1 Gradient Descent

$$C'(w) = \lim_{\epsilon \to 0} \frac{C(w + \epsilon) - C(w)}{\epsilon} \tag{1}$$

1.1 "Double"

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

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

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

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

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

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

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

$$= \frac{1}{n} \sum_{i=1}^{n} 2(x_i w - y_i) x_i' \tag{9}$$
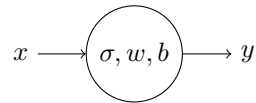
$$\tag{10}$$

$$C(w) = \frac{1}{n} \sum_{1=1}^{n} (x_i w - y_i)^2 \tag{11}$$

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

$$\tag{13}$$

1.2 One neuron model

$$y = \sigma(xw + b) \tag{14}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{15}$$

$$\sigma'(x) = \sigma(x)(1 - \sigma(x)) \tag{16}$$

$$\tag{17}$$

1.2.1 Cost

$$a_i = \sigma(x_i w + b) \tag{18}$$

$$\partial_w a_i = \partial_w \left(\sigma(x_i w + b) \right) \tag{19}$$

$$= a_i(1 - a_i)\partial_w(x_i w + b) \tag{20}$$

$$= a_i(1 - a_i)x_i \tag{21}$$

$$\partial_b a_i = a_i(1 - a_i) \tag{22}$$

$$C = \frac{1}{n} \sum_{i=1}^{n} (a_i - y_i)^2 \tag{23}$$

$$\partial_w C = \partial_w \left(\frac{1}{n} \sum_{i=1}^{n} (a_i - y_i)^2 \right) \tag{24}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \partial_w \left((a_i - y_i)^2 \right) \tag{25}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - y_i)\partial_w (a_i - y_i) \tag{26}$$

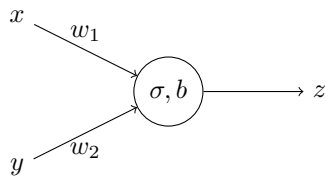$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - y_i)\partial_w a_i \tag{27}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - y_i)(a_i(1 - a_i)x_i) \tag{28}$$

$$\partial_b C = \frac{1}{n} \sum_{i=1}^{n} 2(a_i - y_i)\partial_b a_i \tag{29}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - y_i)(a_i(1 - a_i)) \tag{30}$$

$$\tag{31}$$

1.3 One Neuron Model with 2 inputs

$$z = \sigma(x_i w_1 + y_i w_2 + b) \tag{32}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{33}$$

$$\sigma'(x) = \sigma(x)(1 - \sigma(x)) \tag{34}$$

$$\tag{35}$$

1.3.1 Cost

$$a_i = \sigma(x_i w_1 + y_i w_2 + b) \tag{36}$$

$$\partial_{w_1} a_i = \partial_w \left(\sigma(x_i w_1 + y_i w_2 + b) \right) \tag{37}$$

$$= a_i(1 - a_i)\partial_w(x_i w_1 + y_i w_2 + b) \tag{38}$$

$$= a_i(1 - a_i)x_i \tag{39}$$

$$\partial_{w_2} a_i = a_i(1 - a_i)y_i \tag{40}$$

$$\partial_b a_i = a_i(1 - a_i) \tag{41}$$

$$C = \frac{1}{n} \sum_{i=1}^{n} (a_i - z_i)^2 \tag{42}$$

$$\partial_{w_1} C = \partial_{w_1} \left(\frac{1}{n} \sum_{i=1}^{n} (a_i - z_i)^2 \right) \tag{43}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \partial_{w_1} \left((a_i - z_i)^2 \right) \tag{44}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - z_i)\partial_{w_1} (a_i - z_i) \tag{45}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - z_i)\partial_{w_1} a_i \tag{46}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - z_i)(a_i(1 - a_i)x_i) \tag{47}$$

$$\partial_{w_2} C = \frac{1}{n} \sum_{i=1}^{n} 2(a_i - z_i)(a_i(1 - a_i)y_i) \tag{48}$$

$$\partial_b C = \frac{1}{n} \sum_{i=1}^{n} 2(a_i - z_i)\partial_b a_i \tag{49}$$

$$= \frac{1}{n} \sum_{i=1}^{n} 2(a_i - z_i)(a_i(1 - a_i)) \tag{50}$$

$$\tag{51}$$