

ESL Ex. 5.7

Updated 25/03/25

Based on: https://waxworksmath.com/Authors/G_M/Hastie/WriteUp/Weatherwax_Epstein_Hastie_Solution_Manual.pdf

For a sequence $\{x_i, y_i\}_{i=1}^n$, where $a \leq x_1 < \dots < x_n \leq b$, we want to show that the solution to the constrained minimization problem

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

is a natural cubic spline with knots at each $x_1, \dots, x_n$

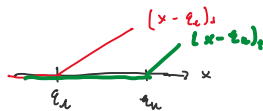
The natural cubic spline is a basis model

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

where the H basis functions are defined as

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

$$\text{and } d_d(x) = \frac{(x - \tau_d)_+^3 - (x - \tau_{d+1})_+^3}{\tau_{d+1} - \tau_d}$$



We will make use of the fact that $q''(x) = q'''(x) = 0$ for $x < x_1 = \tau_1$ and $x \geq x_n = \tau_n$.

To see this, note that:

$$d_d'(x) = 3 \left[\frac{(x - \tau_d)_+^2 - (x - \tau_{d+1})_+^2}{\tau_{d+1} - \tau_d} \right]$$

$$d_d''(x) = 6 \left[\frac{(x - \tau_d)_+ - (x - \tau_{d+1})_+}{\tau_{d+1} - \tau_d} \right] = \begin{cases} 0 & \text{if } x < \tau_d < \tau_{d+1} \\ \frac{(x - \tau_d)}{\tau_{d+1} - \tau_d} & \text{if } \tau_d \leq x < \tau_{d+1} \\ \frac{(x - \tau_d) - (x - \tau_{d+1})}{\tau_{d+1} - \tau_d} & \text{if } x \geq \tau_{d+1} \end{cases}$$

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

Hence

$$N_{d+2}(x) = \begin{cases} 0 & \text{if } x < \tau_d < \tau_{d+1} \\ \dots \\ 1 & \text{if } x \geq \tau_{d+1} \end{cases}$$

and therefore (also $N_1''(x) = N_2''(x) = 0$)

$$q''(x) = \sum_{j=1}^n H_j''(x) \theta_j = \begin{cases} 0 & \text{if } x < \tau_1 \\ \dots \\ 0 & \text{if } x \geq \tau_n \end{cases}$$

Similarly,

$$d_d''(x) = \begin{cases} 0 & \text{if } x < \tau_d \\ \dots \\ 0 & \text{if } x \geq \tau_{d+1} \end{cases}$$

$$q'''(x) = 0 \quad \text{if } x < \tau_1 \text{ or } x \geq \tau_n$$

Recall that $\tau_1 = x_1 > a$ and $\tau_n = x_n < b$.

- (a) 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 on $[a, b]$ 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:
 $\int u dv = uv - \int v du$

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

since $q'(x_i) = q''(x_i) = 0$ for $x < x_1$ and $x > x_N$

integration by parts again

(b) Here

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

and

$$\begin{aligned} \int_a^b \tilde{q}''(x)^2 dx &= \int_a^b h''(x)^2 dx + \int_a^b q''(x)^2 dx + 2 \int_a^b q''(x) h''(x) dx \\ &= \int_a^b q''(x)^2 dx \end{aligned}$$

$= 0$ by (a)

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 $\tilde{q}$ is a minimizer of the penalized RSS, such that

$$\sum_{i=1}^N (y_i - \tilde{q}(x_i))^2 + \lambda \int_a^b \tilde{q}''(t)^2 dt \leq \sum_{i=1}^N (y_i - q(x_i))^2 + \lambda \int_a^b q''(t)^2 dt$$

Now, since both $\tilde{q}$ and q interpolates $\{x_i, y_i\}_{i=1}^N$, that is $\tilde{q}(x_i) = q(x_i)$ for each $x_1 \dots x_N$, the squared errors are the same at both sides of the ineq.
 The above thus reduces to $\lambda \int_a^b \tilde{q}''(t)^2 dt \leq \lambda \int_a^b q''(t)^2 dt$

From (b), we know that

$$\int_a^b \tilde{q}''(t)^2 dt \geq \int_a^b q''(t)^2 dt$$

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

Hence, the minimizer $\tilde{q}$ must be a natural cubic spline with knots in each x_i .