

---

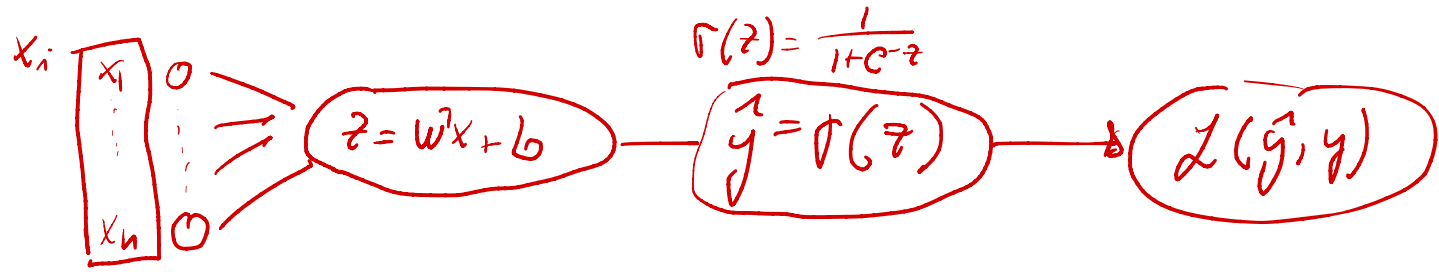
---

---

---

---





$$J = -\frac{1}{n} \sum_{i=1}^n \left[ y_i \log(\hat{y}_i) + (1-y_i) \log(1-\hat{y}_i) \right]$$

$$\frac{\partial J}{\partial \hat{y}} = -2 \left[ y \log(\hat{y}) + (1-y) \log(1-\hat{y}) \right] = -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}}$$

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

$$\boxed{\frac{\hat{y} - y}{\hat{y}(1-\hat{y})}}$$

$$\frac{\partial \mathcal{L}}{\partial z} = \frac{\partial \mathcal{L}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} = \left( \frac{\hat{y} - y}{\hat{y}(1-\hat{y})} \right) \cdot \left( \hat{y}(1-\hat{y}) \right) = \boxed{\hat{y} - y}$$

$$\frac{\partial \hat{y}}{\partial z} = 2 \left[ \frac{1}{(1+e^{-z})} \right] = 2 \left[ (1+e^{-z})^{-1} \right] = 2 (1+e^{-z})^{-2} (e^{-z})$$

$$\boxed{\frac{e^{-z}}{(1+e^{-z})^2}} = \left( \frac{1}{1+e^{-z}} \right) \left( 1 - \frac{1}{1+e^{-z}} \right) = \boxed{\hat{y}(1-\hat{y})}$$

$$\frac{\partial \mathcal{L}}{\partial w} = \underbrace{\frac{\partial \mathcal{L}}{\partial \tilde{y}} \cdot \frac{\partial \tilde{y}}{\partial z}}_{\text{chain rule}} \cdot \frac{\partial z}{\partial w} = \boxed{(\hat{y} - y) \cdot x}$$

$$\frac{\partial z}{\partial w} = \frac{\partial}{\partial w} [w^T x + b] = x$$

$$\frac{\partial \mathcal{L}}{\partial b} = 1$$

$$\frac{\partial \mathcal{L}}{\partial b} = (\hat{y} - y) (1) = \boxed{\hat{y} - y}$$

# Algorithm (Des. Log)

for  $t=1$  to max\_iter :

$$z = w.T @ X + b$$

$$\hat{y} = \sigma(z)$$

$$dz = \hat{y} - y$$

$$dw = \frac{1}{n} X @ dz.T$$

$$db = \frac{1}{n} \text{np.sum}(dz)$$

$$w = w - lr * dw$$

$$b = b - lr * db$$

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

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