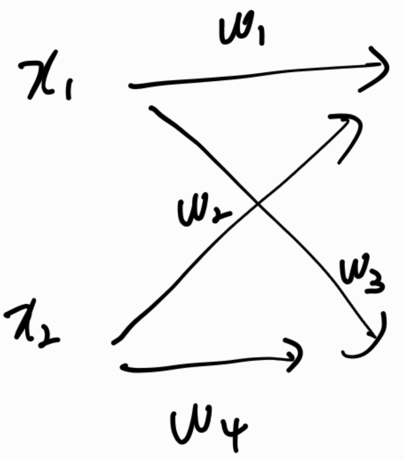
target vs output

activation function $= \dfrac{1}{1+e^{-x}}$

$H(x) = Wx + b$

7-5 예시 $(w_1 x_1 + b_1) + (w_2 x_2 + b_2) = z_1$



$$\frac{1}{1 + e^{(w_1 x_1 + b_1) + (w_2 x_2 + b_2)}} = h_1$$

$$(w_3 x_1 + b_3) + (w_4 x_2 + b_4) = z_2$$

$$\frac{1}{1 + e^{(w_3 x_1 + b_3) + (w_4 x_2 + b_4)}} = h_2$$

$$z_3 = w_5\left(\frac{1}{1 + e^{(w_1 x_1 + b_1) + (w_2 x_2 + b_2)}}\right) + b_5 + w_6\left(\frac{1}{1 + e^{(w_3 x_1 + b_3) + (w_4 x_2 + b_4)}}\right) + b_6$$

$$O_1 = \frac{1}{1 + e^{w_5\left(\frac{1}{1+e^{\cdots}}\right) + b_5 + w_6\left(\frac{1}{1+e^{\cdots}}\right) + b_6}}$$

$$z_4 = w_7\left(\frac{1}{1 + e^{(w_1 x_1 + b_1) + (w_2 x_2 + b_2)}}\right) + b_7 + w_8\left(\frac{1}{1 + e^{(w_3 x_1 + b_3) + (w_4 x_2 + b_4)}}\right) + b_8$$

$$O_2 = \frac{1}{1 + e^{w_7\left(\frac{1}{1+e^{\cdots}}\right) + b_7 + w_8\left(\frac{1}{1+e^{\cdots}}\right) + b_8}}$$

$$E_{o1} = \frac{1}{2}\left(target_{o1} - output_{o1}\right)^2$$

$$= \frac{1}{2}\left(target_{o1} - \frac{1}{1 + e^{w_5\left(\frac{1}{1+e^{\cdots}}\right) + b_5 + w_6\left(\frac{1}{1+e^{\cdots}}\right) + b_6}}\right)^2$$

$$\frac{1}{2}\left(target_{o_1} - \boxed{\frac{1}{1+e^{w_5\left(\frac{1}{1+e^{(w_1x_1+b_1)+(w_2x_2+b_2)}}\right)+b_5+w_6\left(\frac{1}{1+e^{(w_3x_1+b_3)+(w_4x_2+b_4)}}\right)+b_6}}}\right)^2 = E_1$$

$$\frac{1}{1+e^{\boxed{w_5\left(\frac{1}{1+e^{(w_1x_1+b_1)+(w_2x_2+b_2)}}\right)+b_5+w_6\left(\frac{1}{1+e^{(w_3x_1+b_3)+(w_4x_2+b_4)}}\right)+b_6}}} = O_1$$

$$w_5\left(\frac{1}{1+e^{(w_1x_1+b_1)+(w_2x_2+b_2)}}\right)+b_5+w_6\left(\frac{1}{1+e^{(w_3x_1+b_3)+(w_4x_2+b_4)}}\right)+b_6 = z_3$$

$$\frac{\partial E_1}{\partial w_5} = \frac{\partial E_1}{\partial o_1} \times \frac{\partial o_1}{\partial z_3} \times \frac{\partial z_3}{\partial w_5}$$

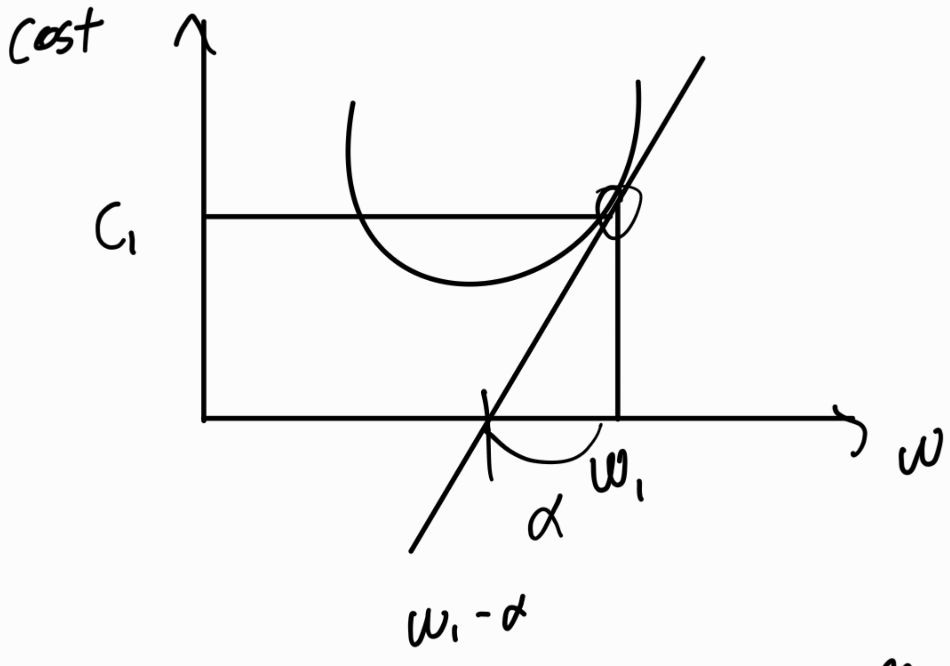
$$-(target_{o_1} - Output_{o_1})$$

$$h_1$$

$$Output_1 \times (1 - Output_1)$$

E total에서 바로 w_5 편미분 하려면

머리 터지니까 O_1 먼저.



$$cost(w,b) = \sum_{i=1}^{n} \frac{1}{n}\left((w x_i + b) - O_i\right)^2$$

b는 고정 가정

$$\frac{\partial\, cost(w,b)}{\partial w} = \sum_{i=1}^{n} \frac{1}{n}\left(2(w x_i + b)x_i\right) = \frac{C_1}{\alpha}$$

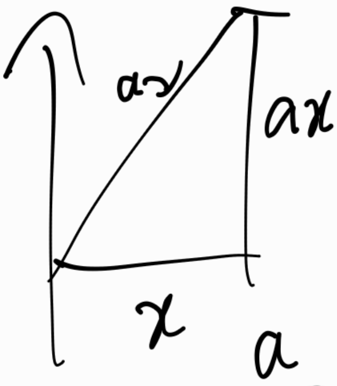
$$y' = (1, \infty),\ \beta = \frac{1}{y'}$$

β = learning rate = $(0, 1)$

$$\beta \quad \frac{2(w x_i + b) x_i}{\text{↳기울기}} = \frac{C_1}{\alpha}\beta = \frac{C_1}{\alpha}\frac{1}{y'} = \frac{1}{x'}$$

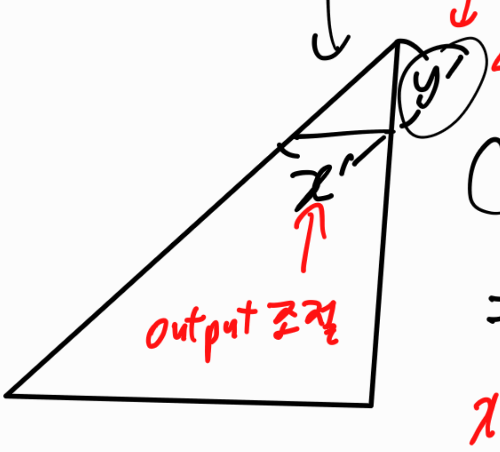
ex)

$$ax = f(x)$$



$$\frac{a}{f(x)} = \frac{1}{x}$$

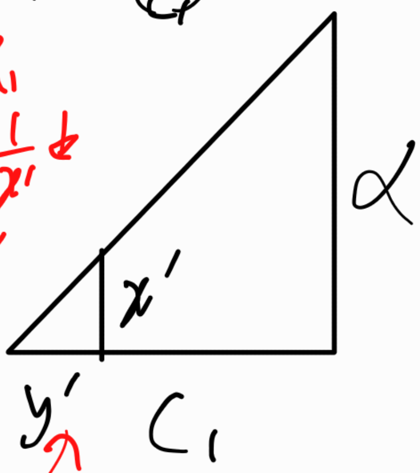
$x \neq 0$
$a \neq 0$.



$$x'\left(\frac{C_1}{\alpha}\right) = y'$$

$$x' = \left(\frac{\alpha}{C_1}\right) y'$$

C_1 $\beta\downarrow$ = $\frac{1}{x'}\downarrow$ = $x'\uparrow$



연결해서 이해하면 쉬움