Solutions

March 3, 2023

Answer Q.1. I replicate figure 5.3 from ESL and the following is the result.

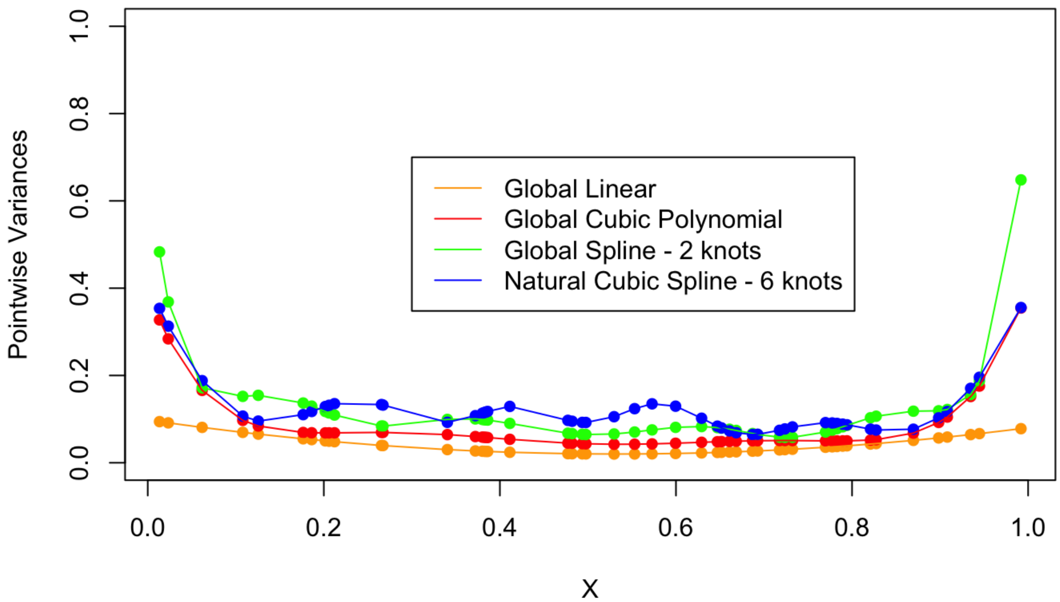

Figure 1: Replication of figure 5.3

Answer Q.2. We need to fit a logistic regression between predictor variable tobacco and response variable chd using South African heart disease dataset. And then calculate the standard deviation and plot the result within 1 standard deviation. I use build in function "glm" to fit logistic regression and the output contains variance covariance matrix. The following figures are the results corresponding to natural spline, B- spline and truncated polynomial spline. The shadowed area corresponds to 1 standard error but they are very small in the figure.

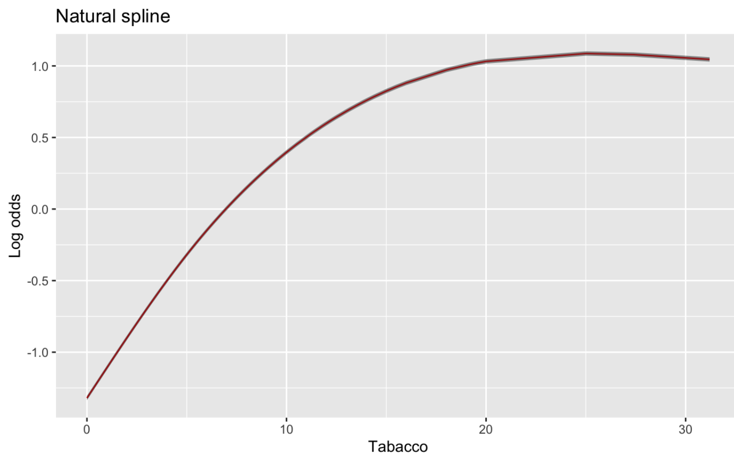

Figure 2: Natural spline

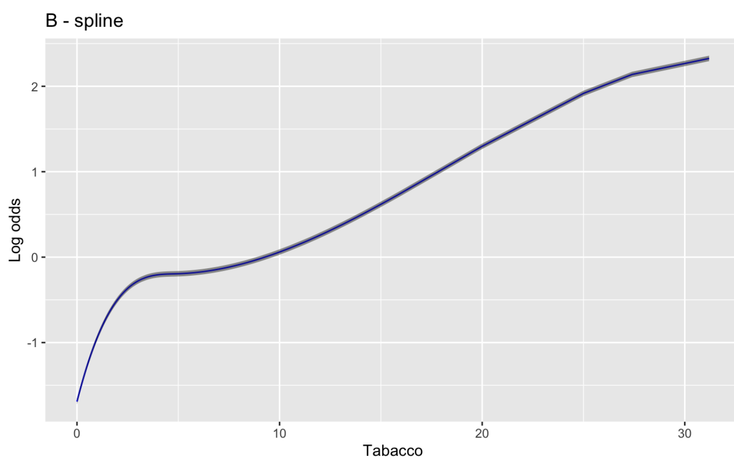

Figure 3: B-spline

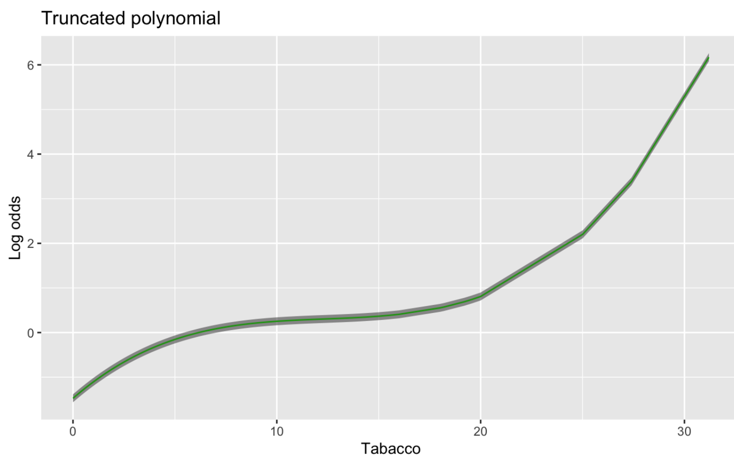

Figure 4: Truncated polynomial spline

Answer Q.3. In this question, I use modify the truncated polynomial function in the note to get the result. I add one more variable called "natural" which takes value TRUE or FALSE. If the input is FALSE, then the code is the same as in the note. Otherwise, we need to use a recursive formula to get what is called $N_{k+2}(x)$. The following is the result I get.

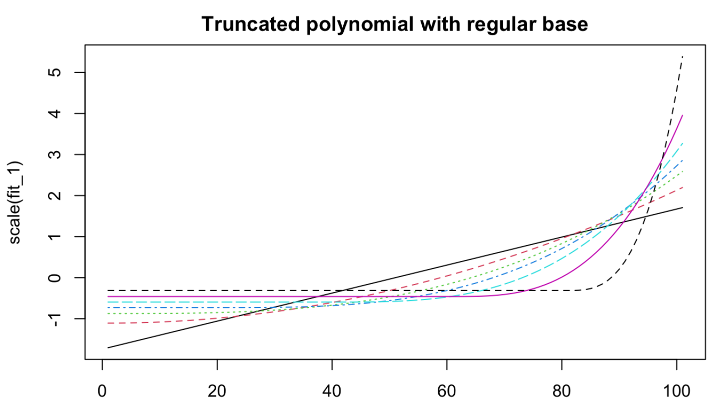

Figure 5: Truncated polynomial with regular base

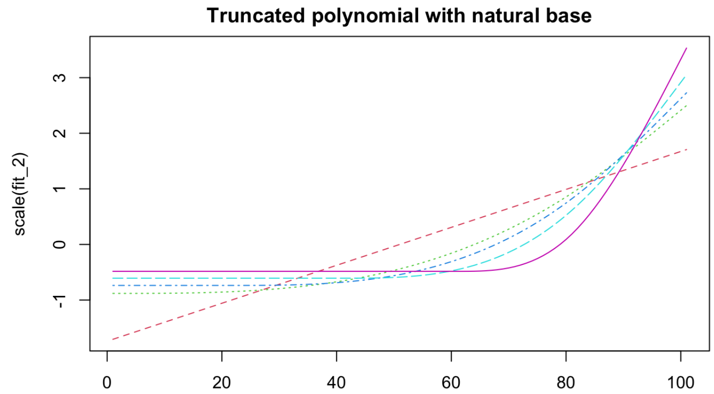

Figure 6: Truncated polynomial with natural base

Answer Q.4. a I simulate 100*100 points and add some noise to them, then define z which has trigonometric relationship between x and y. The following is the result.

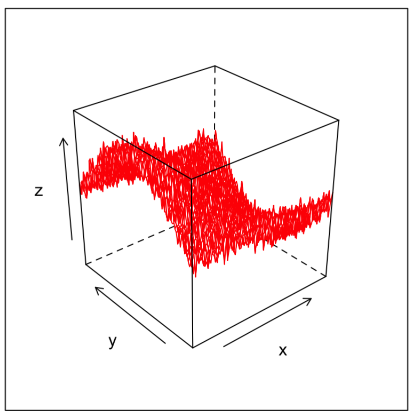

Figure 7: 3D plot

b I have no idea how to solve this problem.

Answer Q.5. ESL 5.4 : Recall that for natural cubic spline, the function is linear beyond the boundary knots. So we first consider the case where $X < \xi_1$, then we have

$$f(X) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$$

We need $\beta_2 = 0$ and $\beta_3 = 0$ to get linearity. Then we consider the case where $X > \xi_K$, then we have

$$
\begin{aligned}
f(X) &= \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \sum_{k=1}^{K} \theta_k (X - \xi_k)^3 \\
&= \beta_0 + \beta_1 X + X^3 \sum_{k=1}^{K} \theta_k - 3X^2 \sum_{k=1}^{K} \theta_k \xi_k + 3X \sum_{k=1}^{K} \theta_k \xi_k^2 - \sum_{k=1}^{k} \theta_k \xi_k^3
\end{aligned}
$$

In order to get linearity, we conclude that $\sum_{k=1}^{K} \theta_k = 0$ and $\sum_{k=1}^{K} \theta_k \xi_k = 0$. To derive a formula like 5.4 and 5.5, suppose that f(X) can written in the form of $\sum_{k=1}^{K} \alpha_k N_k(X)$ and we need to find an expression of α_k. It is obvious that $\alpha_1 = 1$ and $\alpha_2 = X$, then we calculate

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

This is the exact result we want.

Answer Q.6. ESL 5.13 Recall that we have

$$
\hat{f}_\lambda(x_i) = \sum_{j=1}^{n} S_\lambda(i,j) y_j
$$

and we use $\hat{f}_\lambda^{-i}(x_i)$ to denote the case when i-th observation is taken out. Notice that

$$
(1 - S_\lambda(i,i))\hat{f}_\lambda^{-1}(x_i) = \sum_{i \neq j} S_\lambda(i,j) y_j
$$

Combine these two functions together we have

$$\hat{f}_\lambda^{-i}(x_i) = \hat{f}_\lambda^{i}(x_i) + S_\lambda(i,i)\hat{f}_\lambda^{-i}(x_i) - S_\lambda(i,i)y_i$$

Rearrange gives us

$$y_i - \hat{f}_\lambda^{-i}(x_i) = \frac{y_i - \hat{f}_\lambda^{i}(x_i)}{1 - S_\lambda(i,i)}$$

Summing both sides gives us the formula of CV.