0. a. $\hat{\beta}_\lambda = (X^T X + \lambda I)^{-1} X^T y$

Show $\|\beta_\lambda\|\big|_{\lambda>0} \le \|\beta_\lambda\|\big|_{\lambda=0}$

sol: $X = UDV^T = \sum_{j=1}^{p} d_j u_j v_j^T$

$$(X^T X + \lambda I)^{-1} X^T y = (V D^T \underbrace{U^T U}_{I} D V^T + \lambda V V^T)^{-1} V D^T U^T y$$

$$= (V D^T D V^T + \lambda V V^T)^{-1} V D^T U^T y$$

$$= (D^T D + \lambda I)^{-1} V D^T U^T y$$

$$= \sum_{j=1}^{p} \frac{v_j d_j u_j^T}{d_j^2 + \lambda} y$$

$(\hat{\beta}_\lambda|_{\lambda=0} - \hat{\beta}_\lambda|_{\lambda>0})$

$$\sum_{j=1}^{p} \frac{v_j d_j u_j^T}{d_j^2} - \sum_{j=1}^{p} \frac{v_j d_j u_j^T}{d_j^2+\lambda} y = \sum_{j=1}^{p} \frac{d_j^2 + \lambda - d_j^2}{d_j(d_j^2+\lambda)} v_j u_j^T y$$

$$= \sum_{j=1}^{p} \frac{\lambda}{d_j(d_j^2+\lambda)} \boxed{v_j u_j^T y} > 0$$

(we need $\|\beta_\lambda\|\big|_{\lambda>0}$; $\|\beta_\lambda\|\big|_{\lambda=0}$)

$\Rightarrow \|\beta_\lambda\|\big|_{\lambda>0} \le \|\beta_\lambda\|\big|_{\lambda=0}$.

b. ① $\overline{err} = \frac{1}{n} \sum (y_i - \hat{y}_i)^2$

$$= \frac{1}{n} (y - X\hat{\beta})^T (y - X\hat{\beta})$$

$$= \frac{1}{n} (y - X(X^T X + \lambda I)^{-1} X^T y)^T (y - X(X^T X + \lambda I)^{-1} X^T y)$$

$$= \frac{1}{n} y (I - X(X^T X + \lambda I)^{-1} X^T)^T (I - X(X^T X + \lambda I)^{-1} X^T) y$$

$$= \frac{1}{n} y (I - X(X^T X + \lambda I)^{-1} X^T)^2 y$$

Scanned by CamScanner

0.b. ③

$$\frac{1}{n} y(I - X(X^TX+\lambda I)^{-1}X^T)^2 y$$

$$= \frac{1}{n} y(I - UDV^T(VD^TU^TUDV^T+\lambda I)^{-1} VD^TU^T)^2 y$$

$$= \frac{1}{n} y(I - UDV^T(D^TD+\lambda I)^{-1} VD^TU^T)^2 y$$

$$= \frac{1}{n} y(I - DD^T(D^TD+\lambda I)^{-1})^2 y$$

$$= \frac{1}{n} y \sum_{j=1}^{p} \left(1 - \frac{d_j^2}{d_j^2+\lambda}\right)^2 y$$

$$\frac{\partial}{\partial\lambda} \frac{1}{n} y \sum_{j=1}^{p} \left(1 - \frac{d_j^2}{d_j^2+\lambda}\right)^2 y$$

$$= \frac{1}{n} y \sum_{j=1}^{p} 2\left(1 - \frac{d_j^2}{d_j^2+\lambda}\right) y \cdot \frac{d^2}{(d_j^2+\lambda)^2}$$

$$= \frac{1}{n} y \sum_{j=1}^{p} 2\, \frac{d_j^2(d_j^2+\lambda) - d^4}{(d_j^2+\lambda)^3}\, y$$

$$= \frac{1}{n} y \sum_{j=1}^{p} 2\, \frac{d_j^2\lambda}{(d_j^2+\lambda)^3}\, y > 0$$

$\therefore$ $\overline{err}$ is increasing function of λ.

Scanned by CamScanner

1.

For $X < \xi_1$: $f(X) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$

$\because$ linear boundary conditions $(f(X) = AX + B)$

$\therefore \beta_2 = \beta_3 = 0$ ※

For $X > \xi_K$: $f(X) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$

$$+ \sum_{k=1}^{K} \beta_{k+3} \left(X^3 - 3X^2 \xi_k + 3X \xi_k^2 - \xi_k^3 \right)$$

$\because$ linear boundary conditions $(f(X) = AX + B)$

$$\therefore \begin{cases} \beta_2 + \sum_{k=1}^{K} \beta_{k+3} (-3 \cdot \xi_k) = 0 \\ \beta_3 + \sum_{k=1}^{K} \beta_{k+3} = 0 \end{cases} \Rightarrow \begin{cases} \sum_{k=1}^{K} \beta_{k+3} \xi_k = 0 \\ \sum_{k=1}^{K} \beta_{k+3} = 0 \end{cases}$$ ※

2. (a)

$(-2,-1)$
$(0,0)$
$(1,2)$

$$\Rightarrow \begin{cases} \bar{x} = -\frac{1}{3} \\ \bar{y} = \frac{1}{3} \end{cases}$$

$$X_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \left(-\frac{1}{3}\right) = \frac{1}{3} \begin{bmatrix} -5 \\ 1 \\ 4 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} - \frac{1}{3} = \frac{1}{3} \begin{bmatrix} -4 \\ -1 \\ 5 \end{bmatrix}$$

$$\underset{\sim}{X} = [X_1 \; X_2] = \frac{1}{3} \begin{bmatrix} -5 & -4 \\ 1 & -1 \\ 4 & 5 \end{bmatrix}$$

$$\phi_1 = \underset{\|\phi_1'\|^2=1}{\arg\max} \operatorname{Var} (\phi_{11}' X_1 + \phi_{21}' X_2)$$

= eigenvector corresponding to the max eigenvalue of $\underset{\sim}{X}^T \underset{\sim}{X}$

$$\tilde{X} = (\tilde{X}_1, \tilde{X}_2)^T = X - \phi_1 (\phi_1^T X)$$

$$\phi_2 = \underset{\|\phi_2'\|^2=1}{\arg\max} \operatorname{Var} (\tilde{\phi}_{21}' \tilde{X}_1 + \tilde{\phi}_{22}' \tilde{X}_2)$$

= eigenvector corresponding to the 2^{nd} largest eigenvalue of $\underset{\sim}{X}^T \underset{\sim}{X}$

$$\underset{\sim}{X}^T \underset{\sim}{X} = \frac{1}{9} \begin{bmatrix} 42 & 39 \\ 39 & 42 \end{bmatrix} :$$

eigenvalue	eigenvector
$\frac{1}{3}$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$
9	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\begin{cases} Z_1 = \frac{1}{\sqrt{2}} x + \frac{1}{\sqrt{2}} y \\ Z_2 = -\frac{1}{\sqrt{2}} x + \frac{1}{\sqrt{2}} y \end{cases}$$ ※

Scanned by CamScanner
Scanned by CamScanner

2.(b) first principal component:

$$z_i = \frac{1}{\sqrt{2}} x_i + \frac{1}{\sqrt{2}} y_i$$

$$\begin{cases} (-2,-1): & z_1 = -\frac{3}{\sqrt{2}} \\ (0,0): & z_2 = 0 \\ (1,2): & z_3 = \frac{3}{\sqrt{2}} \end{cases}$$

$$\bar{z} = 0, \quad \sigma^2 = \frac{\sum_{i=1}^{3} (z_i - \bar{z})^2}{3} = 3 \#$$

second principal component:

$$z_i = -\frac{1}{\sqrt{2}} x_i + \frac{1}{\sqrt{2}} y_i$$

$$\begin{cases} (-2,-1): & z_1 = \frac{1}{\sqrt{2}} \\ (0,0): & z_2 = 0 \\ (1,2): & z_3 = \frac{1}{\sqrt{2}} \end{cases}$$

$$\bar{z} = \frac{\sqrt{2}}{3}, \quad \sigma^2 = 0.1111$$

#

Scanned by CamScanner

3. initialize : $\hat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} y_i$, $\hat{f}_j \equiv 0$, $\forall i, j$

iterate 1 :

$$\hat{f}_1 \leftarrow \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{\substack{j=1 \\ k\neq j}}^{2} \hat{f}_k (x_{ik})\right)^2 + \lambda_1 \|f_1\|_2^2 + \lambda_2 \|f_2\|_2^2$$

$$= \sum_{i=1}^{n} (y_i - \beta_0)^2$$

$$\hat{f}_1 \leftarrow \hat{f}_1 - \frac{1}{n}\sum_{i=1}^{n} \hat{f}_1 (x_{i1})$$

$$\hat{f}_2 \leftarrow \sum_{i=1}^{n} \left(y_i - \beta_0 - \hat{f}_1 (x_{i1})\right)^2 + \lambda_1 \|f_1\|_2^2$$

$$\hat{f}_2 \leftarrow \hat{f}_2 - \frac{1}{n}\sum_{i=1}^{n} \hat{f}_2 (x_{i2})$$

#

Scanned by CamScanner
Scanned by CamScanner

4. It was mention in the chapter that a cubic regression spline with on knot ξ can be obtained using a basis of the form $x, x^2, x^3, (x-\xi)_+^3$, where $(x-\xi)_+^3 = (x-\xi)^3$ if $x > \xi$ and equals 0 otherwise.

We will now show that a fcn. of the form

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x-\xi)_+^3$$

is indeed a cubic regression spline, regardless of the values of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

(a) Find a cubic polynomial

$$f_1(x) = a_1 + b_1 x + c_1 x^2 + d_1 x^3$$

such that $f(x) = f_1(x)$ for all $x \le \xi$. Express a_1, b_1, c_1, d_1 in terms of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

For $x \le \xi$

$$\Rightarrow \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 = a_1 + b_1 x + c_1 x^2 + d_1 x^3$$

$$\Rightarrow a_1 = \beta_0,\ b_1 = \beta_1,\ c_1 = \beta_2,\ d_1 = \beta_3 \text{ \#}$$

(b) Find a cubic polynomial

$$f_2(x) = a_2 + b_2 x + c_2 x^2 + d_2 x^3$$

Such that $f(x) = f_2(x)$ for all $x > \xi$. Express a_2, b_2, c_2, d_2 in terms of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$. We have now established that $f(x)$ is a piecewise polynomial.

For $x > \xi$,

$$\Rightarrow \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x-\xi)^3$$

$$= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x^3 - 3\xi x^2 + 3\xi^2 x - \xi^3)$$

$$= (\beta_0 - \beta_4 \xi^3) + (\beta_1 + 3\beta_4 \xi^2) x + (\beta_2 - 3\beta_4 \xi) x^2 + (\beta_3 + \beta_4) x^3$$

$$= \quad a_2 \quad + \quad b_2 x \quad + \quad c_2 x^2 \quad + \quad d_2 x^3$$

$$\Rightarrow a_2 = \beta_0 - \beta_4 \xi^3,\ b_2 = \beta_1 + 3\beta_4 \xi^2,\ c_2 = \beta_2 - 3\beta_4 \xi,\ d_2 = \beta_3 + \beta_4 \text{ \#}$$

(c) Show that $f_1(\xi) = f_2(\xi)$. That is, $f(x)$ is continuous at ξ.

$$f_1(\xi) = \beta_0 + \beta_1 \xi + \beta_2 \xi^2 + \beta_3 \xi^3$$

$$f_2(\xi) = (\beta_0 - \beta_4 \xi^3) + (\beta_1 + 3\beta_4 \xi^2)\xi + (\beta_2 - 3\beta_4 \xi)\xi^2 + (\beta_3 + \beta_4)\xi^3$$

$$= \beta_0 - \cancel{\beta_4 \xi^3} + \beta_1 \xi + \cancel{3\beta_4 \xi^3} + \beta_2 \xi^2 - \cancel{3\beta_4 \xi^3} + \beta_3 \xi^3 + \cancel{\beta_4 \xi^3}$$

$$= \beta_0 + \beta_1 \xi + \beta_2 \xi^2 + \beta_3 \xi^3$$

$$= f_1(\xi) \text{ \#}$$

Scanned by CamScanner
Scanned by CamScanner

4. (d) Show that $f_1'(\xi) = f_2'(\xi)$. That is, $f'(x)$ is continuous at ξ.

$$f_1'(x) = b_1 + 2c_1x + 3d_1x^2$$

$$f_1'(\xi) = \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2$$

$$\begin{aligned} f_2'(\xi) &= (\beta_1 + 3\beta_4\xi^2) + 2(\beta_2 - 3\beta_4\xi)\xi + 3(\beta_3 + \beta_4)\xi^2 \\ &= \beta_1 + 3\beta_4\xi^2 + 2\beta_2\xi - 6\beta_4\xi^2 + 3\beta_3\xi^2 + 3\beta_4\xi^2 \\ &= \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2 \\ &= f_1'(\xi) \end{aligned}$$

(e) Show that $f_1''(x) = f_2''(x)$. That is, $f''(x)$ is continuous at ξ.

$$f''(x) = 2c_1 + 6d_1x$$

$$f_1''(\xi) = 2\beta_2 + 6\beta_3x$$

$$\begin{aligned} f_2''(\xi) &= 2(\beta_2 - 3\beta_4\xi) + 6(\beta_3 + \beta_4)\xi \\ &= 2\beta_2 - \cancel{6\beta_4\xi} + 6\beta_3\xi + \cancel{6\beta_4\xi} \\ &= 2\beta_2 + 6\beta_3\xi \\ &= f_1''(\xi) \end{aligned}$$

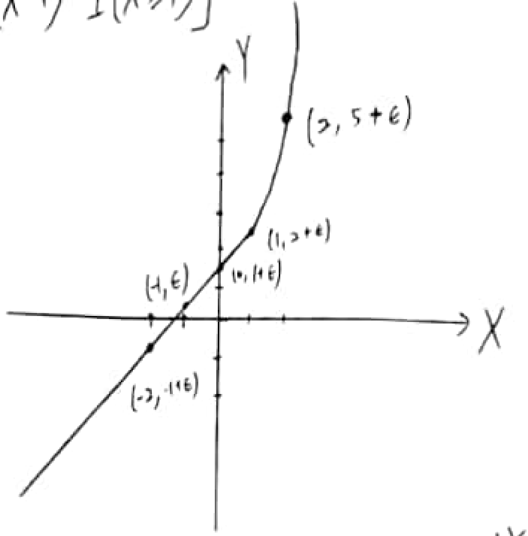
Therefore, $f(x)$ is indeed a cubic spline.

Scanned by CamScanner
Scanned by CamScanner

5. $Y = \beta_0 + \beta_1 X + \beta_2 \left[(X-1)^2 I(X \geq 1) \right] + \epsilon$

(a) $Y = 1 + X + 2\left[(X-1)^2 I(X \geq 1) \right] + \epsilon$

X	Y
-2	$-1+\epsilon$
-1	ϵ
0	$1+\epsilon$
1	$2+\epsilon$
2	$5+\epsilon$



$\#$

(b)

X	Y
-1	-1
1	2
2	8

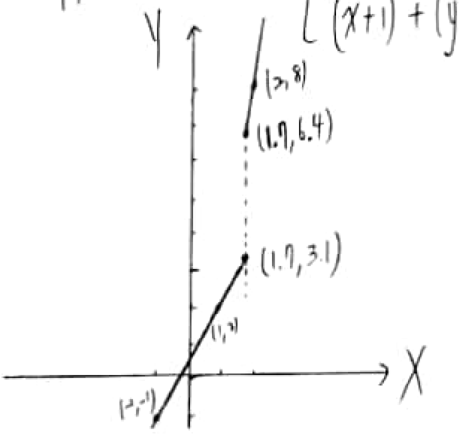
$$\begin{cases} -1 = \beta_0 - \beta_1 \\ 2 = \beta_0 + \beta_1 \\ 8 = \beta_0 + 2\beta_1 + \beta_2 \end{cases} \Rightarrow \begin{cases} \beta_0 = \frac{1}{2} \\ \beta_1 = \frac{3}{2} \\ \beta_2 = \frac{9}{2} \end{cases}$$

$\#$

(c) minimize $K(y_1 - \beta_0 - \beta_1 x_1)^2 + K(y_2 - \beta_0 - \beta_1 x_2)^2$

$$\begin{cases} x_1 = -1, \ y_1 = -1 \\ x_2 = 1, \ y_2 = 2 \end{cases} \Rightarrow y = \frac{1}{2} + \frac{3}{2}x \qquad \begin{cases} x_1 = 1, y_1 = 2 \\ x_2 = 2, y_2 = 8 \end{cases} \Rightarrow y = -4 + 6x$$

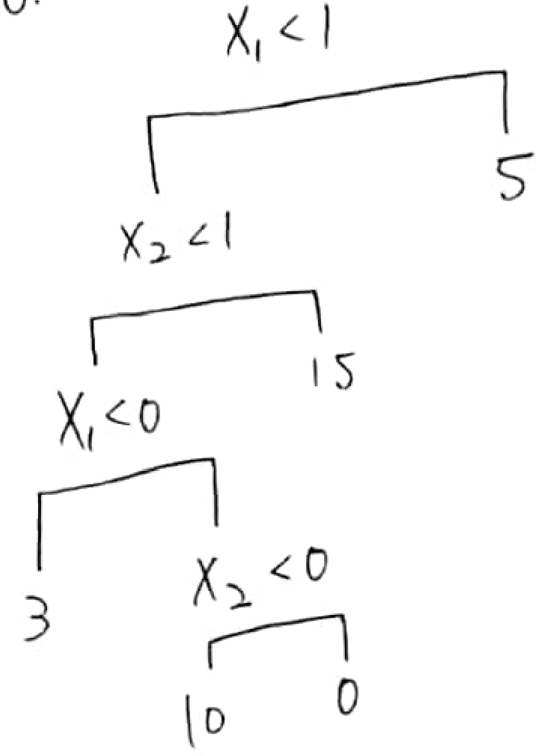
jump point happens when $\begin{cases} y = \frac{1}{2} + \frac{3}{2}x \\ (x+1)^2 + (y+1)^2 = (x-2)^2 + (y-8)^2 \end{cases} \Rightarrow \begin{cases} x = 1.7 \\ y \doteq 3.1 \end{cases}$



$\#$

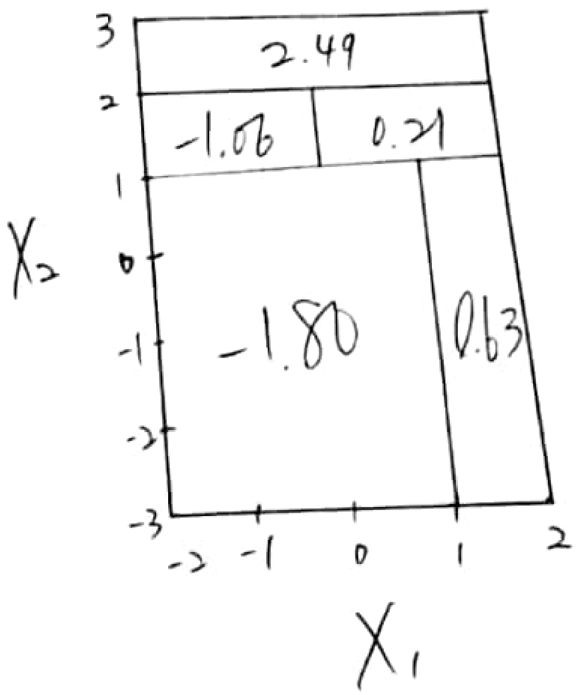
Scanned by CamScanner
Scanned by CamScanner

6. (a)



#

(b)



#

Scanned by CamScanner