

1 Gradient Descent

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

1.1 ”Twice”

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

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

$$= \frac{1}{n} \left((x_0 w - y_0)^2 + (x_1 w - y_1)^2 + \dots + (x_n w - y_n)^2 \right)' \quad (4)$$

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

$$= \frac{1}{n} \sum_{i=1}^n 2(x_i w - y_i)x_i \quad (6)$$

$$(7)$$