1 Gradient Descent

$$\frac{dC(w_i)}{dw_i} = \lim_{\epsilon \to 0} \frac{C(w_i + \epsilon) - C(w_i)}{\epsilon} \tag{1}$$

1.1 "Twice"

$$C(w) = \frac{1}{n} \sum_{i=1}^{n} (x_i w - y_i)^2 \tag{2}$$

$$C'(w) = \left(\frac{1}{n} \sum_{i=1}^{n} (x_i w - y_i)^2 \right)' \tag{3}$$

$$= \frac{1}{n} \left(\sum_{i=1}^{n} (x_i w - y_i)^2 \right)' \tag{4}$$

$$= \frac{1}{n} \left((x_0 w - y_0)^2 + (x_1 w - y_1)^2 + (x_2 w - y_2)^2 + \ldots \right)' \tag{5}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left((x_i w - y_i)^2 \right)' \tag{6}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2 \left(x_i w - y_i \right) (x_i w - y_i)' \tag{7}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2 \left(x_i w - y_i \right) ((x_i w)' - (y_i)') \tag{8}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2 \left(x_i w - y_i \right) (x_i w)' \tag{9}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2 \left(x_i w - y_i \right) x_i \tag{10}$$

$$\tag{11}$$

$$C(w) = \frac{1}{n} \sum_{i=1}^{n} (x_i w - y_i)^2 \tag{12}$$

$$C'(w) = \frac{1}{n} \sum_{i=1}^{n} 2 \left(x_i w - y_i \right) x_i \tag{13}$$

$$\tag{14}$$

1.2 One Neuron Model

$$y = \sigma(xw + b) \tag{15}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

$$\sigma'(x) = \sigma(x)(1 - \sigma(x)) \tag{17}$$

$$\tag{18}$$