# Spling Smoothing and Tuning Parameter Selection

Eun Ryung Lee

SKKU

October 1, 2018

# Introduction

- Instead of just considering input variables $\mathbf{x}$ (potentially mult.),

  augment/replace with transformations = "input features"
- Linear basis expansions

  - maintain linear form in terms of these transformations

$$f(\mathbf{x}) = \sum_{j=1}^{p} \beta_j h_j(\mathbf{x})$$

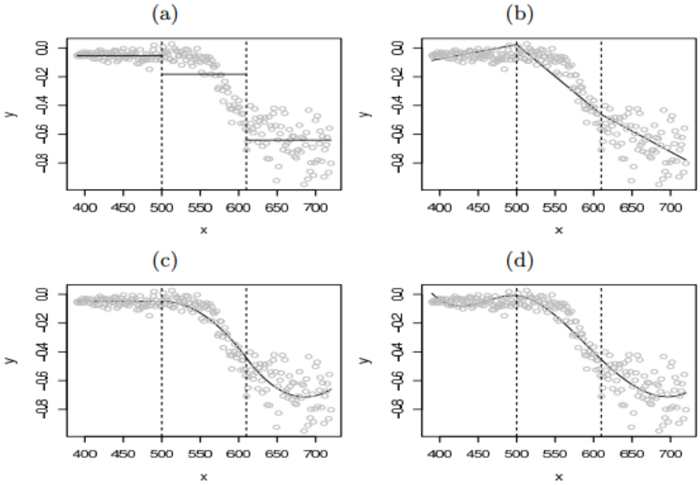
- What transformations should we use?
  - $h_j(\mathbf{x}) = x_j$
  - $h_j(\mathbf{x}) = x_j^2$ or $h_j(\mathbf{x}) = x_j x_k$
  - $h_j(\mathbf{x}) = I(L_j \leq x_j \leq U_j)$
  - $\cdots$

# Piecewise Polynomials Fits

- Assume that $\mathbf{x}$ is univariate
- Polynomial fits are often good locally, but not globally
  - Adjusting coefficients to fit one region can make the function go wild in other regions
- Consider `piecewise polynomial` fits
  - Local behavior can often be well approximated by low-order polynomials

# Piecewise Polynomials Fits

LIDAR Data Example



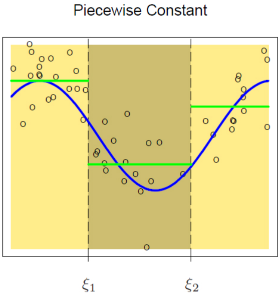
# Piecewise Constant Fits

Example 1: Piecewise constant, with 3 basis functions

- Basis functions

  $h_1(x) =$ , $h_2(x) =$ , $h_3(x) =$

- Resulting model: $f(x) = \sum_{j=1}^{3} \beta_j h_j(x)$
- Fit: Take the mean of the data in each region



# Piecewise Linear Fits

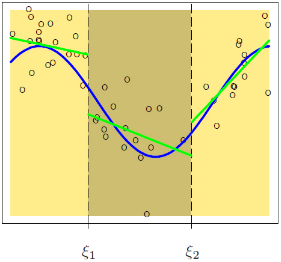
Example 1: Piecewise linear, with 6 basis functions

- Basis functions

$$h_{j+3} = h_j(x)x, \;\; j = 1, 2, 3$$

- Resulting model: $f(x) = \sum_{j=1}^{6} \beta_j h_j(x)$
- Fit: fit the linear model with the data in each region

Piecewise Linear



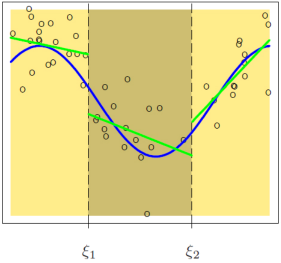
# Linear Splines

- Resulting piecewise linear model:

$$
\begin{aligned}
f(x) \quad = \quad & I(x < \xi_1)(\beta_1 + \beta_4 x) + I(\xi_1 \leq x < \xi_2)(\beta_2 + \beta_5 x) \\
& + I(\xi_2 \leq x)(\beta_3 + \beta_6 x)
\end{aligned}
$$

- Typically prefer continuity
  - Enforce
  - Which implies
  - # of parameters?

Piecewise Linear



# Linear Splines

- More directly, we can use the `truncated power` basis

  $$h_1(x) = 1$$
  $$h_2(x) = x$$
  $$h_3(x) = (x - \xi_1)_+$$
  $$h_4(x) = (x - \xi_2)_+$$

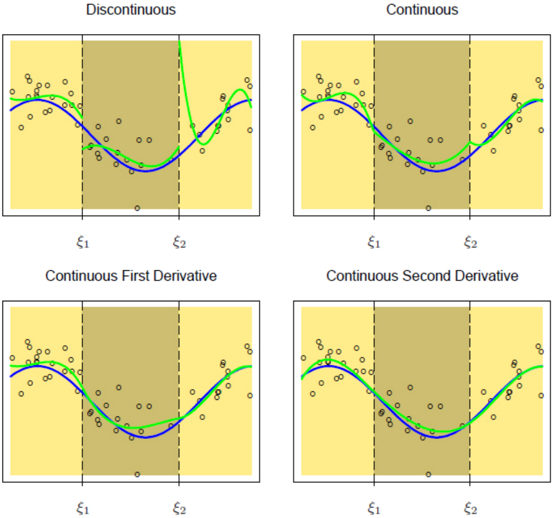
- Resulting model:

$$f(x) = \beta_0 + \beta_1 x + \beta_2 (x - \xi_1)_+ + \beta_3 (x - \xi_2)_+$$

# Quadratic and Cubic Splines

- Quadratic spline
  - $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + b_1(x - \xi_1)_+^2 + b_2(x - \xi_2)_+^2$
  - Has continous first derivative
- Cubic spline
  - $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + b_1(x - \xi_1)_+^3 + b_2(x - \xi_2)_+^3$
  - Has continous first and second derivatives

# Cubic Spline Fit



# Cubic Spline Fit

- Discontinuous cubic spline

$$\begin{aligned} f(x) &= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \\ &\quad b_0 + b_1(x-\xi_1) + b_2(x-\xi_1)^2 + b_3(x-\xi_1)_+^3 \\ &\quad c_0 + c_1(x-\xi_2) + c_2(x-\xi_2)^2 + c_3(x-\xi_2)_+^3 \end{aligned}$$

- Continuous cubic spline

$$\begin{aligned} f(x) &= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \\ &\quad + b_1(x-\xi_1) + b_2(x-\xi_1)^2 + b_3(x-\xi_1)_+^3 \\ &\quad + c_1(x-\xi_2) + c_2(x-\xi_2)^2 + c_3(x-\xi_2)_+^3 \end{aligned}$$

# Cubic Spline Fit

- Continously differentiable cubic spline

$$
\begin{aligned}
f(x) &= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \\
&\quad + b_2 (x - \xi_1)^2 + b_3 (x - \xi_1)_+^3 \\
&\quad + c_2 (x - \xi_2)^2 + c_3 (x - \xi_2)_+^3
\end{aligned}
$$

- Continuously twice differentiable cubic spline

$$
f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + b_3 (x - \xi_1)_+^3 c_3 (x - \xi_2)_+^3
$$

# Cubic Splines as Linear Smoothers

- Cubic spline function with $K$ knots:

$$g(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{k=1}^{K} b_k (x - \xi_k)_+^3$$

- The true model : $Y = f(X) + \epsilon$ with $E(\epsilon) = 0$
- Assume that $f(x) \approx g(x)$ in the region of our interest
- Design matrix
- Estimator
- Linear smoother

# Space of polynomial splines

- Definition

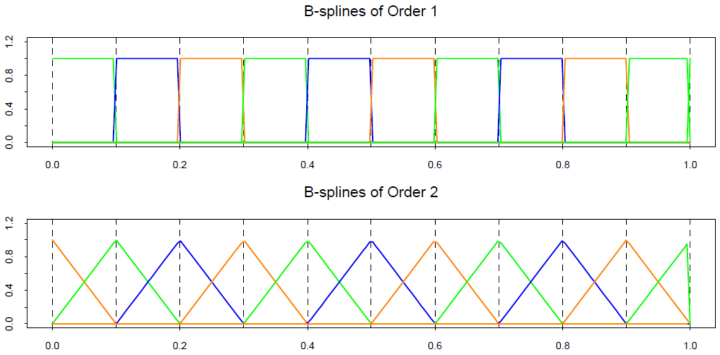
  An degree-$d$ spline with internal knots $\xi_1 < \xi_2 \cdots < \xi_K$ is a piecewise $d$ degree polynomial with $d-2$ continuous derivatives at the knots.

- Basis for the space of polynomial splines of order $M$

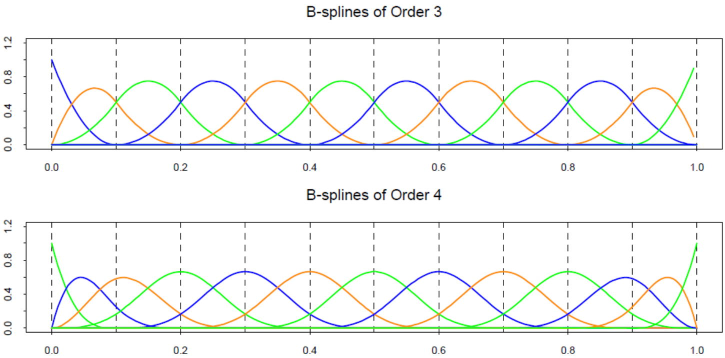
  - Power basis : $1,\ x, x^2, \ldots, x^d, (x-\xi_1)_+^d, \ldots, (x-\xi_K)_+^d$
  - B-spline basis : defined recursively in terms of divided differences starting from the Haar basis functions
  - Dimension : $K + d + 1$

- Choices
  - Order of spline
  - Number of knots
  - Placement of knots

# B-Spline Basis - Order 1 and 2(Linear)



# B-Spline Basis - Order 3(Quadratic) and 4(Cubic)



# Cubic B-Splines as Linear Smoothers

- Cubic B-spline with $K$ knots has the basis expansion:
- The true model : $Y = f(X) + \epsilon$ with $E(\epsilon) = 0$
- Assume that $f(x) \approx g(x)$ in the region of our interest
- Design matrix
- Computational gain:

# Smoothing Splines

Penalized Least Square: for a given $\lambda \geq 0$, find the function $f$ which minimizes

$$\sum_{i=1}^{n} \{Y_i - f(X_i)\}^2 + \lambda \int \{f''(x)\}^2 \, dx \tag{1}$$

over some class $\mathcal{F}$ of smooth functions.

- $\sum_{i=1}^{n} \{Y_i - f(X_i)\}^2$: RSS
- $\int \{f''(x)\}^2 \, dx$ : a penalty for the roughness of the function
- similar to a Ridge regression
- It is known that the minimizer $\hat{f}$ is a natural cubic spline, with knots at each sample point $X_1, \ldots, X_n$

# Natural Spline

- Definition
  A spline function of order $2k-1$ is called a *natural spline of order* $2k-1$ if it is a polynomial of order $k-1$ beyond the boundary knots $\xi_1$ and $\xi_K$.
- Dimension: $K$
- Power basis for the space of natural cubic splines:

$$1,\ x,\ d_i(x) - d_{n-1}(x), \quad i = 1, \ldots, n-2$$

  where

$$d_i(x) = \frac{(x - X_i)_+^3 - (x - X_n)_+^3}{X_n - X_i}.$$

- Several choices of basis for the space of natural splines are provided in Section 5.3.3 of
  Eubank (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker, Inc.

# Estimation

- Recall that the minimizer $\hat{f}$ of the penalized criterion 1 for smoothing spline is a natural cubic spline, with knots at each sample point $X_1, \ldots, X_n$.
- Deriving smoothing spline:

# Smoothing Spline as a Linear Smoother

1. Spline Smoothing
   - Regression Splines
   - Smoothing Splines: Roughness Penalty

2. Regularization/Tuning Parameter Selection

# Regularization/Tuning Parameter in Nonparameter Regression

- Consider (univariate) nonparametric regression model $Y_i = f(X_i) + \epsilon_i$ for $i = 1, \ldots, n$.

- In the above nonparametric regression estimation methods for estimating $f$, there exists a regularization tuning parameter, which controls degree of smoothing i.e.,model complexity:
  - Kernel regression:
  - Regression spline:
  - Smoothing spline:

# Regularization/Tuning Parameter Selection

- Note that the above estimation methods are all a linear smoother.
- What can one use for selecting the tuning parameter in each problem?