1.

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

For $j = 1$:

$$y_i = \hat{\beta}_0 + \beta_{01} + \beta_{11} x_{i1} + \beta_{21} x_{i1}^2 + \beta_{31} x_{i1}^3 + \hat{f}_2(x_{i2}) + \varepsilon_i$$

Let $\underset{\sim}{X}_1 = \begin{bmatrix} 1 & x_{11} & x_{11}^2 & x_{11}^3 \\ \vdots & & \ddots & \vdots \\ 1 & x_{n1} & x_{n1}^2 & x_{n1}^3 \end{bmatrix}$, $\underset{\sim}{y}_1 = \begin{bmatrix} y_1 - \hat{\beta}_0 - \hat{f}_2(x_{12}) \\ y_2 - \hat{\beta}_0 - \hat{f}_2(x_{22}) \\ \vdots \\ y_n - \hat{\beta}_0 - \hat{f}_2(x_{n2}) \end{bmatrix}$

, then the solution for cubic polynomial will be

$\hat{\underline{\beta}}_1 = (\underset{\sim}{X}_1^T \underset{\sim}{X}_1)^{-1} \underset{\sim}{X}_1^T \underset{\sim}{y}_1$

$\Rightarrow \hat{f}_1 \leftarrow \underset{\sim}{X}_1 \beta$

$\hat{f}_1 \leftarrow f_1 - \frac{1}{n}\sum_{i=1}^{n} \hat{f}_1(x_{i1})$

For $j = 2$:

Let $\underset{\sim}{X}_2 = \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ & \vdots \\ 1 & x_{n2} \end{bmatrix}$, $\underset{\sim}{y}_2 = \begin{bmatrix} y_1 - \hat{\beta}_0 - \hat{f}_1(x_{11}) \\ y_2 - \hat{\beta}_0 - \hat{f}_1(x_{21}) \\ \vdots \\ y_n - \hat{\beta}_0 - \hat{f}_1(x_{n1}) \end{bmatrix}$, the solution for ridge regression is

$\hat{\underline{\beta}}_2 = (\underset{\sim}{X}_2^T \underset{\sim}{X}_2 + \lambda \mathbb{I})^{-1} \underset{\sim}{X}_2^T \underset{\sim}{y}_2$

$\Rightarrow \hat{f}_2 \leftarrow \underset{\sim}{X}_2 \hat{\underline{\beta}}_2$

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

For $\hat{\beta}_0$:

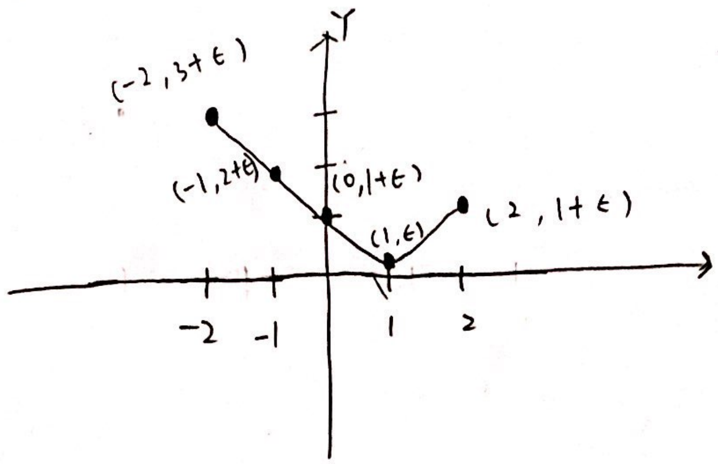
$\hat{\beta}_0 \leftarrow \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}_1(x_{i1}) - \hat{f}_2(x_{i2})\right)$

2.

(a)

$$Y = 1 - b_1(X) + 2\, b_2(X) + \epsilon$$
$$= 1 - X + 2(X-1)^2 I(X \geq 1) + \epsilon$$

X	-2	-1	0	1	2
Y	$3+\epsilon$	$2+\epsilon$	$1+\epsilon$	ϵ	$1+\epsilon$



(b) $\{(-2,-1),\ (1,2),\ (2,7)\}$

plug into $\quad Y = \beta_0 + \beta_1 X + \beta_2 (X-1)^2 I(X \geq 1)$

$$\Rightarrow \begin{cases} -1 = \beta_0 - 2\beta_1 & \text{— } ① \\ 2 = \beta_0 + \beta_1 & \text{— } ② \\ 7 = \beta_0 + 2\beta_1 + \beta_2 & \text{— } ③ \end{cases}$$

② − ① ; ③ − ①

$\Rightarrow \hat{\beta}_1 = 1$ $\qquad\qquad 4\beta_1 + \beta_2 = 8$

$\qquad\qquad\qquad \Rightarrow \hat{\beta}_2 = 4$

$$\therefore \begin{cases} \hat{\beta}_1 = 1 \\ \hat{\beta}_2 = 4 \end{cases}$$

3.

(a)

$$K(x,x') = \langle (x, x^2), (x', x'^2) \rangle$$
$$= xx' + x^2 x'^2 = xx'(1 + xx')$$

(b)

$\because$ Support vectors $(-2,4), (-1,1)$

$w_1 Y_1 + w_2 Y_2 + c = 0$

$\Rightarrow (-2-(-1)) Y_1 + (4-1) Y_2 + c = 0$

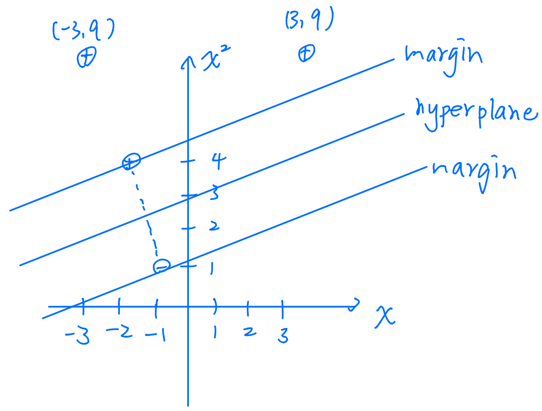
$\Rightarrow -Y_1 + 3 Y_2 + c = 0$

Substitude $\frac{(-2,4)+(-1,1)}{2}$ into it $\Rightarrow c = -9$

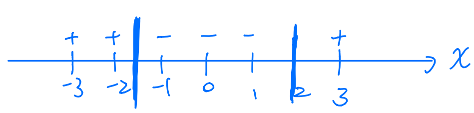
Margin $= \sqrt{(-2-(-1))^2 + (4-1)^2} / 2 = \sqrt{10}/2$

$w_1 = -1,\ w_2 = 3,\ c = -9$

(c)



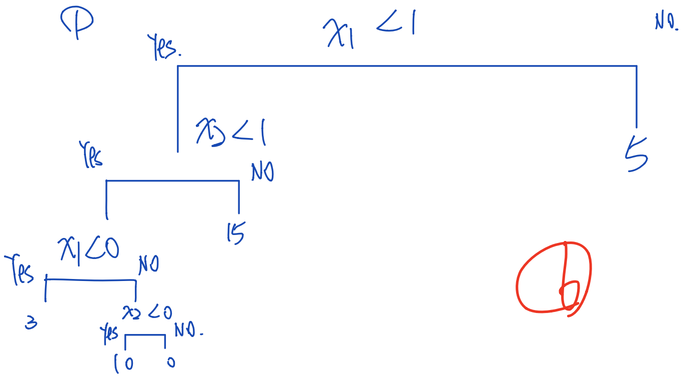
(d)



4.

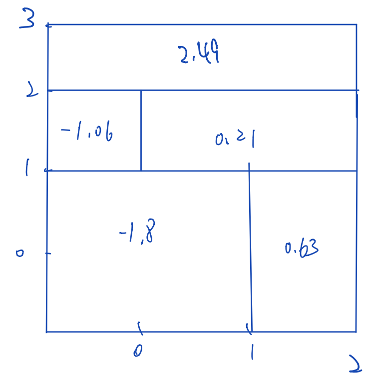
(a)

①


$x_1 < 1$
Yes.
NO.
$x_2 < 1$
Yes
NO
5
$x_1 < 0$
Yes
NO
15
3
$x_2 < 0$
Yes
NO.
10
0

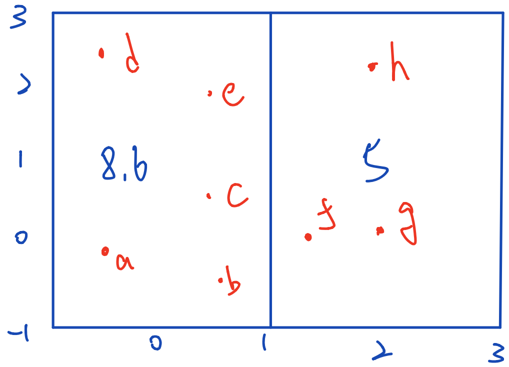
6

②


3
2
1
0
2.49
-1.06
0.21
-1.8
0.63
0
1
2

6

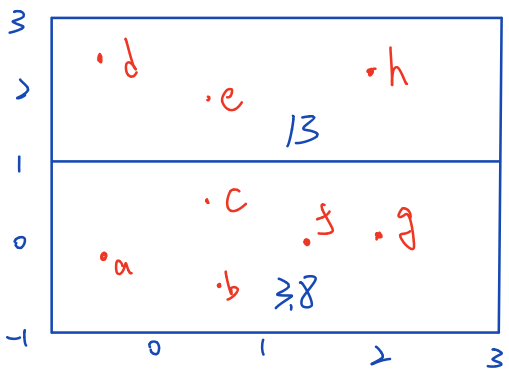
4. (b)



RSS = 233.2

Choose min-RSS and update residual.

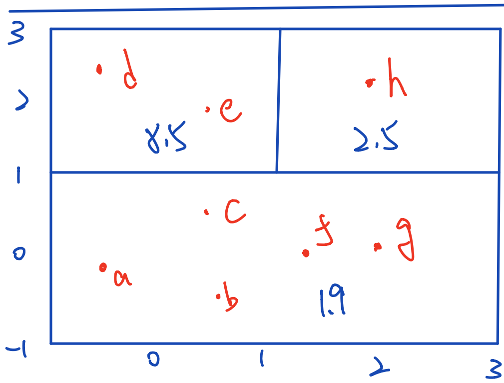
$r = y - \lambda \hat{y}$ —— 4



RSS = 98.8. ✓

⇒ a	1.1
b	8.1
c	-1.9
d	5.5
e	11.5
f	2.1
g	0.1
h	2.5

1

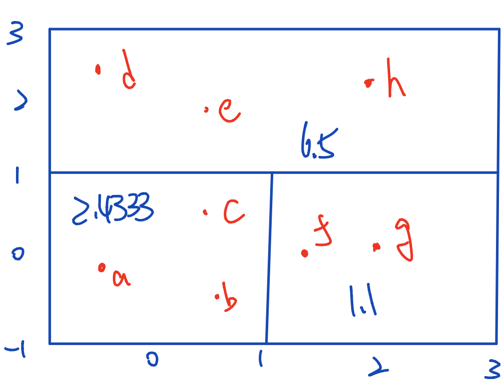


RSS = 74.8 ✓

⇒ Choose min-RSS and update residual.

$r = y - \lambda \hat{y}$

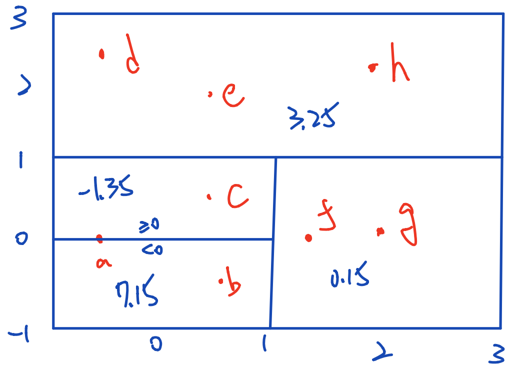
⇒ a	0.15
b	7.15
c	-2.85
d	1.25
e	7.25
f	1.15
g	-0.85
h	1.25



RSS = 96.67.

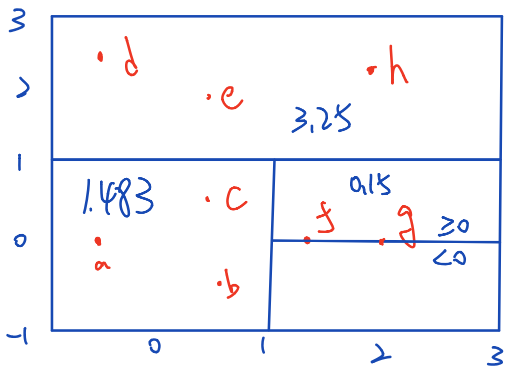
1

$f^{(3)}(x)$



✓

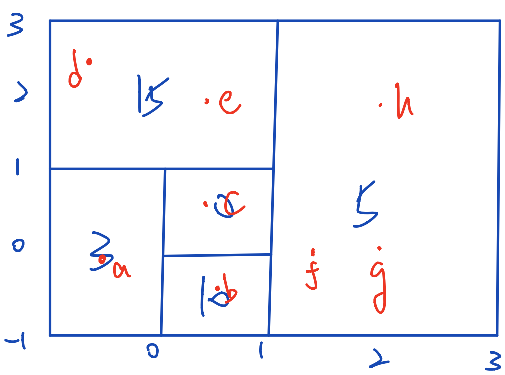
RSS = 20.5 #



RSS = 78.67

①

Recall: $f^{(1)}(x)$



the final tree: $\hat{f}(x) = 0.5 \cdot f^{(1)}(x) + 0.5 f^{(2)}(x)$ #

③