Logistic Regression Overview :

Equations :

$W = \begin{bmatrix} w_1 \\ w_2 \\ . \\ . \\ w_n \end{bmatrix}_{nx1}$ ..... initialize with zeros

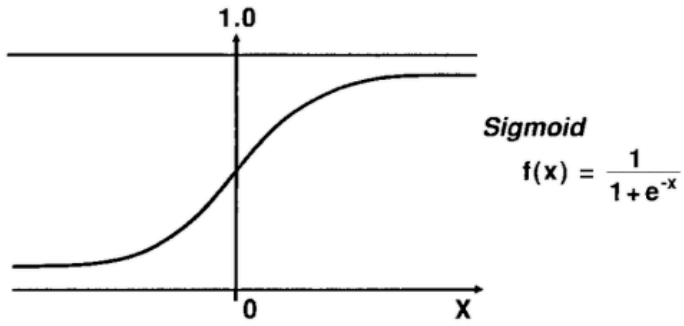
$B = single\ weight/parameter$

$X = \begin{bmatrix} . & . & . \\ . & . & . \\ . & . & . \\ . & . & . \end{bmatrix}_{nxm}$

$Y = \begin{bmatrix} . & . & . & . & . \end{bmatrix}_{1xm}$

$\sigma = \frac{1}{(1+e^{-x})}$ ..... (sigmoid function)

A = $\sigma(W^T * X + b)$ ..... (probabilistic predictions of shape (1 x m))



Cost function :

$cost = -\frac{1}{m} \sum_{i=1}^{m} [y * log(a) + (1 - y) * log(1 - a)]$

Gradient Descent

$dW = \frac{\partial COST}{\partial W} = (A - Y) * X^T$ ..... shape (1 x n)

$dB = \frac{\partial COST}{\partial B} = (A - Y)$

$W = W - \alpha * dW^T$

$B = B - \alpha * dB$