

1 Gradient Descent

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

1.1 “Twice”

sequence of derivating $C(w)$ with respect to w .

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

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

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

$$= \frac{2}{n} \sum_{i=1}^n (x_i w - y_i)(x_i)' \quad (5)$$

$$(6)$$