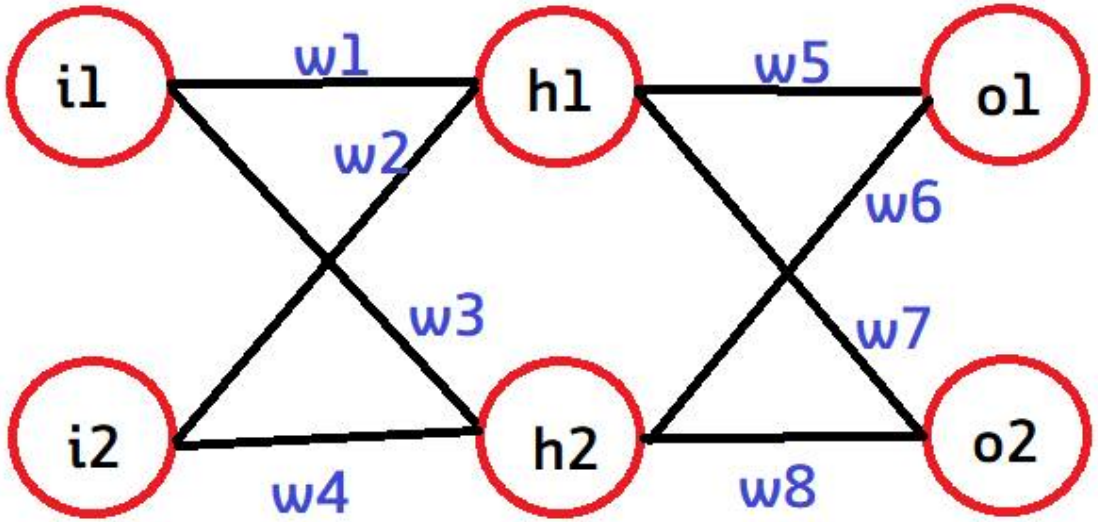
역전파 (Back propagation)



활성화 함수 통과 전 결과값: net / 활성화 함수 통과 후 결과값: out

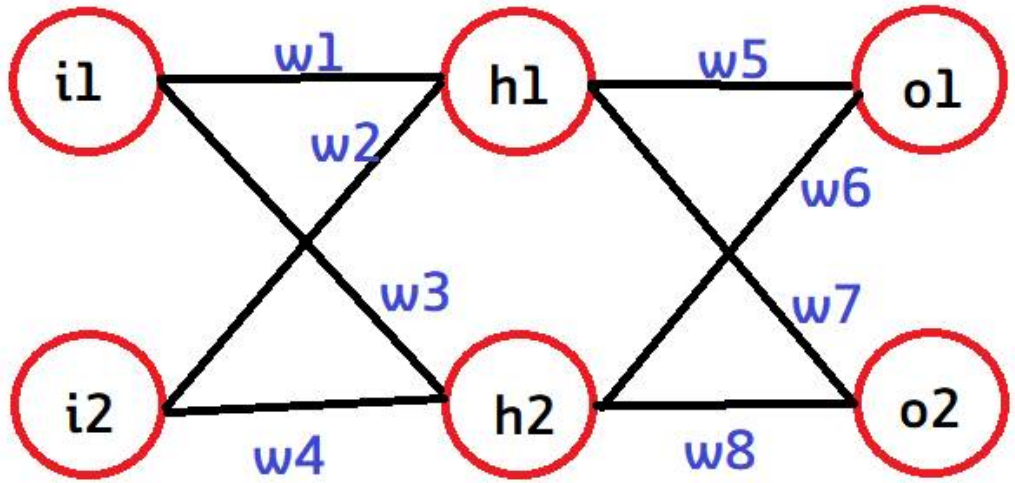
$$net_{o1} = w5 \bullet out_{h1} + w6 \bullet out_{h2}$$

손실함수 $$E = \frac{1}{2}(target_{o1} - out_{o1})^2 + \frac{1}{2}(target_{o2} - out_{o2})^2$$

Chain Rule

$$w_5^+ = w_5 - \gamma \frac{\partial E}{\partial w_5}$$

변화 $w5 \rightarrow net_{o1} \rightarrow out_{o1} \rightarrow E$ 영향

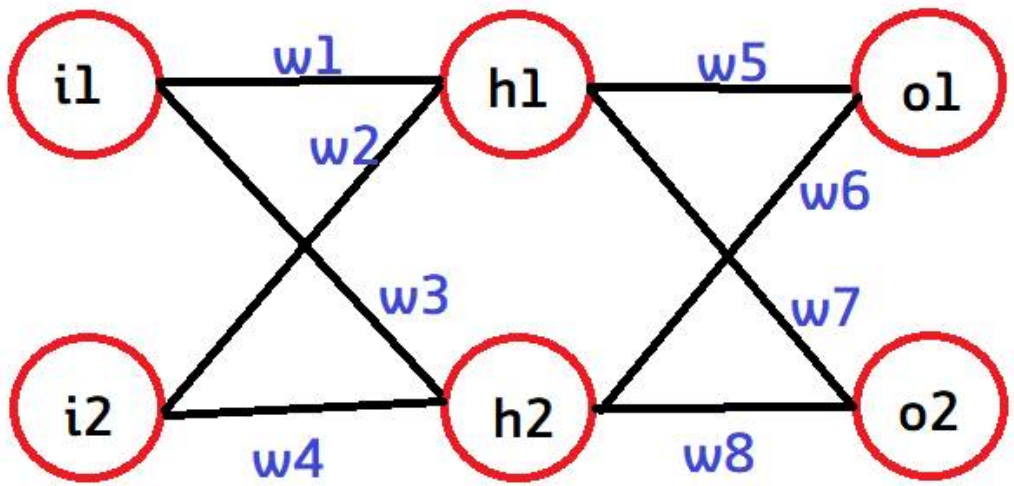


$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial out_{o1}} \frac{\partial out_{o1}}{\partial net_{o1}} \frac{\partial net_{o1}}{\partial w_5}$$

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial out_{o1}} \frac{\partial out_{o1}}{\partial net_{o1}} \frac{\partial net_{o1}}{\partial w_5}$$

$$E = \frac{1}{2}(target_{o1} - out_{o1})^2 + \frac{1}{2}(target_{o2} - out_{o2})^2$$

$$\Rightarrow \frac{\partial E}{\partial out_{o1}} = -(target_{o1} - out_{o1})$$



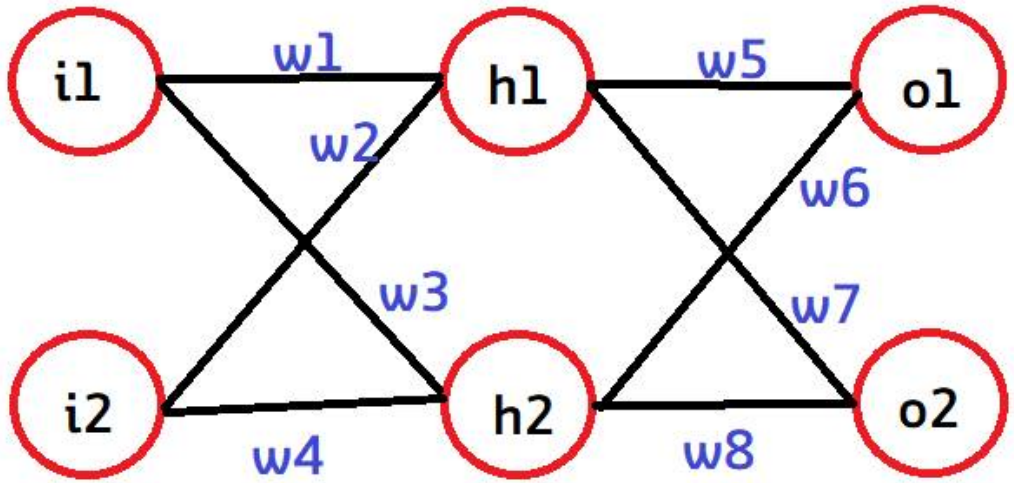
$$out_{o1} = \frac{1}{1 + e^{-net_{o1}}} \quad \text{(sigmoid)}$$

$$\Rightarrow \frac{\partial out_{o1}}{\partial net_{o1}} = out_{o1}(1 - out_{o1})$$

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial out_{o1}} \frac{\partial out_{o1}}{\partial net_{o1}} \frac{\partial net_{o1}}{\partial w_5}$$

$$net_{o1} = w5 \bullet out_{h1} + w6 \bullet out_{h2}$$

$$\Rightarrow \frac{\partial net_{o1}}{\partial w_5} = out_{h1}$$



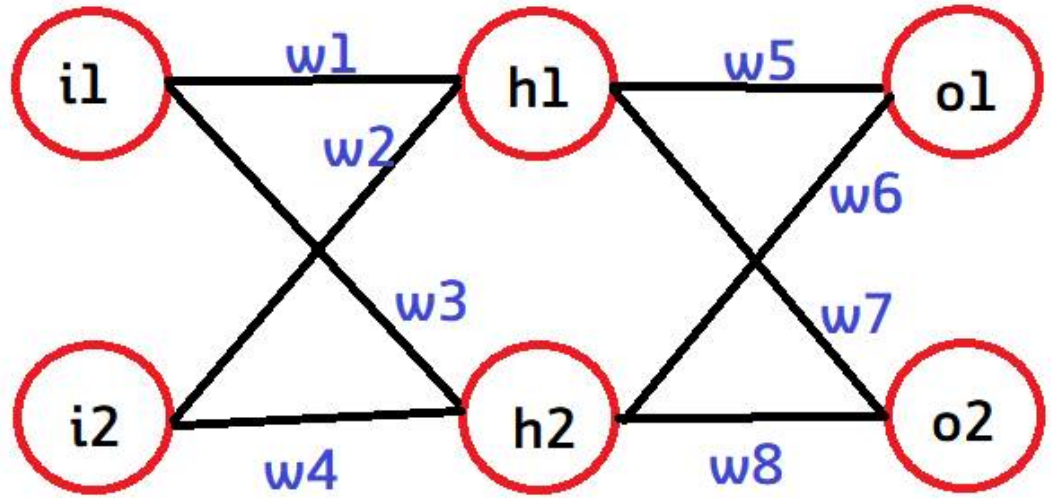
$$w_5^+ = w_5 - \gamma \frac{\partial E}{\partial w_5}$$

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial out_{h1}} \frac{\partial out_{h1}}{\partial net_{h1}} \frac{\partial net_{h1}}{\partial w_1}$$

$$w_1^+ = w_1 - \gamma \frac{\partial E}{\partial w_1}$$

$\frac{\partial out_{h1}}{\partial net_{h1}}$, $\frac{\partial net_{h1}}{\partial w_1}$: 구하기 단순함

$\frac{\partial E}{\partial out_{h1}}$: 구하기 복잡함



$$E = \frac{1}{2}(target_{o1} - out_{o1})^2 + \frac{1}{2}(target_{o2} - out_{o2})^2$$

$$\Rightarrow \frac{\partial E}{\partial out_{h1}} = \frac{\partial E_{o1}}{\partial out_{h1}} + \frac{\partial E_{o2}}{\partial out_{h1}}$$

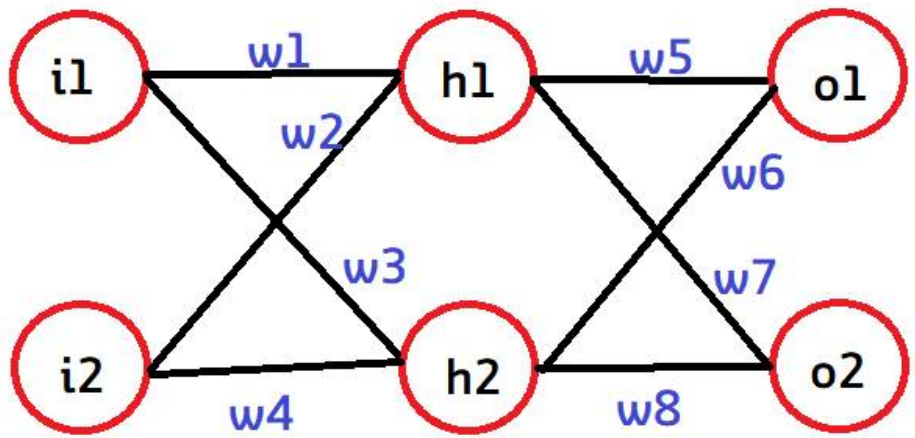
$$\frac{\partial E}{\partial out_{h1}} = \frac{\partial E_{o1}}{\partial out_{h1}} + \frac{\partial E_{o2}}{\partial out_{h1}}$$

$$\frac{\partial E_{o1}}{\partial out_{h1}} = \frac{\partial E_{o1}}{\partial out_{o1}} \frac{\partial out_{o1}}{\partial net_{o1}} \frac{\partial net_{o1}}{\partial out_{h1}} \quad / \quad \frac{\partial E_{o2}}{\partial out_{h1}} = \frac{\partial E_{o2}}{\partial out_{o2}} \frac{\partial out_{o2}}{\partial net_{o2}} \frac{\partial net_{o2}}{\partial out_{h1}}$$

$$\frac{\partial E}{\partial out_{o1}} = -(target_{o1} - out_{o1}) \ , \quad \frac{\partial out_{o1}}{\partial net_{o1}} = out_{o1}(1 - out_{o1})$$

$$net_{o1} = w_5 \bullet out_{h1} + w_6 \bullet out_{h2}$$

$$\Rightarrow \frac{\partial net_{o1}}{\partial out_{h1}} = w_5$$



CNN에서의 chain rule

Input

$$x = \begin{bmatrix} x_{11}\, x_{12}\, x_{13}\, x_{14} \\ x_{21}\, x_{22}\, x_{23}\, x_{24} \\ x_{31}\, x_{32}\, x_{33}\, x_{34} \\ x_{41}\, x_{42}\, x_{43}\, x_{44} \end{bmatrix}, \qquad w = \begin{bmatrix} w_{11}\, w_{12} \\ w_{21}\, w_{22} \end{bmatrix}$$

$$net = \begin{bmatrix} net_{11}\, net_{12}\, net_{13} \\ net_{21}\, net_{22}\, net_{23} \\ net_{31}\, net_{32}\, net_{33} \end{bmatrix} = \begin{bmatrix} w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}\; w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23}\; ... \\ w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32} \qquad ... \qquad ... \\ ... \qquad ... \qquad ... \end{bmatrix}$$

$$net = \begin{bmatrix} net_{11}\, net_{12}\, net_{13} \\ net_{21}\, net_{22}\, net_{23} \\ net_{31}\, net_{32}\, net_{33} \end{bmatrix} = \begin{bmatrix} w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}\; w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23} \ldots \\ w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32} \qquad \ldots \qquad \ldots \\ \ldots \qquad \ldots \qquad \ldots \end{bmatrix}$$

$$\frac{\partial net}{\partial w_{11}} = \begin{bmatrix} \frac{\partial net_{11}}{w_{11}} & \frac{\partial net_{12}}{w_{11}} & \frac{\partial net_{13}}{w_{11}} \\ \frac{\partial net_{21}}{w_{11}} & \frac{\partial net_{22}}{w_{11}} & \frac{\partial net_{23}}{w_{11}} \\ \frac{\partial net_{31}}{w_{11}} & \frac{\partial net_{32}}{w_{11}} & \frac{\partial net_{33}}{w_{11}} \end{bmatrix} = \begin{bmatrix} x_{11}\, x_{12}\, x_{13} \\ x_{21}\, x_{22}\, x_{23} \\ x_{31}\, x_{32}\, x_{33} \end{bmatrix}$$

$$\frac{\partial E}{\partial w_{11}} = \frac{\partial E}{\partial net}\frac{\partial net}{\partial w_{11}} \text{ 이므로} \qquad \frac{\partial E}{\partial w_{11}} = \frac{\partial E}{\partial net_{11}}x_{11} + \ldots + \frac{\partial E}{\partial net_{33}}x_{33}$$

$$w_{11}^{+} = w_{11} - \gamma\frac{\partial E}{\partial w_{11}}$$

참고자료

- 박재우 교수님(연세대학교), *Backpropagation*, (딥러닝 수업자료, 2021)
- 박재우 교수님(연세대학교), *Backpropagation for Convolutional Neural Networks*, (딥러닝 수업자료, 2021)
- 박해선, *혼자 공부하는 머신러닝 + 딥러닝* (한빛미디어, 2021), 423-441