Question 1: Linear spline model.

(1):

$$f(x) = \beta_0 + \beta_1 x + \sum_{k}^{K} u_k (x - K_k)_+$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \text{and} \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_K \end{pmatrix}$$ define the coefficient of the polynomial function and truncated function, respectively.

2):

$$X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \qquad Z = \begin{pmatrix} (x_1 - K_1)_+ & \cdots & (x_1 - K_K)_+ \\ \vdots & & \vdots \\ (x_n - K_1)_+ & \cdots & (x_n - K_K)_+ \end{pmatrix}$$

$$\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad f(x) = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}$$

In the matrix form: $y = f(x) + \varepsilon$

$$= X\beta + Zu + \varepsilon$$

Penalized Spline fitting Criterion:

$$\| y - X\beta - Zu \|^2 + \underbrace{\lambda \|u\|^2}$$

we don't want to penalize β

3) In this case, we assume u is Rvec.

OR: u has prior distribution $\left(u \overset{d}{\sim} N(0, G) \right)$,

$$\mathrm{Cov}(u) = \sigma_u^2 I_k = G$$

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n = R$$

Given.

Let $\lambda = \dfrac{\sigma^2}{\sigma_u^2} \Rightarrow \dfrac{\lambda}{\sigma^2} = \boxed{\dfrac{1}{\sigma_u^2}}$

Recall $\min\limits_{\beta, u} \|y - X\beta - Zu\|^2 + \lambda \|u\|^2$

"$\equiv$" $\min\limits_{\beta, u} \dfrac{1}{\sigma^2} \|y - X\beta - Zu\|^2 + \boxed{\dfrac{\lambda}{\sigma^2}} \|u\|^2$

"$\equiv$" $\min\limits_{\beta, u} \underbrace{(y - X\beta - Zu)^T R^{-1} (y - X\beta - Zu) + u^T G^{-1} u}$

Claim:

which correspound to:

$\max\limits_{\beta, u} p(y, u \mid X, Z)$ ← Joint likelihood of y & u.

Proof: $p(y, u \mid X, Z) = p(y \mid u, X, Z)\, p(u)$

$\qquad\qquad\qquad\qquad \overset{d}{\sim} N(X\beta + Zu, R) \qquad \overset{d}{\sim} N(0, G)$

$$p(y, u \mid X, Z) = (2\pi)^{-\frac{n}{2}} \det(R)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y - X\beta - Zu)^T R^{-1} (y - X\beta - Zu)\right\}$$

$$\times (2\pi)^{-\frac{K}{2}} \det(G)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} u^T G^{-1} u\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[(y - X\beta - Zu)^T R^{-1} (y - X\beta - Zu) + u^T G^{-1} u\right]\right\}.$$

$$\Rightarrow -2 \log p(y, u \mid X, Z) \propto (y - X\beta - Zu)^T R^{-1} (y - X\beta - Zu) + u^T G^{-1} u.$$

"drop terms not func of β & u."

$$\max_{\beta, u} \; p(y, u \mid X, Z) \text{ "}\equiv\text{" } \min_{\beta, u} \; (y - X\beta - Zu)^T R^{-1} (y - X\beta - Zu) + u^T G^{-1} u$$

(4): let $C = [\, X \mid Z \,]$

$$D = \left[\begin{array}{cc|c} 0 & & 0 \\ & 0 & \\ \hline & & 1 \quad \quad 0 \\ 0 & & \ddots \\ & & 0 \quad \quad 1 \end{array}\right] \begin{array}{l} \} \; 2 \times 2 \quad \underset{=1}{(p+1)} \\ \\ \} \; K \end{array}$$

$$\theta = \begin{pmatrix} \beta \\ u \end{pmatrix}$$

$$\| y - X\beta - Zu \|^2 + \lambda \| u \|^2$$

$$= \| y - C\theta \|^2 + \lambda \theta^T D \theta = (y - C\theta)^T (y - C\theta) + \lambda \theta^T D \theta$$

$$\frac{\partial \{ \quad \}}{\partial \theta} = 0$$

$$\Rightarrow \hat{\theta} = [C^T C + \lambda D]^{-1} C^T y.$$

5): $\hat{f} = C\hat{\theta} = C [C^T C + \lambda D]^{-1} C^T y.$

6): $y = X\beta + \underbrace{Zu + \varepsilon}_{:= \varepsilon^*} = X\beta + \varepsilon^*$

$G = \sigma_u^2 I_K$

where $u \sim N(0, G)$, $\varepsilon \sim N(0, R)$ $R = \sigma^2 I_n$

idpt

$$\Rightarrow \mathrm{cov}(\varepsilon^*) = Z^T G Z + R := V$$

$$\Rightarrow y \overset{d}{\sim} N(X\beta, V)$$

7): To find β which maximize $P(y|X,Z)$

$$P(y|X,Z) \propto \exp\left\{-\frac{1}{2}(y-X\beta)^T V^{-1}(y-X\beta)\right\}$$

$$\text{"}\equiv\text{"} \quad \min_{\beta} (y-X\beta)^T V^{-1}(y-X\beta)$$

$$\frac{\partial\{(y-X\beta)^T V^{-1}(y-X\beta)\}}{\partial\beta} = -2X^T V^{-1}y + 2X^T V^{-1}X\beta = 0$$

$$\Rightarrow \quad X^T V^{-1} y = X^T V^{-1} X\beta$$

$$\Rightarrow \quad \hat{\beta} = [X^T V^{-1} X]^{-1} X^T V^{-1} y$$

$$\text{cov}(\hat{\beta}) = [X^T V^{-1} X]^{-1} X^T V^{-1} \underbrace{V}_{\text{cov}(y)} V^{-1} X [X^T V^{-1} X]^{-1}$$

$$= [X^T V^{-1} X]^{-1}$$

#