Introduction to Shape-Restricted Regression Splines

Wenjie Wang

03 March 2016

Introduction

Introduction

- Let's consider the regression model

$$y_i = f(x_i) + \sigma\epsilon_i, \;\; i = 1, 2, \ldots, n,$$

  where the errors $\epsilon_i, \;\; i = 1, \ldots, n$, are i.i.d.
- When a parametric form for f is not avaiable, nonparametric methods provide estimates for f only assuming some sort of smoothness.
- Three of most popular methods require a choice of user-specified paramaters
  - Kernal smoother: bandwidth
  - Smoothing spline: smoothing parameter
  - Regression splines: number and placement of knots
- If the fits are sensitive to these choices, inference about the regression function will be problematic.

Shape-Restricted Regression

- Monotonic regression
  - A Closed-form solution was provided by Brunk (1955).
  - Pool adjacent violators algorithm (PAVA)
  - Estimator is a step function.
- Convex (concave) regression
  - Convex least square estimates proposed by Hildreth (1954).
  - Hanson and Pledger (1976) proved their consistency.
  - Finding least-squares estimator is a quadratic programming problem.
  - Estimator is a piecewise linear function.
- The fits do not require user-specified parameters. However, neither is satisfactory if f is known to be smooth.

Motivating Examples

- Example 1: Age and income
  - The relationship between log(income) and age of 205 Canadian workers is to be modeled nonparametrically.
  - It is believed that the true relationship between log(income) and age should be concave without any dip.
- Example 2: Onion data
  - When onions are planted more densely, the yield per plant is expected to decrease.
  - The log(yield) is supposed to be convex in the density of the planting.
- The example datasets are both from Ruppert, Wand, and Carroll (2003).

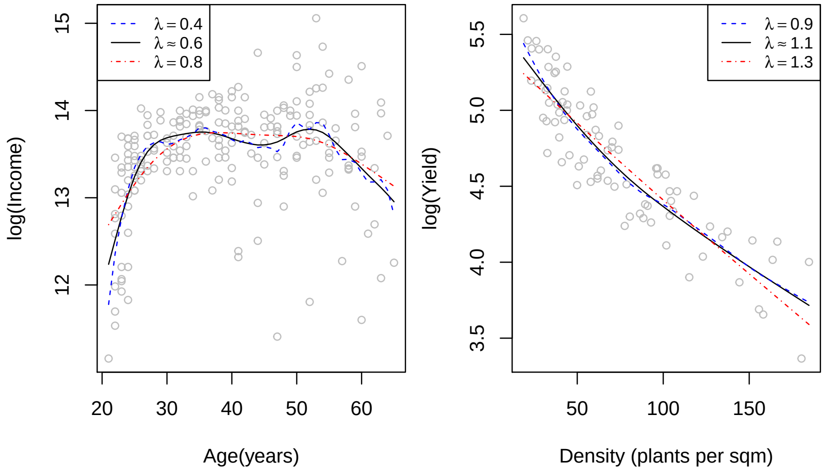

Figure 1: Age–income data (left panel) and onion data (right panel) fitted by smoothing splines with different smoothing parameters.

Smoothness and Shape Assumption

Monotone Regression Splines

- As the integrals of M-splines, the I-splines proposed by Ramsay (1988) are constrained to be monotonic.
- Similar to B-spline, the M-splines can be given recursively. The relationship between B-splines and M-spline of order k can be shown as follows:

$$B_i = (t_{i+k} - t_i)M_i/k, \;\; i = 1, 2, \ldots, n.$$

- To constrain the estimator to be monotone, the coefficients of the basis functions must be nonnegative (the coefficient of the constant function is not constrained).

Shape-Restricted Regression Splines

- Meyer (2008) extended the method to convex restrictions using C-splines where C-splines are integrals of I-splines.
- A convex regression function is estimated using linear combinations of the basis functions with nonnegative coefficients, plus an unrestricted linear combination of the constant function and the identity function $g(x) = x$.
- If the underlying regression function is both increasing and convex, the coefficient on the identity function is restricted to be nonnegative as well.

Estimation Procedures

Algorithm for regression spline fitting

- Ramsay (1988) provided an iterative gradient-based algorithm for estimating f with smoothed monotone regression using I-splines. It converges in "infinitely many" steps, which means that the true solution is approached asymptotically and reached within a user-defined tolerance.
- Meyer (1996) proposed the hinge algorithm for cone projection problem, which can be applied to estimating f under shape and smoothing assumption. In addition, it is guaranteed to produce the solution in a finite number of steps.

Revisit: Shape Restriction Regression

- Consider model in vector form:

$$\boldsymbol{y} = \boldsymbol{\theta} + \sigma\boldsymbol{\epsilon},$$

where $\theta_i = f(x_i)$ and x_i is ordered increasingly, $i = 1, \ldots, n$.

- If f is known to be increasing and convex, the shape restrictions are a set of linear inequality constrains:

$$\theta_1 \leq \theta_2 \leq \ldots \leq \theta_n,$$

and

$$\frac{\theta_2 - \theta_1}{x_2 - x_1} \leq \frac{\theta_3 - \theta_2}{x_3 - x_2} \leq \ldots \leq \frac{\theta_n - \theta_{n-1}}{x_n - x_{n-1}}.$$

Cone Projection

- The shape restrictions can be rewritten in the form $\boldsymbol{A\theta} \geq \mathbf{0}$. The constraint matrix $\boldsymbol{A}$ is $m \times n$ where $m = n - 1$ for monotone and $m = n - 2$ for convex constraints.
- The estimation problem is to find $\boldsymbol{\theta}$ to minimize $||\boldsymbol{y} - \boldsymbol{\theta}||^2$ under constraints $\boldsymbol{A\theta} \geq \mathbf{0}$.
- The m inequality constraints form a convex polyhedral cone $\mathcal{C}$ in $\mathcal{R}^n$ expressed as

$$\mathcal{C} = \{\boldsymbol{\theta} \in \mathcal{R}^n : \boldsymbol{A\theta} \geq 0\}.$$

- Let $\mathcal{V}$ be the null space of $\boldsymbol{A}$. Then $\mathcal{V}$ is contained in $\mathcal{C}$.

- Meyer (1999) showed that the cone can be alternatively written as

$$\mathcal{C} = \left\{ \boldsymbol{\theta} \in \mathcal{R}^n : \boldsymbol{\theta} = \boldsymbol{v} + \sum_{j=1}^{M} b_j \boldsymbol{\delta}_j,\ \ \boldsymbol{v} \in \mathcal{V},\ \ b_1, \ldots, b_M \geq 0 \right\},$$

where $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M$ is the edges or generators of $\mathcal{C}$, $M = m$ if $\boldsymbol{A}$ has full row rank and $M \geq m$ otherwise.

- In addition, the edges $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M$ are orthogonal to $\mathcal{V}$, so the projection of $\boldsymbol{Y}$ onto $\mathcal{C}$ is the sum of the projection onto $\mathcal{V}$ and onto the cone

$$\Omega = \left\{ \boldsymbol{\theta} \in \mathcal{R}^n : \sum_{j=1}^{M} b_j \boldsymbol{\delta}_j,\ \ b_1, \ldots, b_M \geq 0 \right\}$$

The Hinge Algorithm

- The hinge algorithm provides the projection onto Ω by determining the face of the cone on which the projection lands.
- Once the face containing the projection is determined, the projection onto Ω is simply the projection onto the linear space spanned by the edges making up the face.
- For more details and proofs, see Meyer (1999) and Meyer (2013).

Back to Shape-Restricted Regression Splines

- These ideas applied to shaped-restricted regression can be applied to regression splines as well.
- A quadratic spline function is increasing if and only if it is increasing at the knots. Similarly, a cubic spline function is convex if and only if it is convex at the knots (Meyer 2013).
- The spline basis functions are the edges of constraint cone.

Implementation

R Package: **ConSpline** and **coneproj**

- **ConSpline** provides function `conspline` fitting the partial linear model

$$\boldsymbol{y} = \boldsymbol{\theta} + \boldsymbol{Z}\boldsymbol{\alpha} + \boldsymbol{\epsilon},$$

  where $\theta_i = f(x_i),\ i = 1, \ldots, n$, $\boldsymbol{Z}$ is an optional design matrix, f is a smooth function with a user-defined shape: increasing, decreasing, convex, concave, or combinations of monotonicity and convexity.
- **ConSpline** depends on **coneproj** (Liao and Meyer 2014). The latter mainly contains routines for cone projection and quadratic programming.

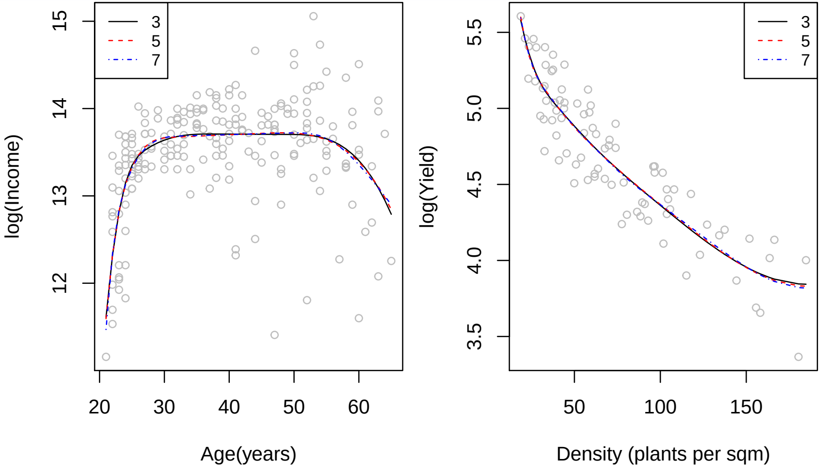

Figure 2: Age–income data (left panel) and onion data (right panel) fitted by shape-restricted regression splines with different number of internal knots.

Summary

Summary

- Constrained estimation is useful for both parametric and nonparametric function estimation.
- More constrain means less freedom, which may lead to more robust estimates.
- When shape assumptions can be combined with smoothness, we may get the best of both worlds.
- Shape-restricted regression splines are a great combination of smoothness and shape assumption.
- Source document of this slide is available at *github*.

Reference

Reference I

Brunk, Hugh D. 1955. "Maximum Likelihood Estimates of Monotone Parameters." *The Annals of Mathematical Statistics*. JSTOR, 607–16.

Hanson, DL, and Gordon Pledger. 1976. "Consistency in Concave Regression." *The Annals of Statistics*. JSTOR, 1038–50.

Hildreth, Clifford. 1954. "Point Estimates of Ordinates of Concave Functions." *Journal of the American Statistical Association* 49 (267). Taylor & Francis Group: 598–619.

Liao, Xiyue, and Mary C Meyer. 2014. "coneproj: An R Package for the Primal or Dual Cone Projections with Routines for Constrained Regression." *Journal of Statistical Software* 61 (1): 1–22.

Meyer, Mary C. 1996. "Shape Restricted Inference with Applications to Nonparametric Regression, Smooth Nonparametric Function Estimation, and Density Estimation." PhD thesis,

Reference II

University of Michigan.

———. 1999. "An Extension of the Mixed Primal–Dual Bases Algorithm to the Case of More Constraints Than Dimensions." *Journal of Statistical Planning and Inference* 81 (1). Elsevier: 13–31.

———. 2008. "Inference Using Shape-Restricted Regression Splines." *The Annals of Applied Statistics*. JSTOR, 1013–33.

———. 2013. "A Simple New Algorithm for Quadratic Programming with Applications in Statistics." *Communications in Statistics-Simulation and Computation* 42 (5). Taylor & Francis: 1126–39.

Ramsay, James O. 1988. "Monotone Regression Splines in Action." *Statistical Science*. JSTOR, 425–41.

Ruppert, David, Matt P Wand, and Raymond J Carroll. 2003. *Semiparametric Regression*. 12. Cambridge university press.