

The set of basis functions in eg 7.9 are:

$$\begin{aligned}
 h_1(x) &= 1 \\
 h_2(x) &= x \\
 h_3(x) &= x^2 \\
 h_4(x) &= x^3 \\
 h_5(x) &= (x - \xi_1)_+^3 \\
 h_6(x) &= (x - \xi_2)_+^3
 \end{aligned}
 \quad
 \left.
 \begin{aligned}
 h_j(x) &= x^{j-1}, & j &= 1, 2, \dots, M=4 \\
 h_{n+l}(x) &= (x - \xi_l)_+^{n-1}, & l &= 1, \dots, K=2
 \end{aligned}
 \right\}
 \text{truncated-power basis functions}$$

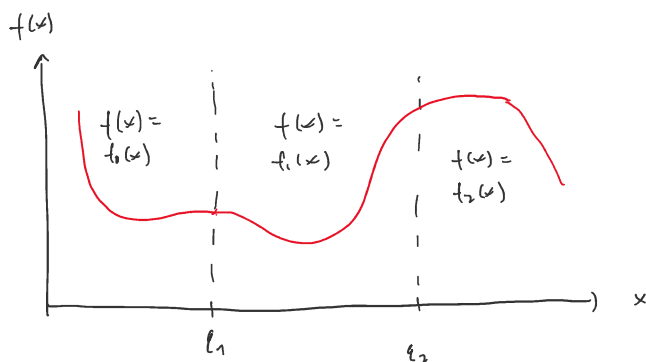
The associated basis-model:

$$\begin{aligned}
 f(x) &= \sum_{j=1}^{M+K} \beta_j h_j(x) \\
 &= \sum_{j=1}^M \beta_j x^{j-1} + \sum_{l=1}^K \beta_{n+l} (x - \xi_l)_+^{n-1}
 \end{aligned}
 \quad (*)$$

with $M=4$ and $K=2$.

We are asked to show that the basis model (*) represents a cubic spline.

A cubic spline is a piecewise third-order (cubic) polynomial with continuous first and second order derivatives.



Thus we need to show that we can rewrite (*) as

$$f(x) = \begin{cases} f_0(x) & \text{if } x < \xi_1 \\ f_1(x) & \text{if } \xi_1 \leq x < \xi_2 \\ f_2(x) & \text{if } \xi_2 \leq x \end{cases} \quad (*)$$

where each $f_i(x)$ is a cubic polynomial:

$$f_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i,$$

such that the following continuity restrictions are satisfied:

$$f_{i-1}(\xi_i) = f_i(\xi_i)$$

$$f'_{i-1}(\xi_i) = f'_i(\xi_i)$$

$$f''_{i-1}(\xi_i) = f''_i(\xi_i)$$

} these restrict the coefficients a_i, b_i, c_i, d_i

(i) Show that (*) defines a set of cubic polynomials:

For $x < \xi_1$:

$$f(x) = \underbrace{\beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3}_{\text{a cubic polynomial}} + \underbrace{\sum_{k=1}^{\infty} \beta_{4+k} (x - \xi_1)^{k-1}}_{=0}$$

$$\equiv f_0(x)$$

For $\xi_1 \leq x < \xi_2$:

$$f(x) = f_0(x) + \beta_{4+1} (x - \xi_1)^3 + 0$$

$$\begin{aligned} (x - \xi_1)^3 &= x^3 - 3\xi_1 x^2 + 3\xi_1^2 x - \xi_1^3 \\ &= (\beta_1 - \xi_1^3) + (\beta_2 + 3\beta_{4+1}\xi_1^2)x + (\beta_3 - 3\beta_{4+1}\xi_1)x^2 + (\beta_4 + \beta_{4+1})x^3 \end{aligned}$$

$$= d_1 + c_1 x + b_1 x^2 + a_1 x^3$$

$$= f_1(x)$$

in
a cubic polynomial

For $x \geq \xi_2$:

$$f(x) = f_1(x) + \underbrace{\beta_{4+2} (x - \xi_2)^3}_{\text{also a cubic polynomial}} = f_2(x)$$

Hence, we can indeed rewrite $f(x)$ as a piecewise polynomial:

$$f(x) = \sum_{j=1}^{4+n} \beta_j h_j(x) = \begin{cases} f_0(x) & x < \xi_1 \\ f_1(x) & \xi_1 \leq x < \xi_2 \\ f_2(x) & x \geq \xi_2 \end{cases}$$

Where

$$f_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i = f_{i-1}(x) + \beta_{4+i} (x - \xi_i)^3$$

which implies

$$d_i = d_{i-1} - \beta_{4+i} \xi_i^3$$

$$c_i = c_{i-1} + 3 \beta_{4+i} \xi_i^2$$

$$b_i = b_{i-1} - 3 \beta_{4+i} \xi_i$$

$$a_i = a_{i-1} - \beta_{4+i}$$

$$\Leftrightarrow \beta_{4+i} = -(a_i - a_{i-1})$$

(ii) Show that $f(x)$, $f'(x)$ and $f''(x)$ are continuous in each knot

This follows from

$$h_{m+1}(\xi_i) = h'_{m+1}(\xi_i) = h''_{m+1}(\xi_i) = 0$$

which implies that for all knots ξ_i :

$$f_i(\xi_i) = f_{i-1}(\xi_i) + \beta_{m+1}(\xi_i - \xi_i)^3 = f_{i-1}(\xi_i)$$

$$f'_i(\xi_i) = f'_{i-1}(\xi_i) + 3\beta_{m+1}(\xi_i - \xi_i)^2 = f'_{i-1}(\xi_i)$$

$$f''_i(\xi_i) = f''_{i-1}(\xi_i) + 6\beta_{m+1}(\xi_i - \xi_i) = f''_{i-1}(\xi_i)$$

ESL Ex. 5.7

For a sequence $\{x_i, y_i\}_{i=1}^n$, where $a < x_1 < \dots < x_n < b$, we want to show that the natural cubic spline $q(x)$ with knots at each x_i minimizes the penalized RSS:

$$RSS(f, \lambda) = \sum_{i=1}^n (y_i - f(x_i))^2 + \lambda \int f''(t)^2 dt$$

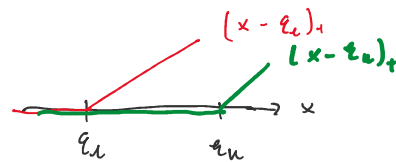
The natural cubic spline is a basis model

$$q(x) = \sum_{j=1}^n N_j(x) \theta_j$$

where the H basis functions are defined as

$$\begin{aligned} H_1(x) &= 1 \\ H_2(x) &= x \\ H_{d+2}(x) &= d_d(x) - d_k(x) \quad \text{for } d = 1 \dots k-2 \end{aligned}$$

$$\text{and } d_d(x) = \frac{(x - \xi_d)_+^3 - (x - \xi_k)_+^3}{\xi_k - \xi_d}$$



We will make use of the fact that $q''(x) = q'''(x) = 0$ for $x < x_1 = \xi_1$ and $x \geq x_n = \xi_k$.

To see this, note that:

$$d_d''(x) = \frac{6[(x - \xi_d)_+ - (x - \xi_k)_+]}{\xi_k - \xi_d} = \begin{cases} 0 & \text{if } x < \xi_d \leq \xi_k \\ \frac{(x - \xi_d)}{\xi_k - \xi_d} & \text{if } \xi_d \leq x < \xi_k \\ \frac{(x - \xi_d) - (x - \xi_k)}{\xi_k - \xi_d} & \text{if } x \geq \xi_k \end{cases}$$

$\underbrace{\hspace{10em}}_{=1}$

Hence

$$N_{d+2}''(x) = \begin{cases} 0 & \text{if } x < \xi_d < \xi_{k-1} \\ \dots & \\ 1-1 & \text{if } x \geq \xi_k \end{cases}$$

and therefore

$$\text{also } N_1''(x) = N_2''(x) = 0$$

and therefore

$$q''(x) = \sum_{j=1}^N H_j''(x) \theta_j = \begin{cases} 0 & \text{if } x < \xi_1 \\ \dots & \\ 0 & \text{if } x \geq \xi_N \end{cases}$$

(so $N_1''(x) = N_2''(x) = 0$)

Similarly,

$$d''_2(x) = \begin{cases} 0 & \text{if } x < \xi_1 \\ \dots & \\ 0 & \text{if } x \geq \xi_N \end{cases}$$

$$q'''(x) = 0 \quad \text{if } x < \xi_1 \quad \text{or} \quad x \geq \xi_N$$

Recall that $\xi_1 = x_1 > a$ and $\xi_N = x_N < b$.

- (u) Suppose $q(x)$ interpolates $\{y_i\}_{i=1}^N$ exactly: $q(x_i) = y_i$ for $i=1 \dots N$
 Let $\tilde{q}$ be any other differentiable function that interpolates the N pairs, such that $\tilde{q}(x_i) = q(x_i)$.
 Define $h(x) = \tilde{q}(x) - q(x)$.

Then

Integration by parts:

$$\int u dv = uv - \int v du$$

$$\begin{aligned} \int_a^b q''(x) h''(x) dx &= [q''(x) h'(x)]_a^b - \int_a^b q'''(x) h'(x) dx \\ &= \underbrace{q''(b) h'(b)}_{=0} - \underbrace{q''(a) h'(a)}_{=0} \\ &\quad - \int_a^{x_1} \underbrace{q'''(x) h'(x)}_{=0} dx + \int_{x_1}^{x_2} q'''(x) h'(x) dx + \dots + \int_{x_{N-1}}^b \underbrace{q'''(x) h'(x)}_{=0} dx \\ &= - \sum_{j=1}^{N-1} \int_{x_j}^{x_{j+1}} q'''(x) h'(x) dx \\ &= - \sum_{j=1}^{N-1} \left([q'''(x) h(x)]_{x_j}^{x_{j+1}} - \int_{x_j}^{x_{j+1}} \underbrace{q^{(4)}(x) h(x)}_{=0} dx \right) \\ &= - \sum_{j=1}^{N-1} [\underbrace{q'''(x_{j+1}) h(x_{j+1})}_{=0} - \underbrace{q'''(x_j) h(x_j)}_{=0}] \end{aligned}$$

since $q''(x) = q'''(x) = 0$ for $x < x_1$ and $x \geq x_N$

$$= 0$$

(b) Here

$$\tilde{q}''(x) = h'(x) + q'(x)$$

and

$$\begin{aligned} \int_a^b \tilde{q}''(x)^2 dx &= \int_a^b h''(x)^2 dx + \int_a^b q''(x)^2 dx + 2 \underbrace{\int_a^b q''(x) h''(x) dx}_{=0 \text{ by (c)}} \\ &\geq \int_a^b q''(x)^2 dx \end{aligned}$$

The equality holds only if $h''(x) = 0$ for all $x \in [a, b]$, which means that $h(x)$ have to be linear in $[a, b]$. In combination with $h(x_i) = 0$ for all x_i , this means that the equality holds only if $h(x) = 0$ for all $x \in [a, b]$, i.e. if $\tilde{q}(x) = q(x)$.

(c) Suppose $\tilde{f}$ is a minimizer of the penalized RSS. Let f be a natural cubic spline with knots at each x_i , such that it has the same function value as $\tilde{q}$ at each knot:

$$f(x_i) = \tilde{f}(x_i) \Rightarrow \sum_{i=1}^N (y_i - f(x_i))^2 = \sum_{i=1}^N (y_i - \tilde{f}(x_i))^2$$

Since $\tilde{f}$ is a minimizer, we have then that

$$\begin{aligned} \text{RSS}(\tilde{f}, \lambda) &\leq \text{RSS}(f, \lambda) \Rightarrow \\ \lambda \int_a^b \tilde{f}''(t) dt &\leq \lambda \int_a^b f''(t) dt \end{aligned}$$

From (b), we know that

$$\int_a^b \tilde{f}''(t) dt \geq \int_a^b f''(t) dt$$

Thus, the equality must hold.

Also from (b), we know that this is only the case if $f(x) = \tilde{f}(x)$ for all $x \in [a, b]$.

Hence, the cubic spline f is a minimizer of $\text{RSS}(f, \lambda)$.