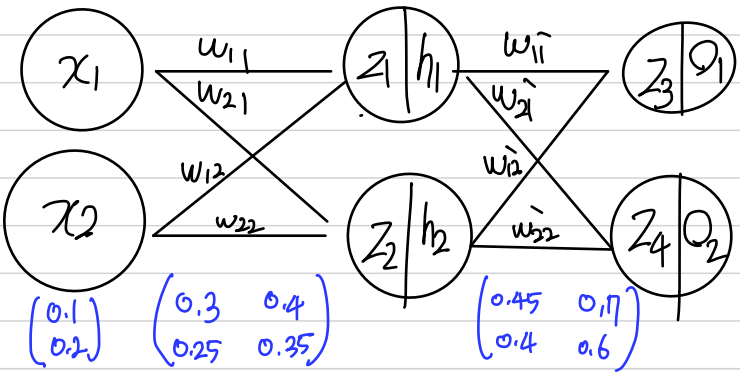
역전파.



$z_1 = x_1 \cdot w_{11} + x_2 \cdot w_{12} = 0.08$

$z_2 = x_1 w_{21} + x_2 \cdot w_{22} = 0.11$

$E_{o1} = \frac{1}{2}(target_1 - o_1)^2 = 0.021933$

$E_{o2} = \frac{1}{2}(target_2 - o_2)^2 = 0.02038$

$E_{total} = E_{o1} + E_{o2} = 0.023917$

Sigmoid $(z_1) = 0.51998 = h_1$ $\quad h_1 \cdot w_{11}' + h_2 \cdot w_{12}' = 0.14498 = z_3$

Sigmoid $(z_2) = 0.527747 = h_2$ $\quad h_1 \cdot w_{21}' + h_2 \cdot w_{22}' = 0.686417 = z_4$

Sigmoid $(z_3) = 0.60944 = o_1$

Sigmoid $(z_4) = 0.66384 = o_2$

역전파 1단계 $w_{11}', w_{21}', w_{12}', w_{22}'$

i) w_{11}'를 update

$\frac{\partial E_{total}}{\partial w_{11}'}$를 구한다. $\Rightarrow \underbrace{\frac{\partial E_{total}}{\partial o_1}}_{i)} \cdot \underbrace{\frac{\partial o_1}{\partial z_3}}_{ii)} \cdot \underbrace{\frac{\partial z_3}{\partial w_{11}'}}_{iii)}$

$= (0.20944)(0.23804)(0.51998)$

$= 0.02592 = \frac{\partial E_{total}}{\partial w_{11}'}$

i) $E_{total} = E_{o1} + E_{o2}$

$= \frac{1}{2}(target_1 - o_1)^2 + \frac{1}{2}(target_2 - o_2)^2$

$\frac{\partial E_{total}}{\partial o_1} = \frac{1}{2} \cdot 2 (target_1 - o_1) \cdot (-1)$

$= -(0.4 - 0.60944) = 0.20944$

ii) $\frac{\partial h(z_3)}{\partial z_3} = h(z_3)(1 - h(z_3))$ $\quad h(x) = sigmoid(x)$

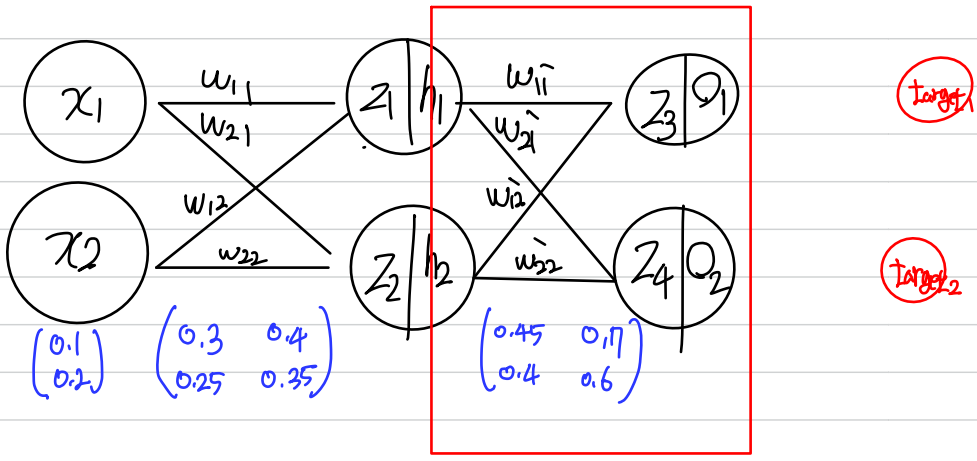
$= 0.60944(1 - 0.60944)$

$= 0.23802$

iii) $z_3 = h_1 \cdot w_{11}' + h_2 \cdot w_{12}'$

$\frac{\partial z_3}{\partial w_{11}'} = h_1 = 0.51998$

$w_{11}' = w_{11}' - \alpha \cdot \frac{\partial E_{total}}{\partial w_{11}'} = 0.45 - (0.1)(0.02592) = 0.447408$



$\begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix}$ $\begin{pmatrix} 0.3 & 0.4 \\ 0.25 & 0.35 \end{pmatrix}$ $\begin{pmatrix} 0.45 & 0.7 \\ 0.4 & 0.6 \end{pmatrix}$

1단계

$\begin{pmatrix} w_{11}' & w_{12}' \\ w_{21}' & w_{22}' \end{pmatrix}$ 갱신완료.

역전파 2단계

ii) $\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$ 를 업데이트,

w_{11} 먼저.

1) $\dfrac{dE_{total}}{dw_{11}} = \underbrace{\dfrac{dE_{total}}{dh_1}}_{i)} \cdot \underbrace{\dfrac{dh_1}{dz_1}}_{ii)} \cdot \underbrace{\dfrac{dz_1}{dw_{11}}}_{iii)}$

항상 이 꼴로 시작.
(원하는 가중치)

i) $\dfrac{dE_{total}}{dh_1} = \dfrac{dE_{10}}{dh_1} + \dfrac{dE_{20}}{dh_1}$ → 마찬가지로 $\dfrac{dE_{10}}{dh_1}$ 반복

$= \dfrac{dE_{10}}{dO_1} \cdot \dfrac{dO_1}{dz_3} \cdot \dfrac{dz_3}{\cancel{dw_{11}'}\, dh_1} \cdot \dfrac{\cancel{dw_{11}'}}{\cancel{dh_1}}$ → 상쇄가

$= \dfrac{d\frac{1}{2}(target_1 - O_1)^2}{dO_1} \cdot \dfrac{dh(z_3)}{dz_3} \cdot \dfrac{d(h_1 \cdot w_{11}' + h_2 \cdot w_{12}')}{d\cancel{w_{11}'}\, h_1}$

$= -(target_1 - O_1) \cdot h(z_3)(1 - h(z_3)) \cdot w_{11}'$

$= (0.20944) \cdot (0.23802) \cdot (0.45)$

$= 0.022433$

$\therefore \dfrac{dE_{total}}{dw_{11}} = \left(0.02243 + \dfrac{dE_2}{dh_1}\right)(0.2496)(0.1)$

$= 0.000808$

$\therefore w_{11} = w_{11} - \alpha \dfrac{dE_{total}}{dw_{11}}$

$= 0.3 - (0.1)(0.000808)$

$= 0.2999192$

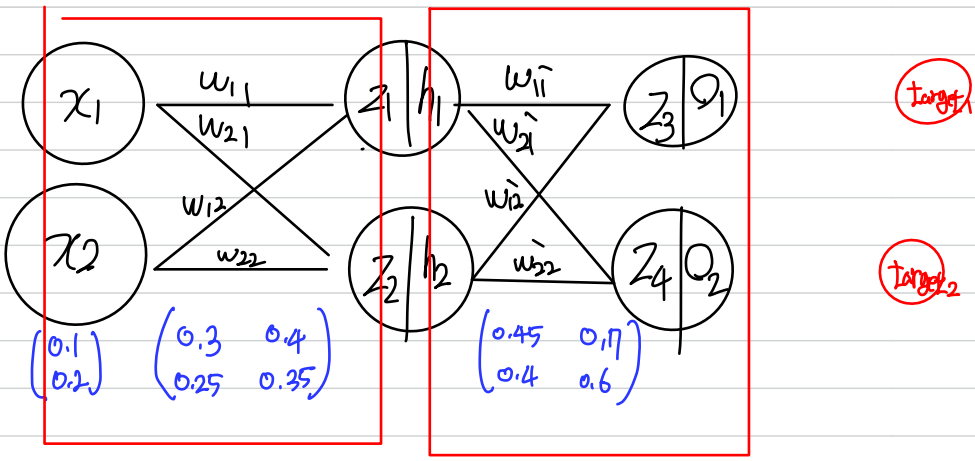
$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}^{old} \longrightarrow \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}^{new}$

$\begin{pmatrix} 0.3 & 0.4 \\ 0.25 & 0.35 \end{pmatrix}$ $\begin{pmatrix} 0.299919 & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$

ii) $\dfrac{dh_1}{dz_1} = \dfrac{d\,h(z_1)}{d(z_1)} = h(z_1)(1 - h(z_1)) = (0.51998)(1 - 0.51998)$

$= 0.249600$

iii) $\dfrac{dz_1}{dw_{11}} = \dfrac{d(x_1 \cdot w_{11} + x_2 \cdot w_{12})}{dw_{11}} = x_1 = 0.1$



2단계 update

$$\begin{pmatrix}0.29959 & 0.249191\\0.39964 & 0.34928\end{pmatrix}$$

1단계 update

$$\begin{pmatrix}w_{11}' & w_{12}'\\w_{21}' & w_{22}'\end{pmatrix}$$ 갱신완료.

$$\begin{pmatrix}0.437038 & 0.38685\\0.69629 & 0.59624\end{pmatrix}$$

init weight

$$\begin{pmatrix}0.3 & 0.4\\0.25 & 0.35\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}z_1\\z_2\end{pmatrix}$$

$$h\begin{pmatrix}z_1\\z_2\end{pmatrix}=\begin{matrix}h_1\\h_2\end{matrix}\quad\begin{pmatrix}0.45 & 0.7\\0.4 & 0.6\end{pmatrix}\begin{pmatrix}h_1\\h_2\end{pmatrix}=\begin{pmatrix}z_3\\z_4\end{pmatrix}$$

$E_{total}=0.023971$

$$h\begin{pmatrix}z_1\\z_2\end{pmatrix}=\begin{matrix}O_1\\O_2\end{matrix}$$

$\frac{1}{2}(O_1-t_1)^2=E_{01}=0.021933$

$\frac{1}{2}(O_2-t_2)^2=E_{02}=0.02038$

1 epoch

update $w=w-\alpha\cdot\frac{dE}{dw}$

update

그래서 기울기가 소실된다는게
이 항이 거의 0에 가깝게 변하질 않아서
학습이 느리거나 제대로 안된다
는거였거나

second weight

$$\begin{pmatrix}0.29959 & 0.24919\\0.39964 & 0.34928\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}z_1\\z_2\end{pmatrix}$$

$$h\begin{pmatrix}z_1\\z_2\end{pmatrix}=\begin{matrix}h_1\\h_2\end{matrix}\quad\begin{pmatrix}0.437038 & 0.38685\\0.69629 & 0.59624\end{pmatrix}\begin{pmatrix}h_1\\h_2\end{pmatrix}=\begin{pmatrix}z_3\\z_4\end{pmatrix}$$

$$h\begin{pmatrix}z_3\\z_4\end{pmatrix}=\begin{matrix}O_1\\O_2\end{matrix}$$

$\frac{1}{2}(O_1-t_1)^2=E_{01}=0.021254$

$\frac{1}{2}(O_2-t_2)^2=E_{02}=0.001981$

$E_{total}=0.023236$