$$C(w) = -\frac{1}{m} \sum y_i \ln(h(x_i)) + (1-y_i) \ln(1-h(x_i))$$

donde $h(x_i) = \dfrac{1}{1+e^{-w^T x_i}}$

Si vemos la función sigmoide por cada dato en x, tenemos

$$\left[\frac{1}{1+e^{-w^T x_1}}, \frac{1}{1+e^{-w^T x_2}}, \frac{1}{1+e^{-w^T x_3}}, \dots, \frac{1}{1+e^{-w^T x_n}}\right]$$

esto se puede representar como

$$h([-w^T x_1, -w^T x_2, \dots, -w^T x_n])$$

si convertimos h en un vector con todos los valores de x en la función sigmoide, podemos representar la suma como una multiplicación de matrices.

Definimos

$X = [x_1, x_2, x_3, \dots, x_n]^T$, podemos decir que

$y^T = [y_1, y_2, \dots, y_n]$ y $h(Xw) = \begin{bmatrix} h(x_1 w) \\ h(x_2 w) \\ \vdots \\ h(x_n w) \end{bmatrix}$, entonces

$$y^T \log(h(Xw)) = [y_1, y_2, \dots, y_n] \begin{bmatrix} \log h(x_1 w) \\ \log h(x_2 w) \\ \log h(x_3 w) \\ \vdots \\ \log h(x_n w) \end{bmatrix}$$

$$= y_1 \log(h(x_1 w)) + y_2 \log(h(x_2 w)) + \dots$$

$$+ y_n \log(h(x_n w))$$

$$= \sum_{i=1}^{n} y_i \log(h(x_i)) \quad [1]$$

lo mismo para la segunda parte

$(1-y)^T$ donde 1 es una matriz de un , tenemos

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1-y_1 \\ 1-y_2 \\ \vdots \\ 1-y_n \end{bmatrix} \Longrightarrow (1-y)^T = \begin{bmatrix} 1-y_1 \\ 1-y_2 \\ \vdots \\ 1-y_n \end{bmatrix}^T$$

$$(1-y)^T \log(1-h(Xw)) = [1-y_1, 1-y_2, \dots, 1-y_n] \begin{bmatrix} \log(1-h(x_1 w)) \\ \log(1-h(x_2 w)) \\ \vdots \\ \log(1-h(x_n w)) \end{bmatrix}$$

$$= (1-y_1)\log(1-h(x_1w)) + (1-y_2)\log(1-h(x_2w)) + \dots$$

$$+ (1-y_n)\log(1-h(x_nw))$$

$$= \sum_{i=1}^{n} (1-y_i)\log(1-h(x_i) \qquad \underline{[2]}$$

$\therefore$ Juntando [1] y [2] tenemos

$$\overbrace{y^T\log(h(Xw)) + (\bar{1}-y)^T\log(1-h(Xw))}^{a} =$$

$$\underbrace{\sum y_i\log(h(x_i)) + (1-y_i)\log(1-h(x_i))}_{a}$$

$$\therefore -\frac{1}{n}a = -\frac{1}{n}a$$