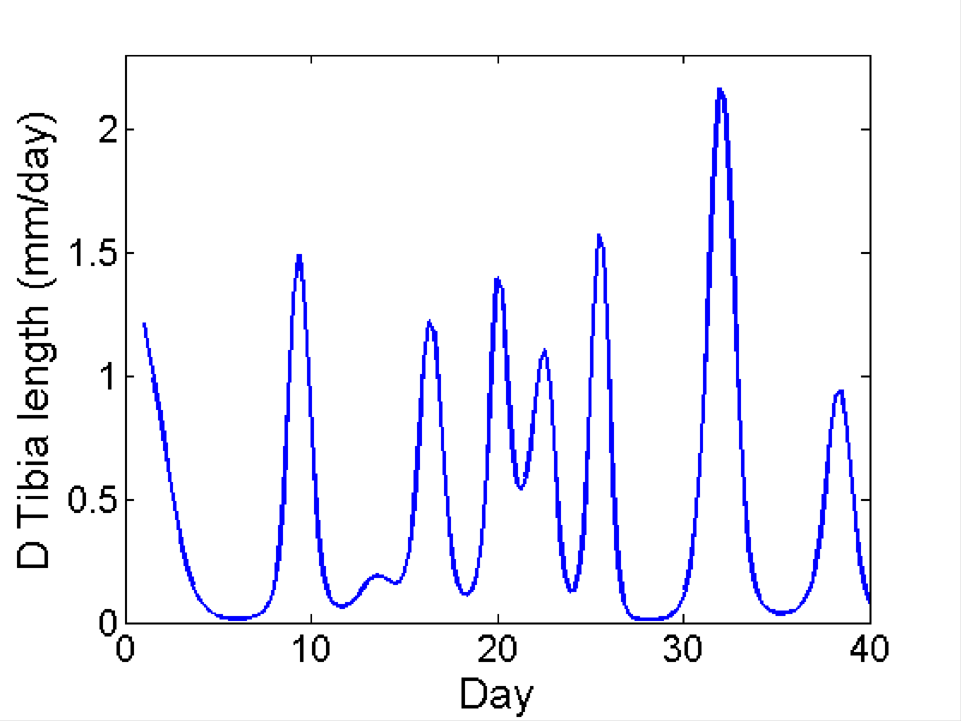
3. Fitting strictly monotone functions

- Growth curves should not in principle lose height, even when the noisy data themselves don't always increase.
- It is difficult to prevent smoothing splines from doing this, especially near adulthood, when the curves flatten.
- Daily measurements of the length of the tibia of a new-born baby, accurate to within about 0.1 mm, challenge our capacity to keep curve monotone.



The solid line is the fit to the data using a monotone smoothing spline.

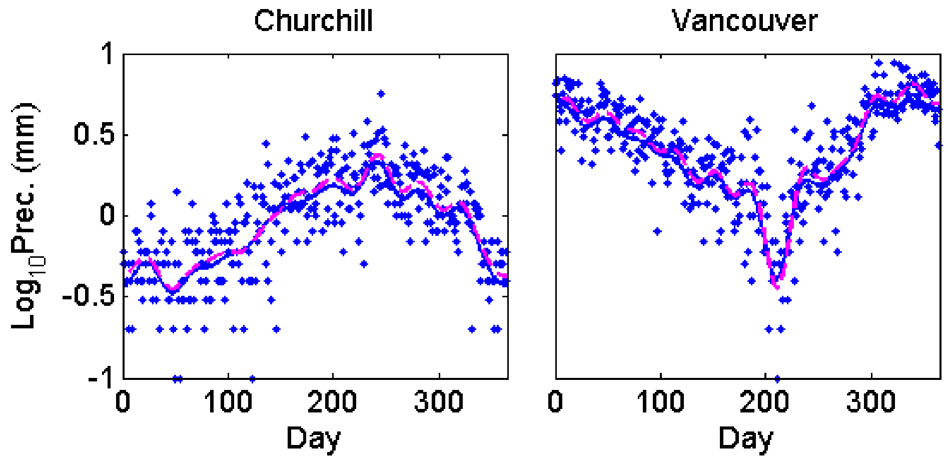
The velocity of the baby's growth.



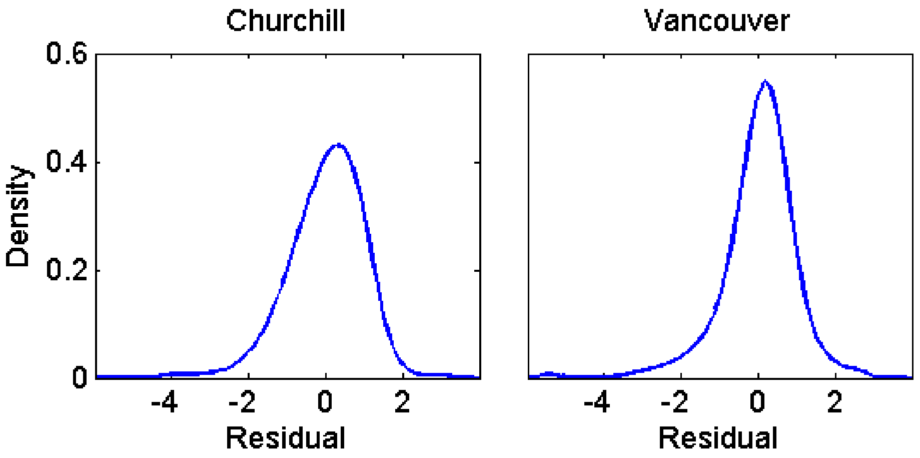
4. Smoothing data with estimation of the density of residuals

- Most approaches to smoothing either implicitly or explicitly assume that the scatter of the data about the curve has a normal or Gaussian distribution, and usually with a constant standard deviation.
- But we often see decidedly non-Gaussian behavior in the actual scatter:
  - long tails
  - pronounced skewness
  - multi-modality
- We also see the nature of the scatter varying over space, time, and other variables.

- We need the technology to smooth data and estimate the density function for the scatter at the same time.
- The density function may need to be a density surface, varying in shape smoothly over t.



Home Page
Title Page
Page 16 of 18
Go Back
Full Screen
Close
Quit


Churchill
Vancouver
Density
0.6
0.4
0.2
0
-4
-2
0
2
Residual
-4
-2
0
2
Residual

Introduction
Positive functions
Fitting strictly . . .