

$$1. a) y = \theta_0 + \theta_1 x.$$

$$X = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} \quad Y = \begin{bmatrix} 6 \\ 10 \\ 16 \end{bmatrix}$$

co-ordinate Descent.

$$(\theta_0, \theta_1) = (0, 0).$$

Fixed $\theta_0 = 0$.

$$RSS = \sum (y_i - \hat{y}_i)^2$$

$$RSS = (6 - \theta_1(1))^2 + (10 - \theta_1(3))^2 + (16 - \theta_1(6))^2$$

$$\frac{\partial (RSS)}{\partial \theta_1} = 0 \Rightarrow 2(6 - \theta_1)(-1) + (2)(10 - \theta_1)(-3) + (2)(16 - 6\theta_1)(-6) = 0$$

$$2\theta_1 - 12 + 18\theta_1 - 60 + 72\theta_1 - 192 = 0$$

$$92\theta_1 = 264$$

$$\theta_1 = \frac{264}{92} = \underline{\underline{2.8695}} \approx 2.87.$$

$$(\theta_0, \theta_1) = (0, 2.87).$$

Fix θ_1 ,

$$\Rightarrow \text{RSS} = (6 - (\theta_0 + (2.87)(1)))^2 \\ + (10 - (\theta_0 + (2.87)(3)))^2 \\ + (16 - (\theta_0 + (2.87)(6)))^2$$

$$\frac{\partial (\text{RSS})}{\partial \theta_0} = 2(6 - 2.87 - \theta_0)(-1) \\ + 2(10 - 2.87 \times 3 - \theta_0)(-1) \\ + 2(16 - 2.87 \times 6 - \theta_0)(-1).$$

$$= 2(3.13 - \theta_0)(-1) + 2(1.39 - \theta_0)(-1) \\ + 2(-1.22 - \theta_0)(-1)$$

$$= 2\theta_0 - 6.26 + 2\theta_0 - 2.78 \\ + 2\theta_0 + 2.44 = 0$$

$$\Rightarrow 6\theta_0 - 6.6 = 0$$

$$\underline{\underline{\theta_0 = 1.1}}$$

$$(\theta_0, \theta_1) = (1.1, 2.87). \quad \text{Iteration ①}$$

$$\text{put } \theta_0 = 1.1 \text{ (constant)}$$

$$\text{RSS} = (6 - (1.1 + \theta_1(1)))^2 + (10 - (1.1 + \theta_1(3)))^2 + (16 - (1.1 + \theta_1(6)))^2$$

$$\frac{\partial(\text{RSS})}{\partial \theta_1} = 2(6 - 1.1 - \theta_1)(-1) + 2(10 - 1.1 - 3\theta_1)(-3) + 2(16 - 1.1 - 6\theta_1)(-6).$$

$$= 2(\theta_1 - 4.9) + 6(3\theta_1 - 8.9) + 12(6\theta_1 - 14.9)$$

$$= 2\theta_1 + 18\theta_1 + 72\theta_1 - 9.8 - 53.4 - 178.8.$$

$$\Rightarrow 92\theta_1 = 242$$

$$\theta_1 = 2.63.$$

$$(\theta_0, \theta_1) = (1.1, 2.63)$$

$$\text{keep } \theta_1 = 2.63 \text{ constant.}$$

$$RSS = (6 - (\theta_0 + (2.63)(1)))^2 + (10 - (\theta_0 + (2.63)(3)))^2 + (16 - (\theta_0 + (2.63)(6)))^2$$

$$\frac{\partial (RSS)}{\partial \theta_0} = 2(6 - 2.63 - \theta_0)(-1) + 2(10 - 2.63 \times 3 - \theta_0)(-1)$$

$$+ 2(16 - 2.63 \times 6 - \theta_0)(-1)$$

$$= 2(\theta_0 - 3.37) + 2(\theta_0 - 2.11) + 2(\theta_0 - 0.22) = 0$$

$$\Rightarrow 6\theta_0 = 11.04$$

$$\theta_0 = 1.84$$

$$(\theta_0, \theta_1) = (1.84, 2.63) \quad \text{Iteration (2)}$$

keep θ_0 constant

$$\frac{\partial (RSS)}{\partial \theta_1} = 0 \Rightarrow 92\theta_1 = 2(6 - 1.84) + 6(10 - 1.84) + 12(16 - 1.84)$$

$$\theta_1 = 2.45$$

$$(\theta_0, \theta_1) = (1.9, 2.45).$$

keep θ_1 constant.

$$\frac{\partial (RSS)}{\partial \theta_0} = 0 \Rightarrow 6\theta_0 = 2(6 - 2.45) + 2(10 - 2.45 \times 3) + 2(16 - 2.45 \times 6)$$

$$6\theta_0 = 15$$

$$\theta_0 = 2.5$$

$$(\theta_0, \theta_1) = (2.5, 2.45). \text{ Iteration (3)}$$

$$b) (\theta_0, \theta_1) = (0, 0) \quad \alpha = 0.01$$

$$X = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} \quad Y = \begin{bmatrix} 6 \\ 10 \\ 16 \end{bmatrix} \begin{matrix} \xrightarrow{\textcircled{2}} \\ \xrightarrow{\textcircled{1}} \\ \xrightarrow{\textcircled{3}} \end{matrix} \text{ sample order.}$$

random samples

$$\text{sample } \textcircled{1} \rightarrow (3, 10)$$

$$RSS = (10 - (\theta_0 + \theta_1(3)))^2$$

$$\frac{\partial RSS}{\partial \theta_0} = 2(10 - \theta_0 - \theta_1(3))(-1) = -20$$

$$\frac{\partial \text{RSS}}{\partial \theta_1} = 2(10 - \theta_0 - 3\theta_1)(-3)$$

$$= -60$$

$$\theta_0 = \theta_0 - \alpha \frac{\partial \text{RSS}}{\partial \theta_0}$$

$$= 0 + (0.01)(20)$$

$$= 0.2$$

$$\theta_1 = 0 - (0.01)(-60) = 0.6$$

$$(\theta_0, \theta_1) = (0.2, 0.6).$$

sample ② (1, 6)

$$\text{RSS} = (6 - (\theta_0 + \theta_1))^2$$

$$\frac{\partial \text{RSS}}{\partial \theta_0} = 2(6 - \theta_0 - \theta_1)(-1) = 2(6 - 0.8)(-1)$$

$$= -10.4$$

$$\frac{\partial \text{RSS}}{\partial \theta_1} = 2(6 - \theta_0 - \theta_1)(-1) = 2(6 - 0.8)(-1)$$

$$= -10.4$$

$$\theta_0 = \theta_0 - \alpha \frac{\partial \text{RSS}}{\partial \theta_0}$$

$$= 0.2 + (0.01)(10.4)$$

$$= 0.304$$

$$\theta_1 = 0.704$$

Sample ③ (6, 16)

$$(\theta_0, \theta_1) = (0.304, 0.704)$$

$$RSS = (16 - (\theta_0 + \theta_1(6)))^2$$

$$\frac{\partial RSS}{\partial \theta_0} = 2(16 - 6(0.704) - (0.304))(-1)$$

$$= -22.9$$

$$\frac{\partial RSS}{\partial \theta_1} = 2(16 - 6(0.704) - (0.304))(-6)$$

$$= -137.6$$

$$\theta_0 = 0.304 - (0.01)(-22.9)$$

$$= 0.304 + 0.229 = 0.533 //$$

$$\theta_1 = 0.704 + 1.37 =$$

$$(\theta_0, \theta_1) = (0.533, \underline{2.074}) = \underline{(2.074)}$$

$$(c) (X^T X - \lambda I)^{-1} X^T Y$$

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 10 \\ 10 & 46 \end{bmatrix}$$

$$\begin{aligned}
 X^T X - I &= \begin{bmatrix} 3 & 10 \\ 10 & 46 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 3 & 10 \\ 10 & 45 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
 (X^T X - I)^{-1} &= \frac{1}{(135 - 100)} \begin{bmatrix} 45 & -10 \\ -10 & 3 \end{bmatrix} \\
 &= \begin{bmatrix} 1.028 & -0.285 \\ -0.285 & 0.086 \end{bmatrix}
 \end{aligned}$$

$$\begin{bmatrix} 1 & 0.428 & -0.428 \\ -0.2 & -0.028 & 0.2286 \end{bmatrix} \Leftarrow (X^T X - I)^{-1} X^T$$

$$\Theta = \begin{bmatrix} 3.428 \\ 2.1714 \end{bmatrix}$$