

STK2100 Proposed Solutions

ESL Ex. 5.4

Vera Kvisgaard

24-03-25

Consider the truncated power series representation for cubic splines with K interior knots. Let

$$f(X) = \sum_{j=0}^p \beta_j X^j + \sum_{k=0}^K \theta_k (X - \xi_k)_+^3 \quad (\text{ESL 5.70})$$

where $p = 3$ and $(X - \xi)_+^3$ is the “positive cubic function”

$$(X - \xi)_+^3 = \begin{cases} 0 & \text{if } X < \xi, \\ (X - \xi)^3 & \text{otherwise} \end{cases}$$

The exercise asks us to:

1. Prove that the linearity-constraints posed on natural cubic splines implies the following restrictions on the coefficients in

$$\beta_2 = \beta_3 = 0 \text{ and } \sum_{k=0}^K \theta_k = \sum_{k=0}^K \theta_k \xi_k = 0 \quad (\text{ESL 5.71})$$

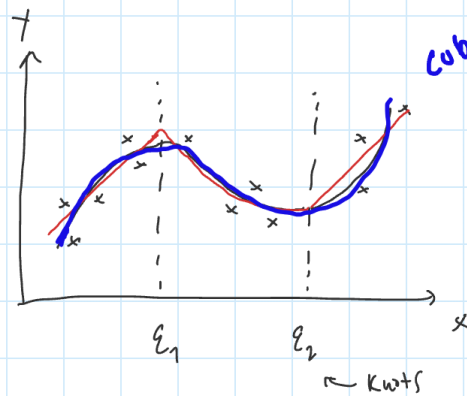
1. Derive the basis representation of natural cubic splines (ESL 5.4)-(ESL 5.5), starting from the restricted truncated power series above (ESL 5.70)-(ESL 5.71).
 - Swapping the starting point, we can also derive (ESL 5.70)-(ESL 5.71) from (ESL 5.4)-(ESL 5.5).

The solutions below are based on this solution manual by Weatherwax and Epstein

1. Prove constraints on the coefficients

A natural cubic spline is a cubic spline with a(n) (additional) constraint: the spline is linear beyond the boundary knots, that is for $X < \xi_1$ and $X > \xi_K$. To find the required restrictions on the coefficients, we visit these domains in turn:

Splines are piecewise polynomials



cubic spline
natural cubic spline

$$f(x) = \begin{cases} f_1(x) & \text{if } x < \xi_1 \\ f_2(x) & \text{if } \xi_1 < x < \xi_2 \\ f_3(x) & \text{if } \xi_2 < x \end{cases}$$

The spline basis representation:

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

↪ a set of basis functions

Figure 1: A natural cubic spline is a cubic spline that is linear beyond the boundary knots

1. For $X < \xi_1$, we have that:

$$f(X) = \sum_{j=0}^p \beta_j X^j = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$$

Thus, for f to be linear in this segment, we need the first two restrictions:
 $\beta_2 = \beta_3 = 0$.

2. For $X > \xi_K$, we have that (when $\beta_3 = \beta_2 = 0$):

$$f(X) = \beta_0 + \beta_1 X + \sum_{k=1}^K \theta_k (X - \xi_k)^3$$

since $(X - \xi_k)_+^3 = (X - \xi_k)^3$ for every k . Expanding these cubic terms, we get:

$$\begin{aligned} f(X) &= \beta_0 + \beta_1 X + \sum_{k=1}^K \theta_k (X^3 - 3\xi_k X^2 + 3\xi_k^2 X - \xi_k^3) \\ &= \beta_0 + \beta_1 X + X^3 \cdot \sum_{k=1}^K \theta_k - X^2 \cdot 3 \sum_{k=1}^K \theta_k \xi_k + X \cdot 3 \sum_{k=1}^K \theta_k \xi_k^2 - \sum_k \theta_k \xi_k^3 \end{aligned}$$

Thus, for f to be linear in this segment, we need the two last constraints:
 $\sum_{k=1}^K \theta_k = \sum_{k=1}^K \theta_k \xi_k = 0$

2a. Derive basis representation of natural cubic splines from the restricted cubic spline representation

First, consider the linear terms in equation (ESL 5.70-5.71), namely $\beta_0 + \beta_1 X$.

These are captured by the two first basis functions $N_1(X) = 1$ and $N_2(X) = X$.

Second, for the non-linear part, we can write

$$\sum_{k=1}^K \theta_k (X - \xi_k)_+^3 = \sum_{k=1}^{K-1} \theta_k (X - \xi_k)_+^3 + \theta_K (X - \xi_K)_+^3 = \sum_{k=1}^{K-1} \theta_k [(X - \xi_k)_+^3 - (X - \xi_K)_+^3]$$

The last equality follows from the linearity constraints in (ESL 5.71), that implies
 $\theta_K = -\sum_{k=1}^{K-1} \theta_k$.

Using

$$d_k(X) = \frac{(X - \xi_k)_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_k} \quad (\text{ESL 5.5})$$

we can re-write the above as

$$\sum_{k=1}^K \theta_k (X - \xi_k)_+^3 = \sum_{k=1}^{K-1} \theta_k (\xi_K - \xi_k) d_k(x)$$

or

$$\sum_{k=1}^K \theta_k (X - \xi_k)_+^3 = \sum_{k=1}^{K-2} \theta_k (\xi_K - \xi_k) d_k(x) + \theta_{K-1} (\xi_K - \xi_{K-1}) d_{K-1}(x)$$

Now, using the linearity constraints again, we can show that

$$\theta_{K-1} (\xi_K - \xi_{K-1}) = \theta_{K-1} \xi_K - \theta_{K-1} \xi_{K-1} = - \sum_{k=1}^{K-2} \theta_k (\xi_K - \xi_k).$$

This follows from

$$\begin{aligned} \theta_{K-1} \xi_{K-1} &= - \left[\sum_{k=1}^{K-2} \theta_k \xi_k + \theta_K \xi_K \right] \\ &= - \left[\sum_{k=1}^{K-2} \theta_k \xi_k - \xi_K \left(\sum_{k=1}^{K-2} \theta_k + \theta_{K-1} \right) \right] \\ &= \sum_{k=1}^{K-2} \theta_k (\xi_K - \xi_k) + \theta_{K-1} \xi_K. \end{aligned}$$

Hence,

$$\sum_{k=1}^K \theta_k (X - \xi_k)_+^3 = \sum_{k=1}^{K-2} \theta_k (\xi_K - \xi_k) d_k(x) - \sum_{k=1}^{K-2} \theta_k (\xi_K - \xi_k) d_{K-1}(x) = \sum_{k=1}^{K-2} \theta_k N_{k+2}(X)$$

Thus, under the linearity constraints in (ESL 5.71), we have that

$$f(X) = \sum_{j=0}^p \beta_j X^j + \sum_{k=0}^K \theta_k (X - \xi_k)_+^3 = \beta_0 N_1(X) + \beta_2 N_2(X) + \sum_{k=1}^{K-2} \theta_k N_{k+2}(X)$$

where the basis functions N_k are those given in (ESL 5.4):

$$\begin{aligned} N_1(X) &= 1 \\ N_2(X) &= X \\ N_{k+2}(X) &= d_k(X) - d_{K-1}(X), \quad k = 1, \dots, K-2 \end{aligned}$$

2b. Derive the restricted cubic spline representation from the basis representation

In the above, we have shown that the basis (ESL 5.4-5.5) is implied by the truncated power basis representation of a cubic spline (ESL 5.70) under the required linearity constraints in (ESL 5.71).

We can swap the starting-point and end-point, and show that we can derive the restricted cubic spline model in (ESL 5.70- 5.71) from the basis representation (ESL 5.4-5.5).

The basis representation of a natural cubic spline implies

$$f(X) = \tilde{\beta}_1 N_1(X) + \tilde{\beta}_2 N_2(X) + \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} N_{k+2}(X)$$

We want to find coefficients $\tilde{\beta}_j$ such that this model corresponds to that in (ESL 5.70- 5.71).

First, consider the linear parts of the two models. These are equal when

$$\beta_0 + \beta_1 X = \tilde{\beta}_1 N_1(X) + \tilde{\beta}_2 N_2(X) = \tilde{\beta}_1 + \tilde{\beta}_2 X,$$

and therefore $\tilde{\beta}_1 = \beta_0, \tilde{\beta}_2 = \beta_1$.

Next, consider the non-linear terms. Using $N_{k+2}(X) = d_k(X) - d_{K-1}(X)$, these can be decomposed into a *blue* and a *red* term as follows:

$$\sum_{k=1}^{K-2} \tilde{\beta}_{k+2} N_{k+2}(X) = \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} d_k(X) - d_{K-1}(X) \sum_{k=1}^{K-2} \tilde{\beta}_{k+2},$$

where (using (ESL 5.5))

$$\sum_{k=1}^{K-2} \tilde{\beta}_{k+2} d_k(X) = \sum_{k=1}^{K-2} \frac{\tilde{\beta}_{k+2}}{\xi_K - \xi_k} (X - \xi_k)_+^3 - (X - \xi_K)_+^3 \sum_{k=1}^{K-2} \frac{\tilde{\beta}_{k+2}}{\xi_K - \xi_k}$$

and

$$d_{K-1}(X) \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} = \frac{(X - \xi_{K-1})_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_{K-1}} \sum_{k=1}^{K-2} \tilde{\beta}_{k+2}.$$

Define $\theta_k = \tilde{\beta}_{k+2}/(\xi_K - \xi_k)$ for $k = 1, \dots, K-2$. Then

$$\sum_{k=1}^{K-2} \tilde{\beta}_{k+2} d_k(X) = \sum_{k=1}^{K-2} \theta_k (X - \xi_k)_+^3 - (X - \xi_K)_+^3 \sum_{k=1}^{K-2} \theta_k$$

and

$$\begin{aligned} d_{K-1}(X) \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} &= \frac{(X - \xi_{K-1})_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_{K-1}} \sum_{k=1}^{K-2} \theta_k (\xi_K - \xi_k) \\ &= \frac{(X - \xi_{K-1})_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_{K-1}} \left[\xi_K \sum_{k=1}^{K-2} \theta_k - \sum_{k=1}^{K-2} \theta_k \xi_k \right] \end{aligned}$$

Now, recall the constraints on the coefficients in (ESL 5.71):

$$\sum_{k=1}^K \theta_k = 0 \quad \Leftrightarrow \quad \sum_{k=1}^{K-2} \theta_k = -(\theta_{K-1} + \theta_K) \sum_{k=1}^K \theta_k \xi_k = 0 \quad \Leftrightarrow \quad \sum_{k=1}^{K-2} \theta_k \xi_k = -(\theta_{K-1} \xi_{K-1} + \theta_K \xi_K)$$

Using these constraints, we can rewrite the blue and the red part as follows:

$$\begin{aligned}\sum_{k=1}^{K-2} \tilde{\beta}_{k+2} d_k(X) &= \sum_{k=1}^{K-2} \theta_k (X - \xi_k)_+^3 + (\theta_{K-1} + \theta_K) (X - \xi_K)_+^3 \\ &= \sum_{k=1}^{K-2} \theta_k (X - \xi_k)_+^3 + \theta_K (X - \xi_K)_+^3 + \theta_{K-1} (X - \xi_K)_+^3\end{aligned}$$

and

$$\begin{aligned}d_{K-1}(X) \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} &= \frac{(X - \xi_{K-1})_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_{K-1}} [(\theta_{K-1} \xi_{K-1} + \theta_K \xi_K) - (\theta_{K-1} + \theta_K) \xi_K] \\ &= \frac{(X - \xi_{K-1})_+^3 - (X - \xi_K)_+^3}{\xi_K - \xi_{K-1}} [\theta_{K-1} (\xi_{K-1} - \xi_K)] \\ &= -\theta_{K-1} ((X - \xi_{K-1})_+^3 - (X - \xi_K)_+^3) \\ &= -(\theta_{K-1} (X - \xi_{K-1})_+^3 - \theta_{K-1} (X - \xi_K)_+^3)\end{aligned}$$

Putting everything *blue* and everything *red* back together, we have that

$$\sum_{k=1}^{K-2} \tilde{\beta}_{k+2} N_{k+2}(X) = \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} d_k(X) - d_{K-1}(X) \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} = \sum_{k=1}^K \theta_k (X - \xi_k)_+^3$$

Hence,

$$\tilde{\beta}_1 N_1(X) + \tilde{\beta}_2 N_2(X) + \sum_{k=1}^{K-2} \tilde{\beta}_{k+2} N_{k+2}(X) = \beta_0 + \beta_1 X + \sum_{k=1}^K \theta_k (X - \xi_k)_+^3$$

where $\sum_k \theta_k = 0$ and $\sum_k \theta_k \xi_k = 0$. $\times$