Linear Regression

a) Find the line equation.

b) Compare with the closed-form

Student	x_i	y_i
1	95	85
2	85	95
3	80	70
4	70	65
5	60	70

the line equation is $y = ax + b$

$\hat{y}_1 = 95a + b$

$\hat{y}_2 = 85a + b$

$\hat{y}_3 = 80a + b$

$\hat{y}_4 = 70a + b$

$\hat{y}_5 = 60a + b$

$$\Rightarrow \text{Error} = E(a,b) = \sum_{i=1}^{5} (\hat{y}_i - y_i)^2$$

$$= (95a + b - 85)^2 + (85a + b - 95)^2 + (80a + b - 70)^2$$

$$+ (70a + b - 65)^2 + (60a + b - 70)^2$$

to be optimized $\begin{cases} \frac{\partial E}{\partial a} = 0 \\ \frac{\partial E}{\partial b} = 0 \end{cases}$

$$E(a,b) = (95a+b-85)^2 + (85a+b-95)^2 + (80a+b-70)^2$$
$$+ (70a+b-65)^2 + (60a+b-70)^2$$

$$\frac{\partial E}{\partial a} = 2(95)(95a+b-85) + 2(85)(85a+b-95) + 2(80)(80a+b-70)$$
$$+ 2(70)(70a+b-65) + 2(60)(60a+b-70) = 0$$

$$\Rightarrow (95^2+85^2+80^2+70^2+60^2)a + (95+85+80+70+60)b$$
$$-(95\times85 + 85\times95 + 80\times70 + 70\times65 + 60\times70) = 0$$

$$\Rightarrow \boxed{3115a + 39b = 3050}$$

$$E(a,b) = (95a+b-85)^2 + (85a+b-95)^2 + (80a+b-70)^2$$
$$+ (70a+b-65)^2 + (60a+b-70)^2$$

$$\frac{\partial E}{\partial b} = 2(95a+b-85) + 2(85a+b-95) + 2(80a+b-70)$$
$$+ 2(70a+b-65) + 2(60a+b-70) = 0$$

$$\Rightarrow (95+85+80+70+60)a + 5b - (85+95+70+65+70) = 0$$

$$\Rightarrow \boxed{390a + 5b = 385}$$

$$\Rightarrow \begin{cases} 3115a + 39b = 3050 \\ 390a + 5b = 385 \end{cases} \rightarrow \begin{cases} a = \frac{47}{73} \\ b = \frac{1955}{73} \end{cases}$$

Closed form solution

$$y = Ax + B$$

$$A = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$B = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

385 31150 390 30500

$$A = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

5 31150 152100

$$\Rightarrow A = \frac{97750}{3650} = \frac{1955}{73} = b \checkmark$$

Student	x_i	y_i
1	95	85
2	85	95
3	80	70
4	70	65
5	60	70