

1.a

Defino el conjunto de datos X_i, y_i

con X_i : features

y_i : target binario ($y_i \in \{0, 1\}$)

$$P(y_i = 1 / X_i, w) = \sigma(w^T x_i) ; \sigma(z) = \frac{1}{1 + e^{-z}} \quad (\text{función sigmoide})$$

↳ Esta es la probabilidad de que el target sea 1 dados x_i, w .

Lo que nos deja:

$$P(y_i = 0 / X_i, w) = 1 - \sigma(w^T x_i)$$

Apliko el principio de Verosimilitud:

$$L(w) = \prod_{i=1}^N P(y_i / X_i, w) = \prod_{i=1}^N \sigma(w^T x_i)^{y_i} \cdot (1 - \sigma(w^T x_i))^{1-y_i}$$

A esta fórmula le aplico el logaritmo para facilitar las cuentas necesarias para optimizarla (hallar w / maximiza la probabilidad)

$$\log L(w) = \sum_{i=1}^N y_i \log(\sigma(w^T x_i)) + (1 - y_i) \log(1 - \sigma(w^T x_i))$$

Aplicando L_2 : $R(w) = \frac{\lambda}{2} \|w\|^2$ (λ hiperparámetro que controla la reg.)

Para el problema a

uno de minimización: defino $\ell(w) = -\frac{1}{N} \log L(w)$

• Se dividió por N para trabajar el promedio

Ahora si, la función de costo es $\ell(w) = -\frac{1}{N} \log L(w) + R(w)$

$$\hookrightarrow \ell(w) = -\frac{1}{N} \sum_{i=1}^N \left[y_i \log(\sigma(w^T x_i)) + (1 - y_i) \log(1 - \sigma(w^T x_i)) \right] + \frac{\lambda}{2} \|w\|^2$$

• Función de Costo: Binary Cross Entropy con Regularización L_2