Chennuri Prateek
16110042.

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

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

using Normal Eq. Method.

$$\hat{y}_i = \theta_0 + \theta_1 x_i$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$

$$\hat{Y} = X\theta$$

error $(\epsilon_i) = y_i - \hat{y}_i$

let $\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}$

$$Y = X\theta + \epsilon \quad \text{— (i)}$$

we need to minimize $\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2$

$\Rightarrow$ we need to minimize $\epsilon^T \epsilon$.

$\epsilon = Y - X\theta$ (from (i)).

$$\varepsilon_1^T\varepsilon_1 = (Y - X\theta)^T (Y - X\theta).$$

$$\varepsilon_1^T\varepsilon_1 = (Y^T - \theta^T X^T)(Y - X\theta)$$

$$\varepsilon_1^T\varepsilon_1 = Y^TY - Y^TX\theta - \theta^TX^TY + \theta^TX^TX\theta.$$

$$\varepsilon_1^T\varepsilon_1 = Y^TY - 2Y^TX\theta + \theta^TX^TX\theta$$

$$\frac{\partial}{\partial\theta}(\varepsilon_1^T\varepsilon_1) = 0 - 2X^TY + 2X^TX\theta = 0$$

$$2X^TY = 2X^TX\theta$$

$$X^TY = X^TX\theta$$

$$\theta = (X^TX)^{-1}X^TY.$$

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

$$X^TX = \begin{bmatrix} 3 & 10 \\ 10 & 46 \end{bmatrix}$$

$$|X^TX| = 3 \times 46 - 100 = 138 - 100$$

$$= 38$$

$$\operatorname{adj}(X^TX) = \begin{bmatrix} 46 & -10 \\ -10 & 3 \end{bmatrix}$$

$$(X^TX)^{-1} = \frac{\operatorname{adj}(X^TX)}{|X^TX|} = \frac{1}{38}\begin{bmatrix} 46 & -10 \\ -10 & 3 \end{bmatrix}.$$

$$\theta = (X^TX)^{-1} X^T y.$$

$$\theta = \frac{1}{38}\underbrace{\begin{bmatrix} 46 & -10 \\ -10 & 3 \end{bmatrix}}_{2\times 2}\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 6 \end{bmatrix}}_{2\times 3}\underbrace{\begin{bmatrix} 6 \\ 10 \\ 16 \end{bmatrix}}_{3\times 1}$$

$$\theta = \frac{1}{38}\begin{bmatrix} 46 & -10 \\ -10 & 3 \end{bmatrix}\begin{bmatrix} 32 \\ 132 \end{bmatrix}$$

$$= \frac{1}{38}\begin{bmatrix} 46\times 32 - 1320 \\ -320 + 132\times 3 \end{bmatrix}$$

$$= \frac{1}{38}\begin{bmatrix} 152 \\ 76 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}.$$

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

$$y_i = \theta_0 + \theta_1 x_i.$$

initial Guess is $(\theta_0, \theta_1) = (0, 0)$.

$$\hat{Y} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\varepsilon_i = y_i - \hat{y}_i$$

$$MSE = \frac{\sum_{i=1}^{N} |\varepsilon_i|^2}{N} \Rightarrow MSE = \frac{|\varepsilon_1|^2 + |\varepsilon_2|^2 + |\varepsilon_3|^2}{3}$$

$$= \frac{(y_1 - (\theta_0 + \theta_1 x_1))^2 + (y_2 - (\theta_0 + \theta_1 x_2))^2 + (y_3 - (\theta_0 + \theta_1 x_3))^2}{3.}$$

$$= \frac{1}{3}\left[\sum_{i=1}^{3} \left(y_i^2 + (\theta_0 + \theta_1 x_i)^2 - 2y_i(\theta_0 + \theta_1 x_i)\right)\right]$$

$$\frac{\partial MSE}{\partial \theta_0} = \frac{1}{3}\left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i) - 2y_i\right)\right]$$

$$\frac{\partial MSE}{\partial \theta_1} = \frac{1}{3}\left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i)(x_i) - 2y_i x_i\right)\right].$$

Initial guess $(\theta_0, \theta_1) = (0,0)$ First iteration

$$\Rightarrow \frac{\partial MSE}{\partial \theta_0} = \frac{1}{3}\left[\sum_{i=1}^{3}\left(2(\theta_0+\theta_1 x_i) - 2y_i\right)\right]$$

$$= \frac{1}{3}\left[\sum_{i=1}^{3}\left(2(0+0x_i) - 2y_i\right)\right]$$

$$= \frac{1}{3}\left[\sum_{i=1}^{3}(-2y_i)\right]$$

$$= \frac{-2}{3}\left[\sum_{i=1}^{3} y_i\right] = \frac{-2}{3}\left[6+10+16\right]$$

$$= \frac{-2}{3}(32)$$

$$= -\frac{64}{3}.$$

$$\frac{\partial MSE}{\partial \theta_1} = \frac{1}{3}\left[\sum_{i=1}^{3}\left(2(\theta_0+\theta_1 x_i)(x_i) - 2y_i x_i\right)\right].$$

$$= \frac{1}{3}\left[\sum_{i=1}^{3}\left(2(0)(x_i) - 2y_i x_i\right)\right]$$

$$= -\frac{2}{3}\left[\sum_{i=1}^{3} x_i y_i\right]$$

$$= -\frac{2}{3}\left[(1)(6) + (3)(10) + (6)(16)\right]$$

$$= -\frac{2}{3}\left[6+30+96\right] = \frac{132}{3}(-2) = \underline{\underline{-88}}$$

updating θ_0, θ_1

$$\theta_0 = \theta_0 - \alpha \frac{\partial(MSE)}{\partial \theta_0}$$

$$= 0 - (0.1)\left[-\frac{64}{3}\right] = 2.134$$

$$\theta_1 = \theta_1 - \alpha \frac{\partial(MSE)}{\partial \theta_1}$$

$$= 0 - (0.1)(-88) = 8.8.$$

Second Iteration

$(\theta_0, \theta_1) = (2.134, 8.8).$

$$\frac{\partial(MSE)}{\partial \theta_0} = \frac{1}{3}\left[\sum_{i=1}^{3}\left(2(\theta_0 + \theta_1 x_i) - 2y_i\right)\right]$$

$$= 41.6$$

$$\frac{\partial(MSE)}{\partial \theta_1} = \frac{1}{3}\left[\sum_{i=1}^{3}\left(2(\theta_0 + \theta_1 x_i)(x_i) - 2y_i x_i\right)\right]$$

$$= \frac{1}{3}\left[2\left[2.134 \times 10 + 8.8[46]\right] - 2(132)\right]$$

$$= \frac{2}{3}\left[21.34 + 404.8 - 132\right] = 196.09.$$

updating (θ_0, θ_1).

$$\theta_0 = \theta_0 - \alpha\left(\frac{\partial MSE}{\partial \theta_0}\right)$$

$$= 2.134 - (0.1)(41.6)$$

$$= -2.02667$$

$$\theta_1 = \theta_1 - \alpha\left(\frac{\partial MSE}{\partial \theta_1}\right)$$

$$= 8.8 - (0.1)(196.09)$$

$$= 8.8 - 19.609$$

$$= -10.809.$$

$$(\theta_0, \theta_1) = (-2.026, -10.809)$$

Third iteration

$$\frac{\partial (MSE)}{\partial \theta_0} = \frac{1}{3}\left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i) - 2y_i\right)\right]$$

$$= -97.44$$

$$\frac{\partial (MSE)}{\partial \theta_1} = \frac{1}{3}\left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i)(x_i) - 2y_i x_i\right)\right]$$

$$= -432.984$$

$$\begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} -2.0267 \\ -10.8089 \end{bmatrix} - 0.1 \begin{bmatrix} -97.44 \\ -432.984 \end{bmatrix}$$

$$= \begin{bmatrix} 7.7 \\ 32.4895 \end{bmatrix}$$

$$(\theta_0, \theta_1) = (7.7, 32.4895)$$

Fourth iteration

$$\frac{\partial (MSE)}{\partial \theta_0} = \frac{1}{3} \left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i) - 2y_i \right) \right]$$

$$= 210.7$$

$$\frac{\partial (MSE)}{\partial \theta_1} = \frac{1}{3} \left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i)(x_i) - 2y_i x_i \right) \right]$$

$$= 959.797$$

$$\begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} 7.7 \\ 32.4895 \end{pmatrix} - 0.1 \begin{pmatrix} 210.7 \\ 959.797 \end{pmatrix}$$

$$\begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} -13.35 \\ -63.49 \end{pmatrix}.$$

$$(\theta_0, \theta_1) = (-13.35, -63.49)$$

fifth iteration.

$$\frac{\partial (MSE)}{\partial \theta_0} = \frac{1}{3}\left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i) - 2y_i\right)\right]$$

$$= -471.3054.$$

$$\frac{\partial (MSE)}{\partial \theta_1} = \frac{1}{3}\left[\sum_{i=1}^{3} \left(2(\theta_0 + \theta_1 x_i)(x_i) - 2y_i x_i\right)\right]$$

$$= -2124.0464.$$

$$\begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} 33.7786 \\ 148.914 \end{pmatrix}$$

(c) $\theta_1 = \dfrac{Cov(X,Y)}{Var(X)}$.

$\theta_0 = \bar{y} - \theta_1 \bar{x}$

$Cov(X,Y) = E(XY) - E[X]E[Y]$.

$E[X] = \dfrac{1+3+6}{3} = 10/3$.

$E[Y] = \dfrac{6+10+16}{3} = \dfrac{32}{3}$

$E[XY] = \dfrac{132}{3}$.

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$
$$= \frac{132}{3} - \frac{10}{3} \times \frac{32}{3}$$
$$= \frac{76}{9}.$$

$$Var(X) = E(X^2) - (E(X))^2$$
$$= \frac{1+9+36}{3} - \left(\frac{10}{3}\right)^2$$
$$= 38/9$$

$$\theta_1 = \frac{Cov(X, Y)}{Var(X)} = \underline{\underline{2}}$$

$$\theta_0 = E[Y] - \theta_1 E[X]$$

$$= \frac{32}{3} - \frac{2 \times 10}{3}$$

$$= \frac{12}{3} = 4$$

$$\underline{\underline{(\theta_0, \theta_1) = (4, 2)}}.$$

2. (b) $X = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \\ 4 & 8 \end{bmatrix}$

$Y = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$

$(X^T X)^{-1} X^T Y = \theta.$

where $X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 3 & 6 \\ 1 & 4 & 8 \end{bmatrix}$

$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 3 & 6 \\ 1 & 4 & 8 \end{bmatrix}$

$$= \begin{array}{c} \begin{bmatrix} 4 & 10 & 20 \\ 10 & 30 & 60 \\ 20 & 60 & 120 \end{bmatrix} \\ \quad\quad C_2 \quad\quad C_3 \end{array}$$

dependent on each other.

$\Rightarrow$ Multicollinearity $\Rightarrow$ $X^T X$ is not invertible