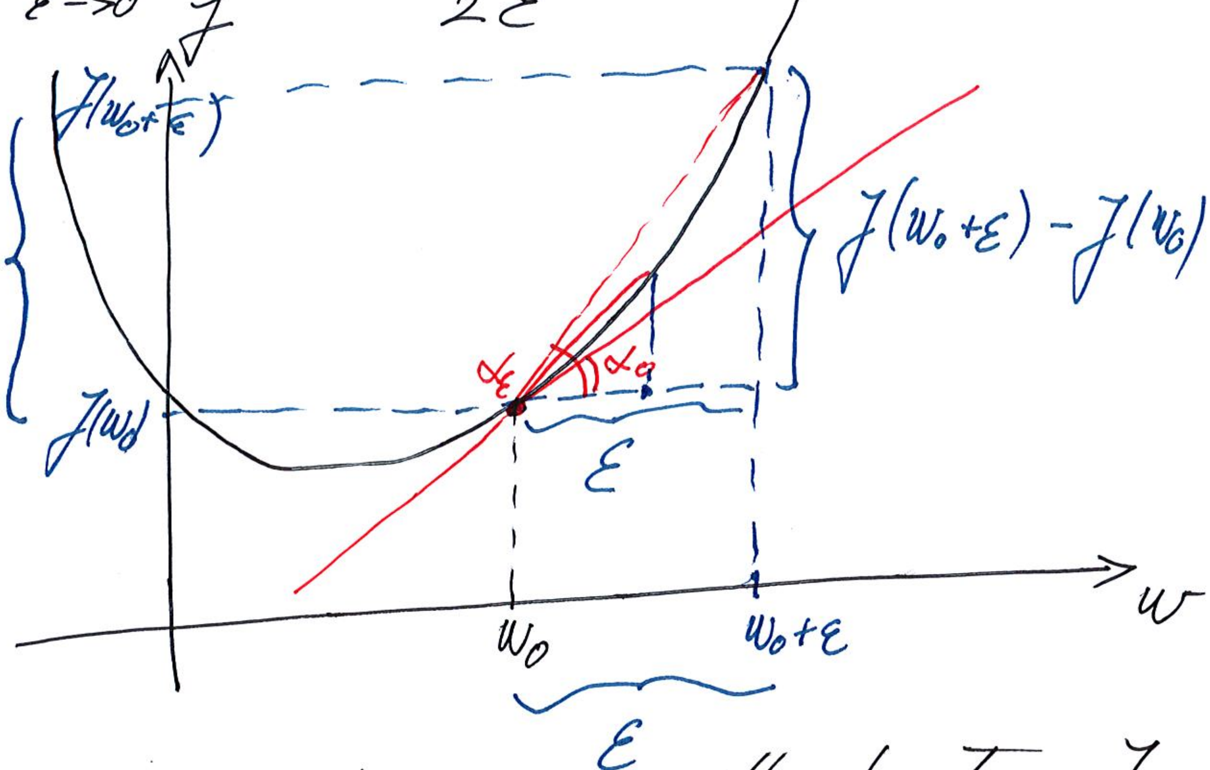
$J: R \longrightarrow R \qquad J(w) = \dots$

J is differentiable at $w = w_0$ iff $(\exists)$ $\lim\limits_{\varepsilon \to 0} \frac{J(w_0+\varepsilon) - J(w_0)}{\varepsilon}$ $\quad J'(w_0) \quad \frac{\partial J}{\partial w}(w_0)$

$$\lim_{\varepsilon \to 0} \frac{J(w_0+\varepsilon) - J(w_0-\varepsilon)}{2\varepsilon}$$



$$\tan(\alpha_\varepsilon) \xrightarrow[\varepsilon \to 0]{} \tan(\alpha_0) = J'(w_0).$$

$J'(w_0)$ gives ~~the~~ how fast the function J grows at w_0, as we increase w.

$J'(w_0) \simeq \underbrace{\text{the rate of}}_{\text{speed}}$ increase in J

$J: \mathbb{R}^M \longrightarrow \mathbb{R}$. $\quad w = [w_1, w_2, \dots, w_M] \quad \frac{\partial J}{\partial w} = \left[\frac{\partial J}{\partial w_1}, \dots, \frac{\partial J}{\partial w_M}\right]$

Start with $w^0 = 0$ (or some other guess, or at random).

Set time $t = 0$

repeat

Compute $\frac{\partial J}{\partial w}(w^0)$

$w^{t+1} = w^t - \eta \cdot \frac{\partial J}{\partial w}(w^t)$ $\rightarrow \left[\frac{\partial J}{\partial w_1}, \ldots, \frac{\partial J}{\partial w_M}\right]$

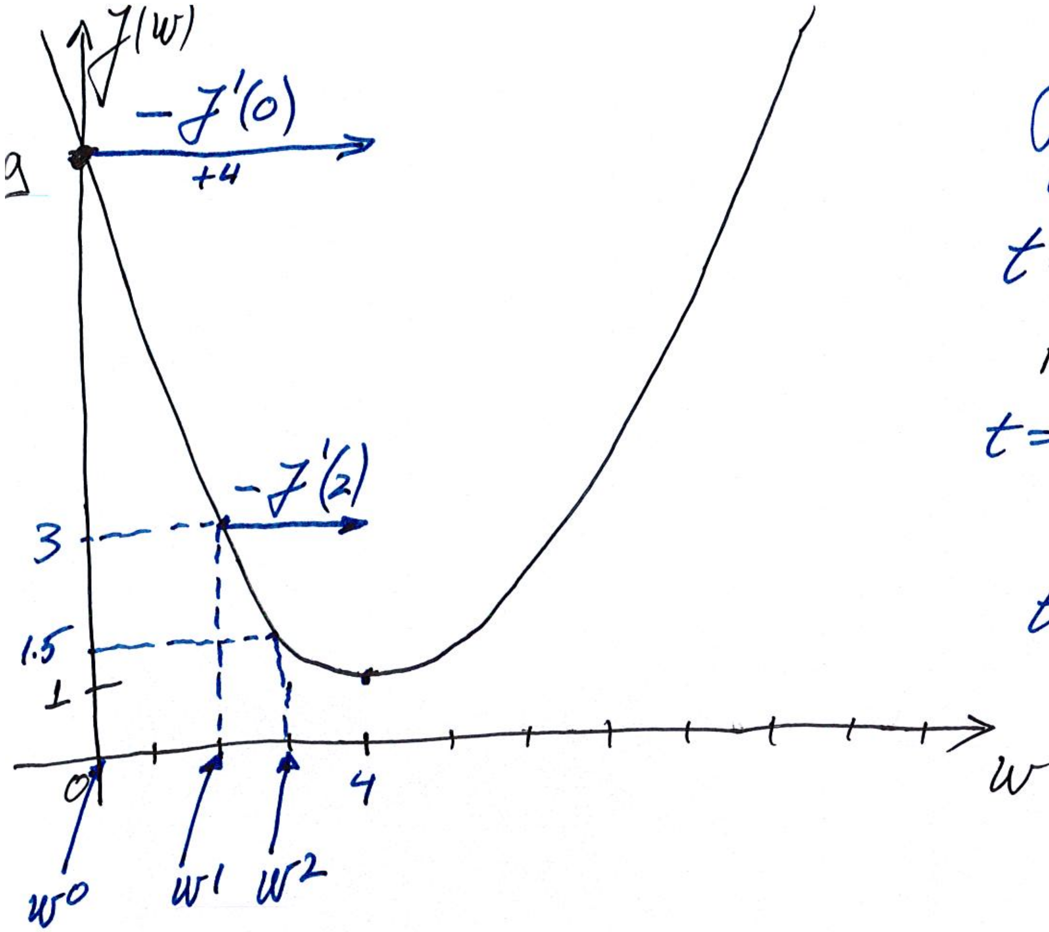
$t \leftarrow t+1$

until convergence or max epochs

$\hookrightarrow$ J does not change much $\dfrac{\overbrace{|J(w^t) - J(w^{t-1})|}^{\text{absolute diff.}}}{|J(w^{t-1})|} < \varepsilon$

0.001

relative diff.



$J(w) = \frac{1}{2}(w-4)^2 + 1 \qquad \frac{\partial J}{\partial w} = w - 4$

Use GD to find its minimum w/ $\eta = 0.5$

$t=0: \; w^0 = 0 \qquad J(w^0) = 9$

First GD step / gradient update

$t=1: \; w^1 = w^0 - 0.5(0-4) \qquad J'(w^0) = -4$

$= 0 + 2 = 2 \qquad J(w^1) = 3$

$t=2: \; w^2 = w^1 - 0.5(w^1 - 4)$

$= 2 - 0.5(2-4) \qquad J'(w^1) = -2$

$= 3$

$J(w^2) = 1.5$

Imagine $w^0 = 0, \; \eta = 2.$