

## Logistic Regression.

Let  $\vec{W}$  be the weight vector  
 $b$  be the bias and  
 $\vec{x}$  be the input vector  
 $y$  be the true label

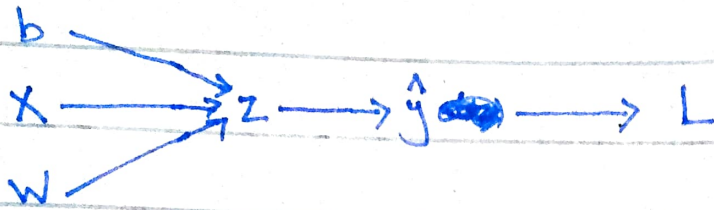
$$z = W^T X + b$$

$\hat{y} = \sigma(z) \rightarrow \text{prediction} \rightarrow (5)$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Error function is defined as

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



$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} \cdot \frac{\partial z}{\partial w} \rightarrow \textcircled{1}$$

Calculating each term separately

$$z = \vec{w}^T \vec{x} + b$$

$$\boxed{\frac{\partial z}{\partial w} = \vec{x}} \rightarrow \textcircled{2}$$

$$\boxed{\frac{\partial \hat{y}}{\partial z} = \sigma(z) [1 - \sigma(z)]} \rightarrow \textcircled{3}$$

$$\frac{\partial L}{\partial \hat{y}} = - \left[ \frac{\partial}{\partial \hat{y}} y \log \hat{y} + \frac{\partial}{\partial \hat{y}} (1-y) \log (1-\hat{y}) \right]$$

$$= - \left[ \frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} (-1) \right]$$

$$\boxed{\frac{\partial L}{\partial \hat{y}} = \frac{1-y}{1-\hat{y}} - \frac{y}{\hat{y}}} \rightarrow \textcircled{4}$$

putting ② ③ ④ in ①

$$\frac{\partial L}{\partial W} = \left[ \frac{1-y}{1-\hat{y}} - \frac{y}{\hat{y}} \right] \sigma'(z) [1-\sigma(z)] \vec{X}$$

by eq ⑤

$$= \left[ \frac{1-y}{1-\hat{y}} - \frac{y}{\hat{y}} \right] \hat{y} [1-\hat{y}] \vec{X}$$

$$= [\hat{y} [1-y] - [1-\hat{y}] y] \vec{X}$$

$$= [\hat{y} - y\hat{y} - y + y\hat{y}] \vec{X}$$

$$\boxed{\frac{\partial L}{\partial W} = [\hat{y} - y] \vec{X}} \rightarrow \text{⑥}$$

Gradient Step is defined by  
Learning Rate

$$\boxed{W_1 = W_0 - \alpha \frac{\partial L}{\partial W}}$$



Now Calculating for bias  $b$ .

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z} \frac{\partial z}{\partial b}$$

$$z = W^T X + b$$

$$\frac{\partial z}{\partial b} = 1$$

$$\frac{\partial L}{\partial b} = [\hat{y} - y] \cdot 1$$

gradient step for bias.

$$b_1 = b_0 - \alpha \frac{\partial L}{\partial b}$$