



deeplearning.ai

# Basics of Neural Network Programming

---

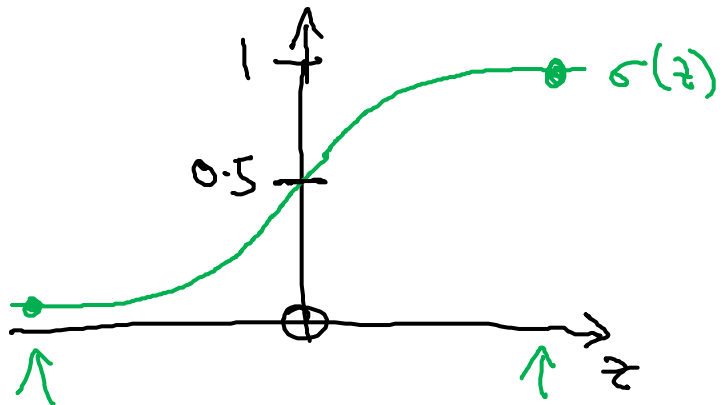
## Logistic Regression

# Logistic Regression

Given  $x$ , want  $\hat{y} = \frac{P(y=1|x)}{0 \leq \hat{y} \leq 1}$   
 $x \in \mathbb{R}^{n_x}$

Parameters:  $\underline{w} \in \mathbb{R}^{n_x}$ ,  $\underline{b} \in \mathbb{R}$ .

Output  $\hat{y} = \sigma(\underbrace{w^T x + b}_z)$



$$x_0 = 1, \quad x \in \mathbb{R}^{n_x+1}$$
$$\hat{y} = \sigma(\theta^T x)$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{n_x} \end{bmatrix} \quad \left. \begin{array}{l} \} b \leftarrow \\ \} w \leftarrow \end{array} \right\}$$

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

$$\text{If } z \text{ large } \sigma(z) \approx \frac{1}{1+0} = 1$$

If  $z$  large negative number

$$\sigma(z) = \frac{1}{1 + e^{-z}} \approx \frac{1}{1 + \text{Big num}} \approx 0$$



deeplearning.ai

# Basics of Neural Network Programming

---

## Logistic Regression cost function

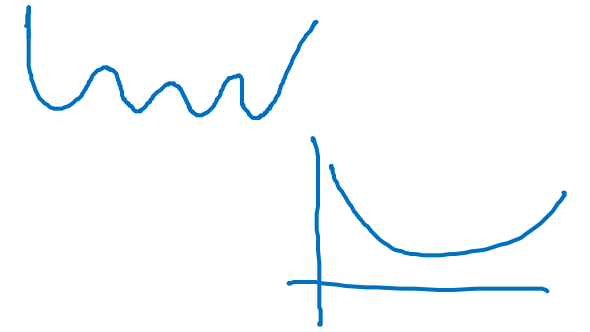
# Logistic Regression cost function

$$\rightarrow \hat{y}^{(i)} = \sigma(w^T \underline{x}^{(i)} + b), \text{ where } \sigma(z^{(i)}) = \frac{1}{1+e^{-z^{(i)}}} \quad z^{(i)} = w^T x^{(i)} + b$$

Given  $\{(\underline{x}^{(1)}, y^{(1)}), \dots, (\underline{x}^{(m)}, y^{(m)})\}$ , want  $\hat{y}^{(i)} \approx \underline{y}^{(i)}$ .

$x^{(i)}$   
 $y^{(i)}$   
 $z^{(i)}$   $i$ -th example.

Loss (error) function:  $\mathcal{L}(\hat{y}, y) = \frac{1}{2} (\hat{y} - y)^2$



$$\mathcal{L}(\hat{y}, y) = - (y \log \hat{y}) + (1-y) \log (1-\hat{y}) \leftarrow$$

If  $y=1$ :  $\mathcal{L}(\hat{y}, y) = -\log \hat{y} \leftarrow$  Want  $\log \hat{y}$  large, want  $\hat{y}$  large.

If  $y=0$ :  $\mathcal{L}(\hat{y}, y) = -\log (1-\hat{y}) \leftarrow$  Want  $\log (1-\hat{y})$  large ... want  $\hat{y}$  small

Cost function:  $J(w, b) = \frac{1}{m} \sum_{i=1}^m \mathcal{L}(\hat{y}^{(i)}, y^{(i)}) = \frac{1}{m} \sum_{i=1}^m [y^{(i)} \log \hat{y}^{(i)} + (1-y^{(i)}) \log (1-\hat{y}^{(i)})]$