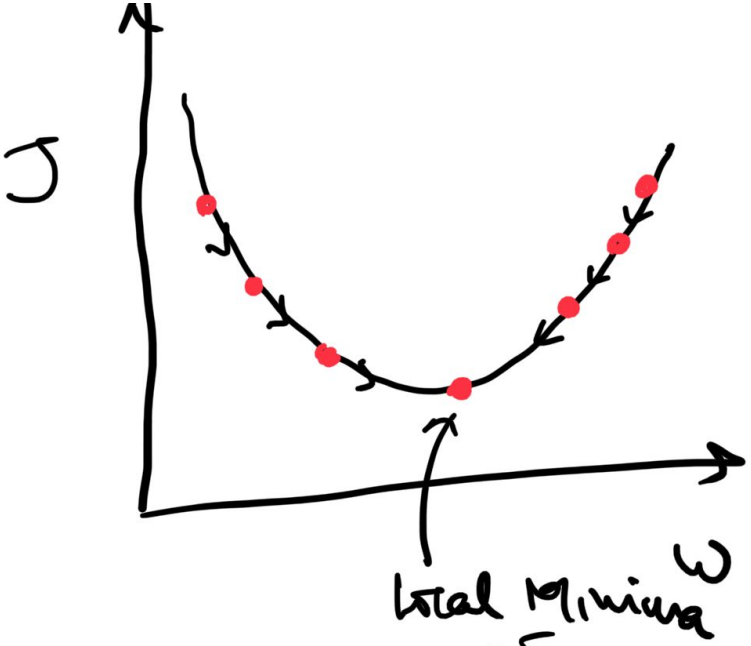
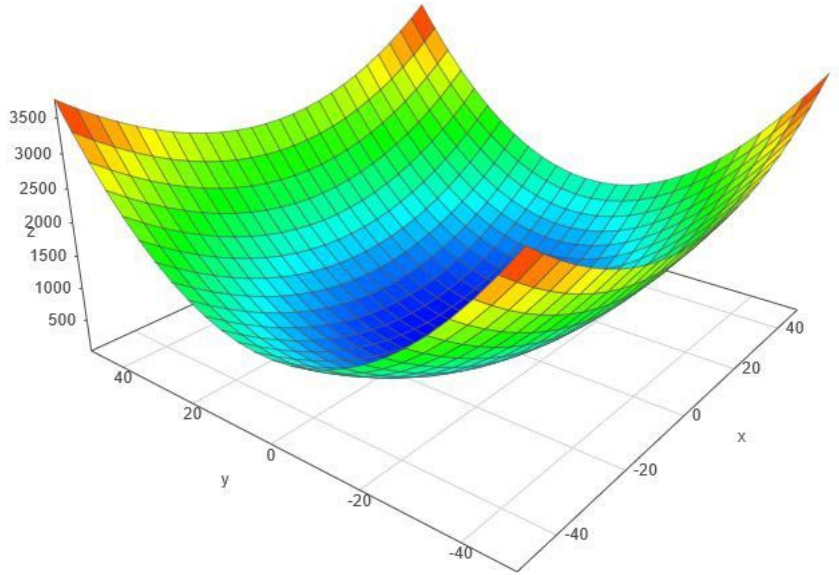
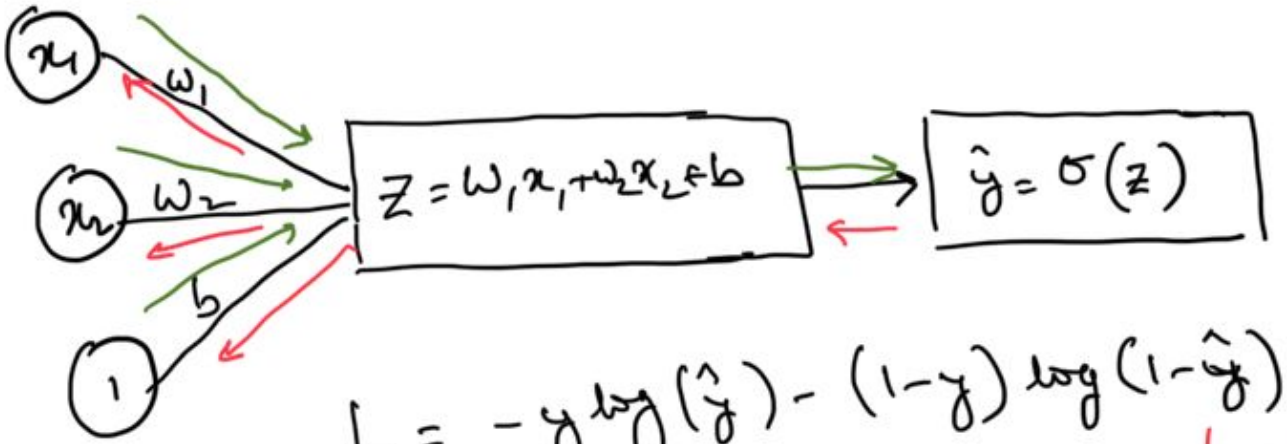




$$w = w - \alpha \frac{\partial J}{\partial w}$$

α = Learning Rate

$$b = b - \alpha \frac{\partial J}{\partial b}$$



$$L = -y \log(\hat{y}) - (1-y) \log(1-\hat{y})$$

↳ Loss function

Loss is at an individual record level.

$$C = \frac{1}{m} * \sum_{i=1}^{m} L$$

↳ Cost function

Average across the data set for the loss

By Gradient descent

$$w_1 = w_1 - \alpha \frac{\partial L}{\partial w_1} \rightarrow (1)$$

$$w_2 = w_2 - \alpha \frac{\partial L}{\partial w_2} \rightarrow (2)$$

$$b = b - \alpha \frac{\partial L}{\partial b} \rightarrow (3)$$

$\hat{y} \Rightarrow a$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial w_1}$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial w_2}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial b}$$

$$L = -y \log a - (1-y) \log (1-a)$$

$$\frac{\partial L}{\partial a} = \frac{-y}{a} - \frac{(1-y)}{(1-a)} \times (-1)$$

$$= \frac{-y}{a} + \frac{(1-y)}{(1-a)} = \frac{a(1-y) - (1-a)y}{a(1-a)}$$

$$= \frac{a - ay - y + ay}{a(1-a)} = \frac{a-y}{a(1-a)} \rightarrow (4)$$

$$\boxed{\frac{\partial L}{\partial a} = \frac{a-y}{a(1-a)}} \rightarrow (4)$$

$a = \sigma(z)$

$$\frac{\partial a}{\partial z} = \sigma(z)\,(1-\sigma(z)) = a\,(1-a) \quad \rightarrow (5)$$

$$\frac{\partial z}{\partial w_1} = x_1 \qquad \frac{\partial z}{\partial w_2} = x_2 \,, \quad \frac{\partial z}{\partial b} = 1 \quad \rightarrow (6)$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial w_1}$$

$$= \frac{(a-y)}{a(1-a)} \times a(1-a) \times x_1$$

$$= (a-y)\, x_1$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial w_2}$$

$$= \frac{(a-y)}{a(1-a)} \times a(1-a) \times x_2$$

$$= (a-y)\, x_2$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial a} * \frac{\partial a}{\partial z} * \frac{\partial z}{\partial b}$$

$$= \frac{(a-y)}{A(1-a)} * A(1-a) * 1$$

$$= (a-y)$$

Final Gradients

$$\frac{\partial L}{\partial w_1} = (a-y)x_1$$

$$\frac{\partial L}{\partial w_2} = (a-y)x_2$$

$$\frac{\partial L}{\partial b} = (a-y)$$