〈Intuition1〉 - derivative

Gradient Descent Algorithm : repeat until convergence {

$$w := w - \alpha \frac{\partial}{\partial w} J(w,b)$$

$$b := b - \alpha \frac{\partial}{\partial b} J(w,b) \quad \}$$

α (learning rate) : controls how big of step you take when updating model's parameter w, b

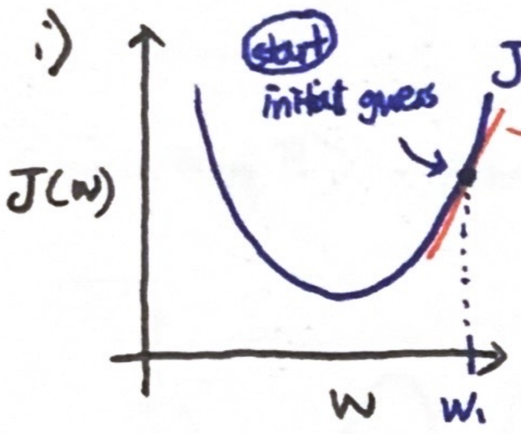
$\frac{\partial}{\partial w} J(w,b)$, $\frac{\partial}{\partial b} J(w,b)$: determines direction for the next step

how?

To simplify, just use $J(w)$ $(b=0)$

$\Rightarrow$ gradient descent $\left[w := w - \alpha \frac{\partial}{\partial w} J(w) \right]$

i)



slope line = derivative of function J at w_1 point

$$\left[w := w - \alpha \frac{\partial}{\partial w} J(w) \right]$$

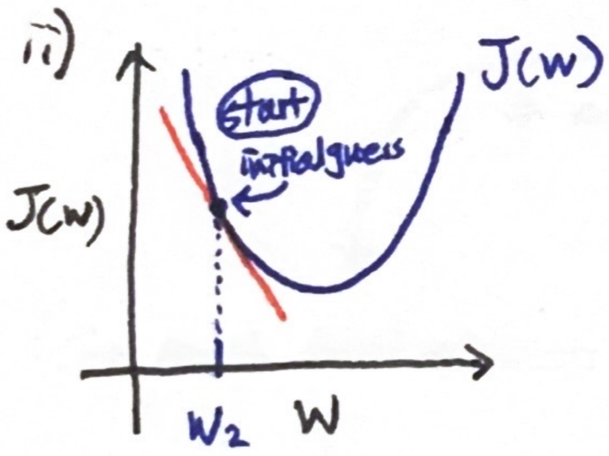
$\Downarrow$ positive number (∵ tangent line points up right)

updated $w = w - \alpha \cdot$ (positive number)

∴ updated w gets smaller than previous w

= decreasing value of w getting closer to the minimum for J

ii)



$$\left[w := w - \alpha \frac{\partial}{\partial w} J(w) \right]$$

$\Downarrow$ negative number (∵ tangent line points down right)

updated $w = w - \alpha \cdot$ (negative number)

∴ updated w gets bigger than previous w

= increasing value of w getting closer to the minimum for J