1 a)

$$E(y) = \int_S (Ax + b)\, p(x)\, dx$$

$$= \int_S Ax\, p(x) + b\, p(x)\, dx$$

$$= A \int_S x\, p(x)\, dx + b \int_S p(x)\, dx = AE(x) + b$$

b) $\text{cov}(x) = E((x - E(x))(x - E(x))^T$

$$\text{cov}(Ax + b) = E((Ax + \cancel{b} - \underbrace{E(Ax+b)}_{(AE(x) + \cancel{b})})(Ax + \cancel{b} - \underbrace{E(Ax+b)}_{(AE(x) + \cancel{b})})^T$$

$$= E((Ax - AE(x))(Ax - AE(x))^T A^T$$

$$AE(((\cancel{x} - E(x))(x - E(x)^T) A^T$$

$$= AE\ \text{cov}(x)\, A^T = A \Sigma A^T$$

2. $X^T X \theta = X^T y$ $\qquad X = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 8 \end{bmatrix}$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 & 9 \\ 9 & 29 \end{bmatrix}$$

$$X^T y = \begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 2 & 3 & 4\end{bmatrix}\begin{bmatrix}1\\3\\6\\8\end{bmatrix} \begin{bmatrix}18\\56\end{bmatrix}$$

now we apply crame's rule

$$\theta_0^* = \frac{\det\begin{bmatrix}18 & 9\\ 56 & 29\end{bmatrix}}{\det\begin{bmatrix}4 & 9\\ 9 & 29\end{bmatrix}} = \frac{18}{35} \qquad \theta_1^* = \frac{\begin{bmatrix}4 & 18\\ 9 & 56\end{bmatrix}}{\begin{bmatrix}4 & 9\\ 9 & 29\end{bmatrix}} = \frac{62}{35}$$

$$y = \frac{18}{35} + \frac{62}{35}x$$

b) $\theta^* = (x^T x)^{-1} x^T y$

$$= \begin{bmatrix}4 & 9\\ 9 & 29\end{bmatrix}^{-1}\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 2 & 3 & 4\end{bmatrix}\begin{bmatrix}1\\3\\6\\8\end{bmatrix}$$

↓

calculator

↓

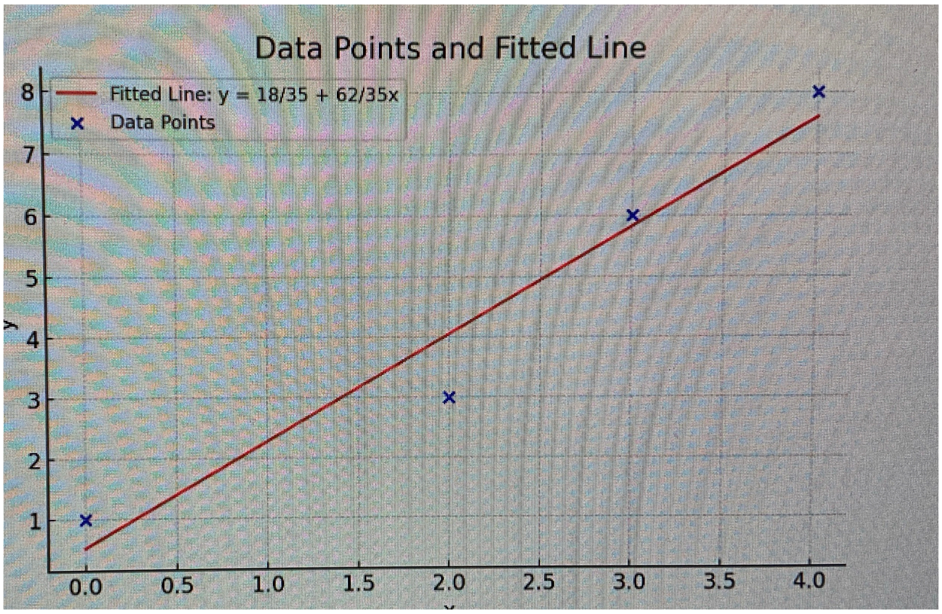
$$\frac{1}{35}\begin{bmatrix}29 & -9\\ -9 & 4\end{bmatrix}\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 2 & 3 & 4\end{bmatrix}\begin{bmatrix}1\\3\\6\\8\end{bmatrix}$$

↓

$$\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 2 & 3 & 4\end{bmatrix}\begin{bmatrix}1\\3\\6\\8\end{bmatrix}$$

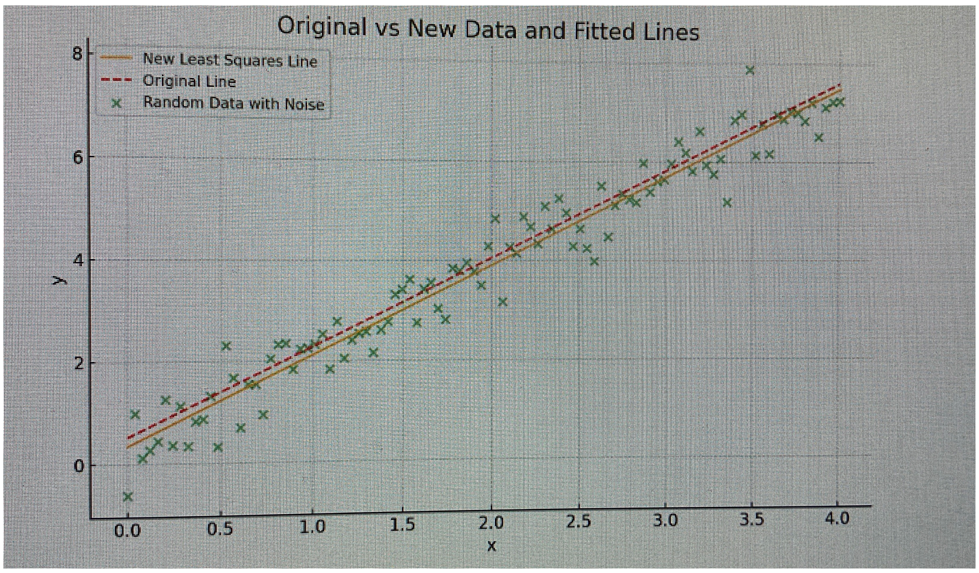
$$\frac{1}{35}\begin{bmatrix}29 & -9\\ -9 & 4\end{bmatrix}\begin{bmatrix}29 & 11 & 2 & -7\\ -9 & -1 & 3 & 7\end{bmatrix}\begin{bmatrix}18\\62\end{bmatrix} = \begin{bmatrix}\frac{18}{35}\\ \frac{62}{35}\end{bmatrix}$$ → same as above

c)



Data Points and Fitted Line
Fitted Line: y = 18/35 + 62/35x
Data Points
y
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
1
2
3
4
5
6
7
8

d)



Original vs New Data and Fitted Lines
New Least Squares Line
Original Line
Random Data with Noise
y
x
0
2
4
6
8
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0