(Ans 4)

linear regression model for the case of a single input and output variable:

$$y = f_w(x) = w_0 + w_1 x.$$

We have n observations
given the training data (x_i, y_i) for $i = 1, 2, 3, \ldots, N$.
with 1 input feature, the labels X and Y will be

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{(n \times 1)} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{(n \times 1)} \qquad W = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

But if we want to find both w_0 and w_1 we need to insert a constant feature column (usually 1) to X.

$$\therefore X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{(n \times 2)} \qquad W = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}_{(2 \times 1)}$$

Error associated with the regressor will be:

$$E = \sum_{i=1}^{N} f_w(x_i) - y_i$$

$$E = \begin{bmatrix} f_w(x_1) - y_1 \\ \vdots \\ f_w(x_N) - y_N \end{bmatrix} = \begin{bmatrix} x_1^T W \\ \vdots \\ x_N^T W \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = XW - Y.$$

Cost Function :

$$J(W) = \frac{1}{2} \sum_{i=1}^{N} \left(f_w(x_i) - y_i \right)^2 .$$

$$= \frac{1}{2} E^T E$$

$$= \frac{1}{2} (XW - Y)^T (XW - Y)$$

then to minimize the cost function (i.e error) we need to equate its derivative to 0.

$$\therefore \nabla_W (J(W)) = \nabla_W \frac{1}{2} E^T E$$

$$= \nabla_W \left(\frac{1}{2} (XW - Y)^T (XW - Y) \right)$$

$$= \frac{1}{2} \nabla_W \left(W^T X^T X W - W^T X^T Y - Y^T X W + Y^T Y \right)$$

$$\text{---------} \left((AB)^T = B^T A^T \right) \& \left((A-B)^T = A^T - B^T \right)$$

$$= \frac{1}{2} \nabla_W \left(W^T X^T X W - W^T X^T Y - W^T X^T Y + Y^T Y \right)$$

$$= \frac{1}{2} \nabla_W \left(W^T X^T X W - 2 W^T X^T Y + Y^T Y \right)$$

$$= \frac{1}{2} \cdot 2 \left(X^T X W - X^T Y \right)$$

$$= X^T X W - X^T Y = 0$$

$$\Rightarrow X^T X W = X^T Y$$

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

X is a $(n \times 2)$ dimensional vector.

$\therefore X^T$ will be a $(2 \times n)$ dimensional vector.

Y is a $(n \times 1)$ dimensional vector.

$$\therefore W = (X^T X)^{-1} X^T Y$$

~~$(n \times 2) \cdot$~~

$$= [(2 \times n) \cdot (n \times 2)]^{-1} \cdot (2 \times n) \cdot (n \times 1)$$

$$= (2 \times 2)^{-1} (2 \times n) \cdot (n \times 1)$$

$$= (2 \times 2) \cdot (2 \times n) \cdot (n \times 1)$$

$$= (2 \times n) \cdot (n \times 1)$$

$$= 2 \times 1$$

W is a (2×1) dimensional vector.

$$\Rightarrow \boxed{W = \begin{bmatrix} W_0 \\ W_1 \end{bmatrix}}$$