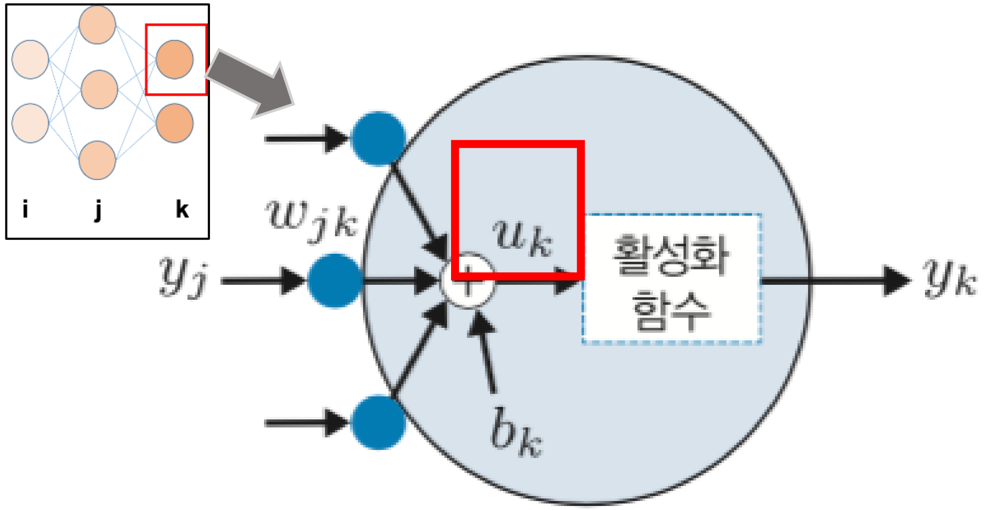
Machine Learning

오류역전파

출력층 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n

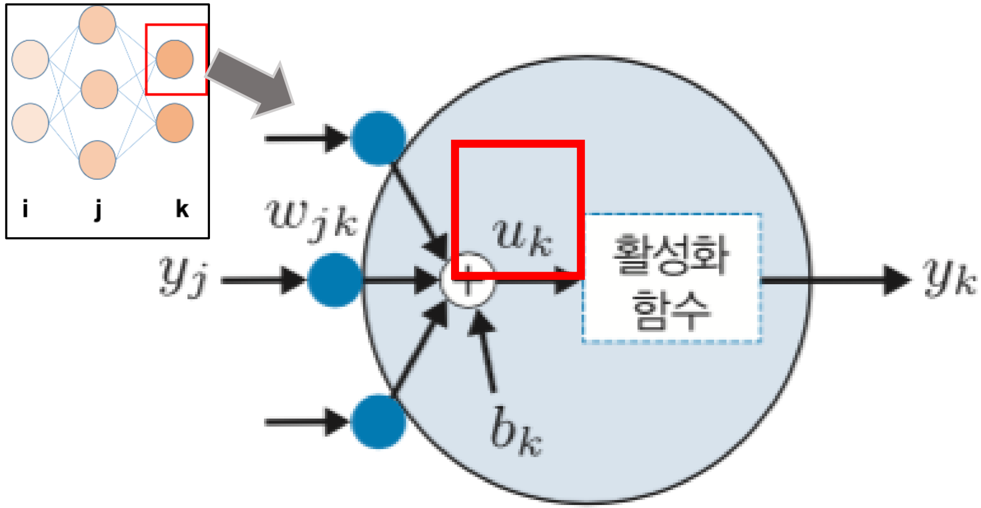


$$u_k = y_1 w_{1k} + y_2 w_{2k} + \dots + y_m w_{mk}$$

$$= \sum_{q=1}^{m} y_q w_{qk} + b_k$$

출력층 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n



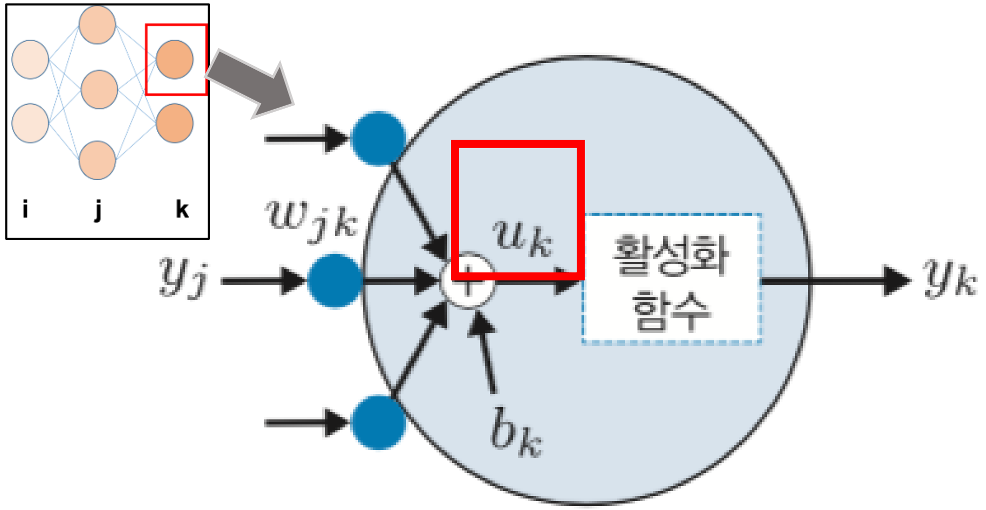
$$u_k = y_1w_{1k}+y_2w_{2k}+\ldots+y_mw_{mk}$$
$$= \sum_{q=1}^{m} y_q w_{qk} + b_k$$

$$\partial w_{jk} = \frac{\partial E}{\partial w_{jk}}$$

$$\partial w_{jk} = \frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial u_k}\frac{\partial u_k}{\partial w_{jk}}$$

출력층 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n



$$\partial w_{jk} = \frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial u_k}\boxed{\frac{\partial u_k}{\partial w_{jk}}}$$

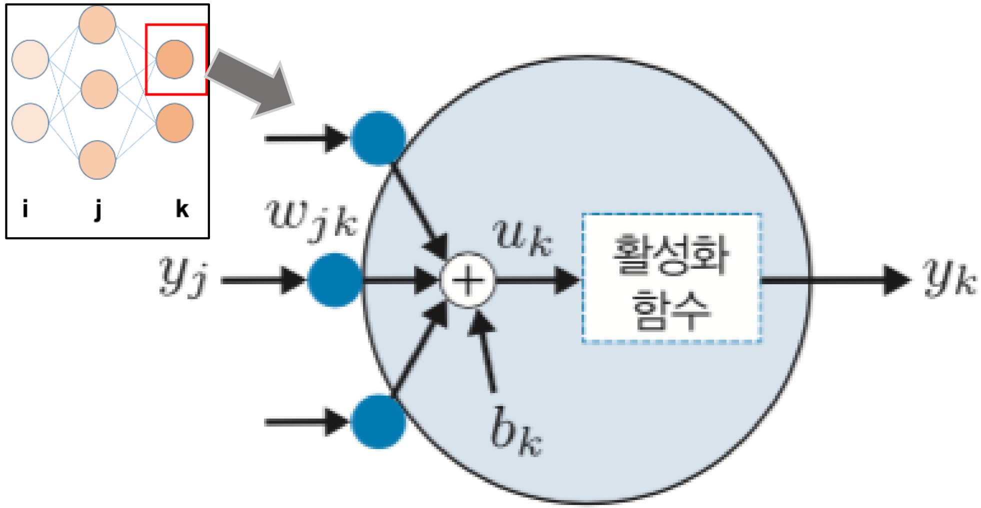
$$\frac{\partial u_k}{\partial w_{jk}} = \frac{\partial(\sum_{q=1}^{m} y_q w_{qk} + b_k)}{\partial w_{jk}}$$

$$= \frac{\partial}{\partial w_{jk}}(y_1 w_{1k} + \ldots + y_j w_{jk} + \ldots + y_m w_{mk})$$

$$= y_j$$

출력층 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n



$$\partial w_{jk} = \frac{\partial E}{\partial w_{jk}} = \boxed{\frac{\partial E}{\partial u_k}} \frac{\partial u_k}{\partial w_{jk}}$$

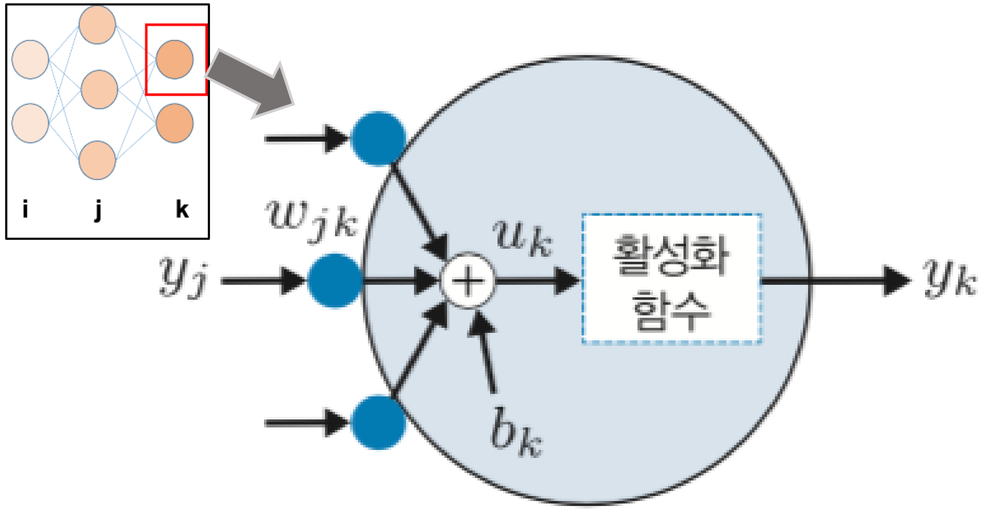
$$\frac{\partial u_k}{\partial w_{jk}} = y_j$$

$$\delta_k = \frac{\partial E}{\partial u_k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial u_k}$$

$$\partial w_{jk} = y_j \delta_k$$

출력층 기울기: bias

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n



$$\partial w_{jk} = \frac{\partial E}{\partial w_{jk}} = \boxed{\frac{\partial E}{\partial u_k}} \frac{\partial u_k}{\partial w_{jk}}$$

$$\frac{\partial u_k}{\partial w_{jk}} = y_j \qquad \partial w_{jk} = y_j \delta_k$$

$$\partial b_k = \frac{\partial E}{\partial b_k} = \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial b_k}$$

$$\frac{\partial u_k}{\partial b_k} = \frac{\partial\left(\sum_{q=1}^{m} y_q w_{qk} + b_k\right)}{\partial b_k}$$

$$= \frac{\partial}{\partial b_k}\left(y_1 w_{1k} + \ldots + y_m w_{mk} + b_k\right)$$

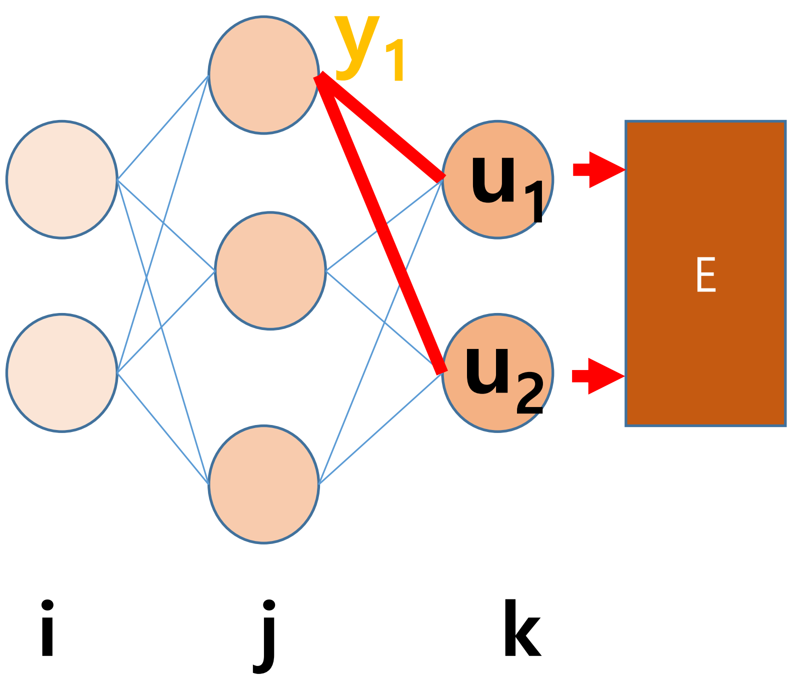
$$= 1$$

$$\partial b_k = \frac{\partial E}{\partial b_k} = \frac{\partial E}{\partial u_k} \frac{\partial u_k}{\partial b_k} = \delta_k * 1 = \delta_k$$

출력층의 입력값 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n

$$\partial y_j = \frac{\partial E}{\partial y_j}$$



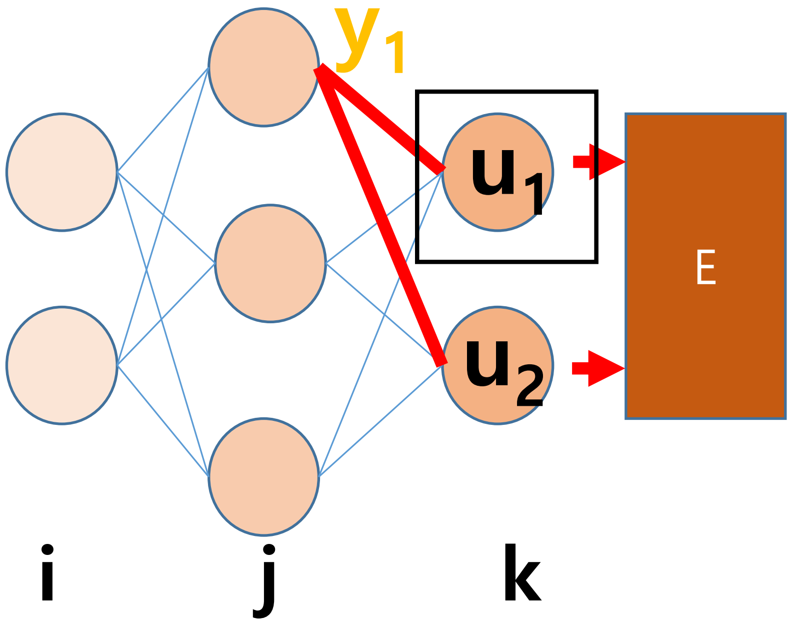
$$\partial y_j = \frac{\partial E}{\partial y_j}$$

$$= \sum_{r=1}^{n} \frac{\partial E}{\partial u_r} \frac{\partial u_r}{\partial y_j}$$

출력층의 입력값 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n

$$\partial y_j = \frac{\partial E}{\partial y_j}$$



$$\partial y_j = \sum_{r=1}^{n} \frac{\partial E}{\partial u_r} \boxed{\frac{\partial u_r}{\partial y_j}}$$

$$\frac{\partial u_r}{\partial y_j} = \frac{\partial\left(\sum_{q=1}^{m} y_q w_{qr} + b_r\right)}{\partial y_j}$$

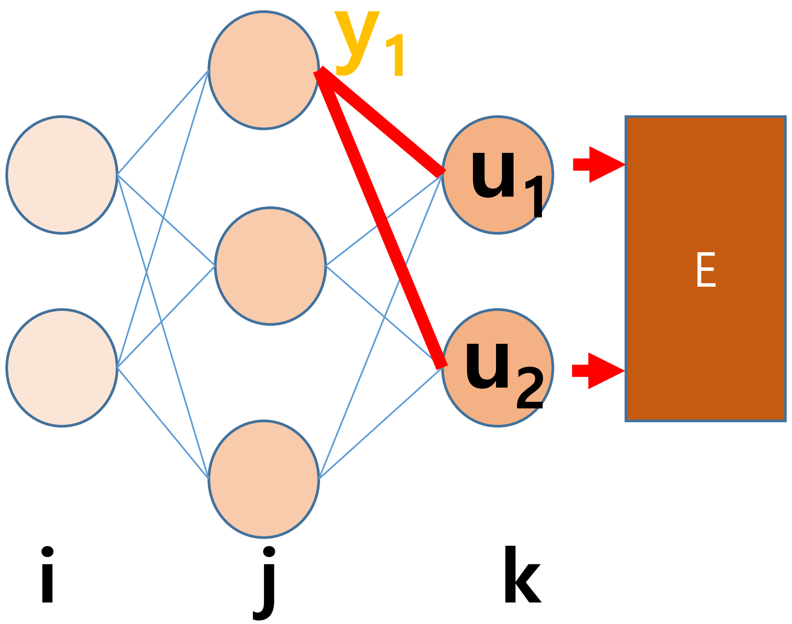
$$= \frac{\partial}{\partial y_j}\left(y_1 w_{1r} + \ldots + y_j w_{jr} + \ldots + y_m w_{mr} + b_r\right)$$

$$= w_{jr}$$

출력층의 입력값 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n

$$\partial y_j = \frac{\partial E}{\partial y_j}$$



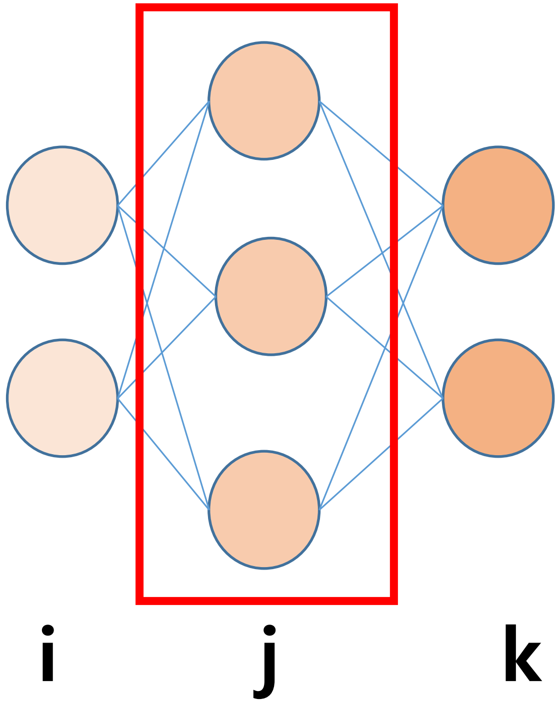
$$\partial y_j = \sum_{r=1}^{n} \frac{\partial E}{\partial u_r} \boxed{\frac{\partial u_r}{\partial y_j}}$$

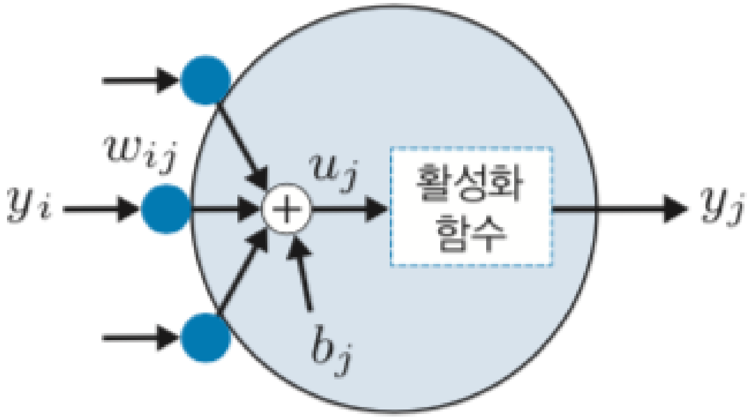
$$\frac{\partial u_r}{\partial y_j} = w_{jr}$$

$$\partial y_j = \sum_{r=1}^{n} \delta_r w_{jr}$$

은닉층의 가중치 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n





$$\partial w_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial u_j} \boxed{\frac{\partial u_j}{\partial w_{ij}}}$$

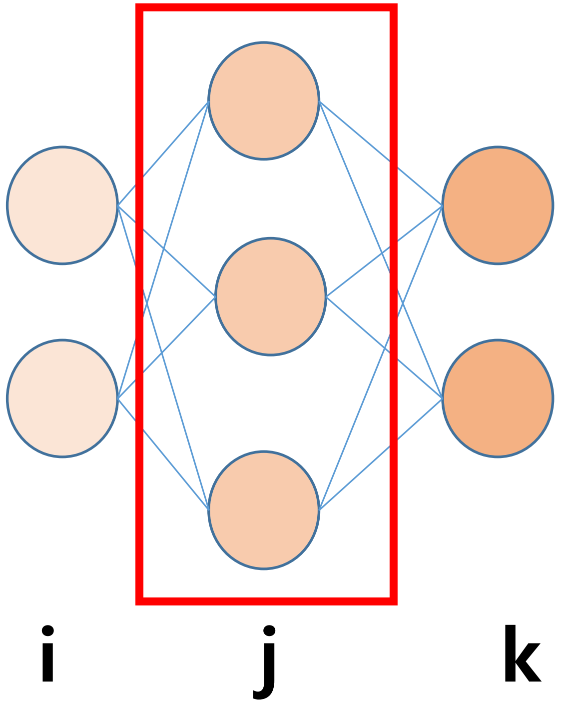
$$\frac{\partial u_j}{\partial w_{ij}} = \frac{\partial\left(\sum_{p=1}^{l} y_p w_{pj} + b_j\right)}{\partial w_{ij}}$$

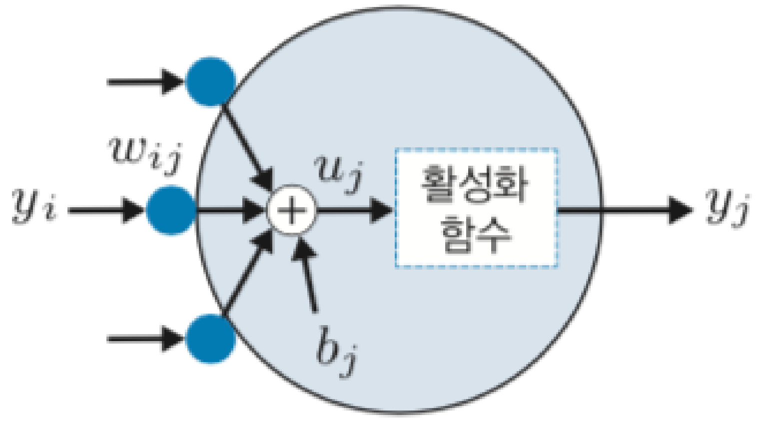
$$= \frac{\partial}{\partial w_{ij}}\left(y_1 w_{1j} + \ldots + y_i w_{ij} + \ldots + y_l w_{lj} + b_j\right)$$

$$= y_i$$

은닉층의 가중치 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n



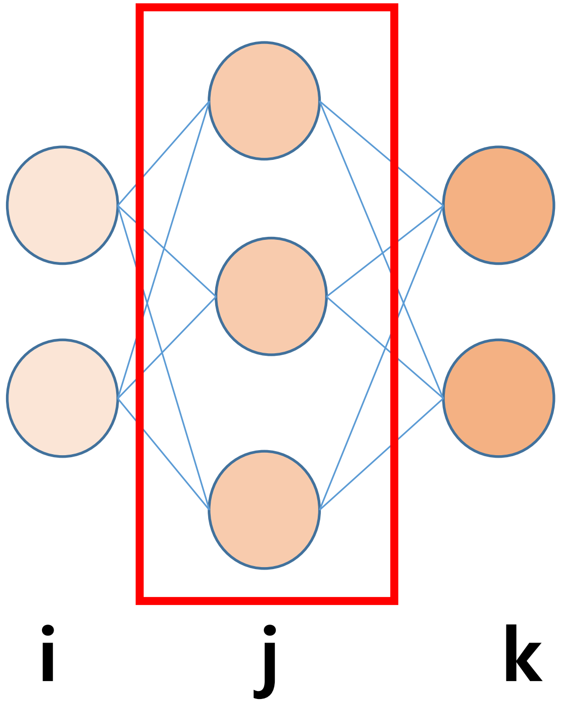


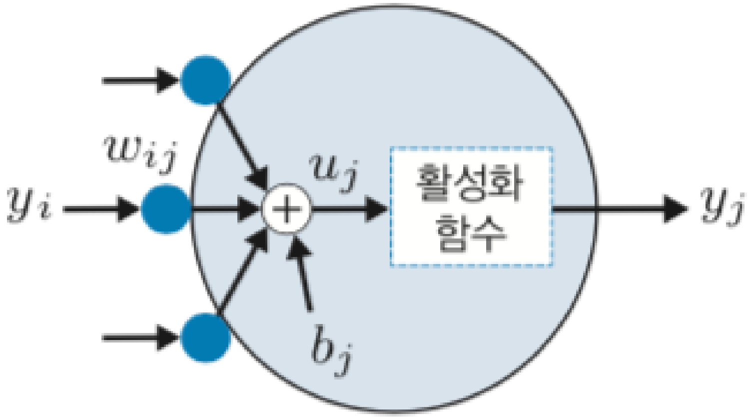
$$\partial w_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial u_j} \boxed{\frac{\partial u_j}{\partial w_{ij}}}$$

$$\frac{\partial u_j}{\partial w_{ij}} = y_i$$

은닉층의 가중치 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n





$$\partial w_{ij} = \frac{\partial E}{\partial w_{ij}} = \boxed{\frac{\partial E}{\partial u_j}} \frac{\partial u_j}{\partial w_{ij}}$$

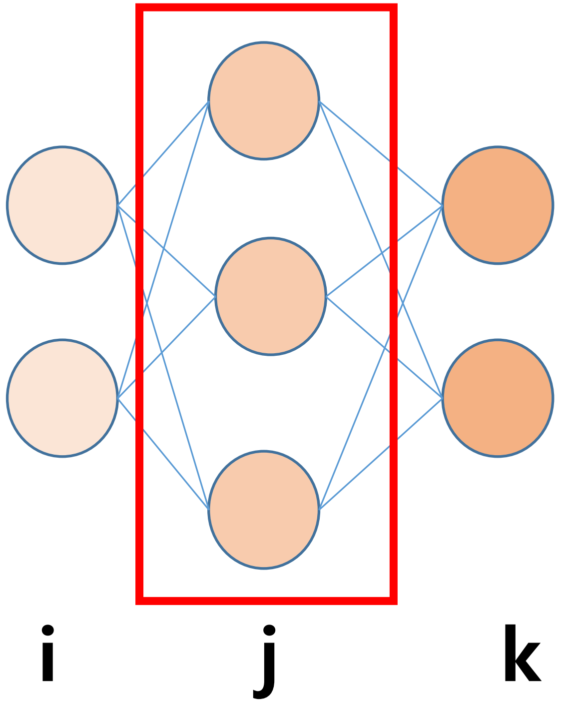
$$\frac{\partial u_j}{\partial w_{ij}} = y_i$$

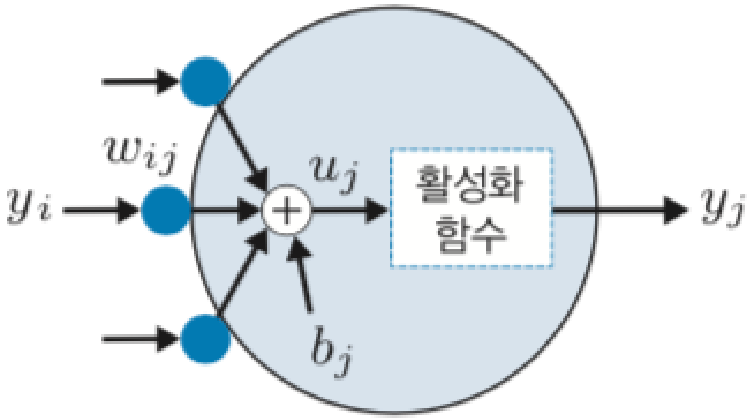
$$\frac{\partial E}{\partial u_j} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial u_j}$$

$$\delta_j = \frac{\partial E}{\partial u_j} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial u_j} = \partial y_j \frac{\partial y_j}{\partial u_j}$$

은닉층의 가중치 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n





$$\partial w_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial u_j}\frac{\partial u_j}{\partial w_{ij}}$$

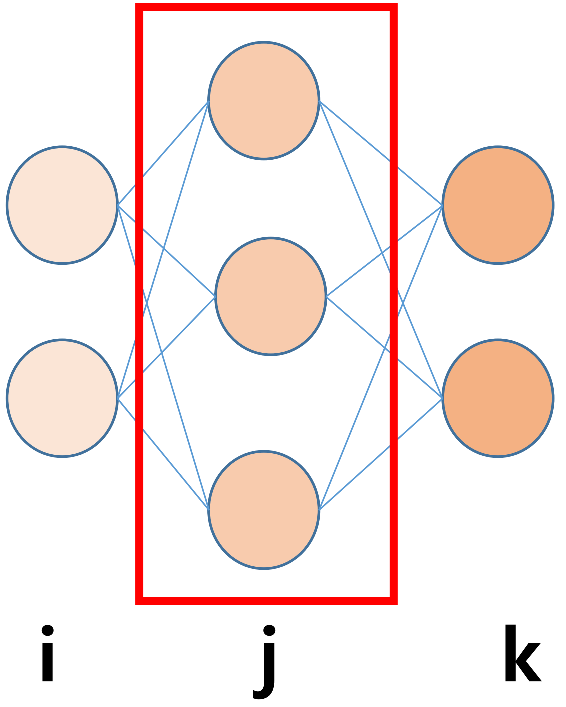
$$\frac{\partial E}{\partial u_j} = \partial y_j \frac{\partial y_j}{\partial u_j} = \delta_j$$

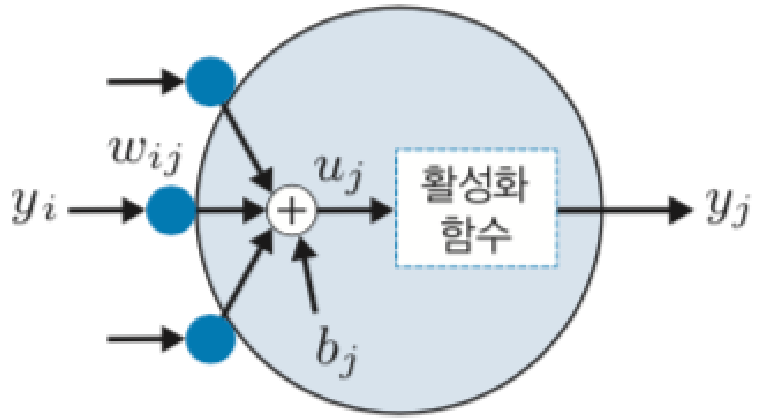
$$\frac{\partial u_j}{\partial w_{ij}} = y_i$$

$$\boxed{\partial w_{ij} = \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial u_j}\frac{\partial u_j}{\partial w_{ij}} = y_i\delta_j}$$

은닉층의 bias 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n





$$\partial b_j = \frac{\partial E}{\partial b_j} = \frac{\partial E}{\partial u_j}\frac{\partial u_j}{\partial b_j}$$

$$\frac{\partial u_j}{\partial b_j} = \frac{\partial(\sum_{p=1}^{l} y_p w_{pj} + b_j)}{\partial b_j}$$

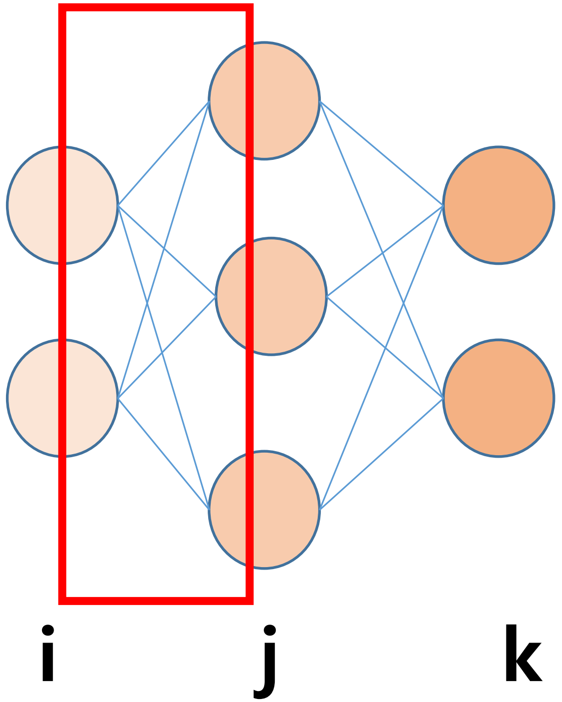
$$= \frac{\partial}{\partial b_j}(y_1 w_{1j} + \ldots + y_i w_{ij} + b_j)$$

$$= 1$$

$$\boxed{\partial b_j = \delta_j}$$

은닉층 앞층의 입력 기울기

layer	첨자	노드 수
Input	i	l
Hidden	j	m
Output	k	n



$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$

기울기 요약: 출력층

$$\partial w_{jk} = \frac{\partial E}{\partial w_{jk}}$$

$$\partial b_k = \frac{\partial E}{\partial b_k}$$

$$\partial y_j = \frac{\partial E}{\partial y_j}$$

$$\delta_k = \frac{\partial E}{\partial u_k} = \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial u_k}$$ 식 5-15

$$\partial w_{jk} = y_j \delta_k$$ 식 5-16

$$\partial b_k = \delta_k$$ 식 5-17

$$\partial y_j = \sum_{r=1}^{n} \delta_r w_{jr}$$ 식 5-18

기울기 요약: 은닉층

$$\partial w_{ij} = \frac{\partial E}{\partial w_{ij}}$$

$$\partial b_j = \frac{\partial E}{\partial b_j}$$

$$\partial y_i = \frac{\partial E}{\partial y_i}$$

$$\delta_j = \frac{\partial E}{\partial u_j} = \partial y_j \frac{\partial y_j}{\partial u_j}$$ 식 5-19

$$\partial w_{ij} = y_i \delta_j$$ 식 5-20

$$\partial b_j = \delta_j$$ 식 5-21

$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$ 식 5-22

회귀 문제 기울기 구하기

- **Loss Function: 오차 제곱합**

- **은닉층 활성화 함수: Sigmoid Function**

- **출력층 활성화 함수: 항등 함수**

회귀 문제 기울기 구하기: 오차 제곱합

$$\delta_k = \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial u_k}$$

$$\begin{aligned}
\frac{\partial E}{\partial y_k} &= \frac{\partial}{\partial y_k}(\frac{1}{2}\sum_k (y_k - t_k)^2) \\
&= \frac{\partial}{\partial y_k}(\frac{1}{2}(y_0 - t_0)^2 + \frac{1}{2}(y_1 - t_1)^2 + \cdots + \frac{1}{2}(y_k - t_k)^2 + \cdots + \frac{1}{2}(y_n - t_n)^2) \\
&= y_k - t_k
\end{aligned}$$

식 5-24

회귀 문제 기울기 구하기: 출력층

(오차제곱합, 항등 함수)

$$\delta_k = \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial u_k}$$

$$\begin{aligned}
\frac{\partial E}{\partial y_k} &= \frac{\partial}{\partial y_k}\left(\frac{1}{2}\sum_k (y_k - t_k)^2\right) \\
&= \frac{\partial}{\partial y_k}\left(\cancel{\frac{1}{2}(y_0 - t_0)^2} + \cancel{\frac{1}{2}(y_1 - t_1)^2} + \cdots + \frac{1}{2}(y_k - t_k)^2 + \cdots + \cancel{\frac{1}{2}(y_n - t_n)^2}\right) \\
&= y_k - t_k
\end{aligned}$$

식 5-24

$$\frac{\partial y_k}{\partial u_k} = \frac{\partial u_k}{\partial u_k} = 1$$

$$\delta_k = y_k - t_k$$

회귀 문제 기울기 구하기: 출력층

(오차제곱합, 항등 함수)

$$\delta_k = y_k - t_k$$

$$\partial w_{jk} = y_j \delta_k$$

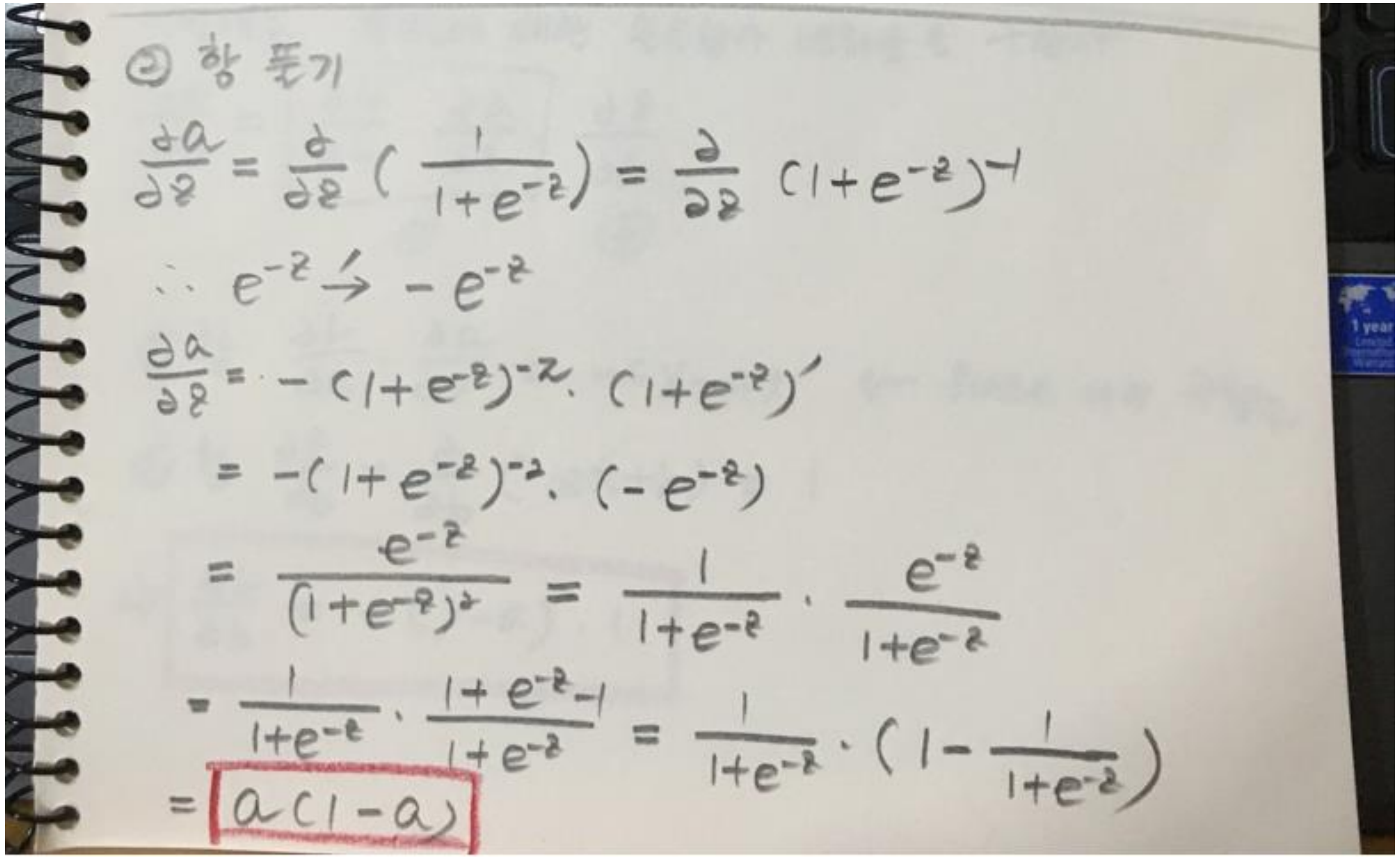
$$\partial b_k = \delta_k$$

$$\partial y_i = \sum_{r=1}^{n} \delta_r w_{jr}$$

회귀 문제 기울기 구하기: 은닉층 (sigmoid)

$$\delta_j = \partial y_j \frac{\partial y_j}{\partial u_j}$$

$$\frac{\partial y_j}{\partial u_j} = \left(1 - \mathrm{y_j}\right)\mathrm{y_j}$$



회귀 문제 기울기 구하기: 은닉층 (sigmoid)

$$\delta_j = \frac{\partial E}{\partial u_j} = \partial y_j \frac{\partial y_j}{\partial u_j}$$

$$\partial w_{ij} = y_i \delta_j$$

$$\partial b_j = \delta_j$$

$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$

$$\delta_j = \partial y_j (1 - \mathrm{y_j}) \mathrm{y_j}$$

$$\partial w_{ij} = y_i \delta_j$$

$$\partial b_j = \delta_j$$

$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$

분류 문제에서 기울기

$$\delta_k = \frac{\partial E}{\partial u_k}$$

$$E = -\sum_k t_k \log(y_k)$$

식 5-29

$$y_k = \frac{\exp(u_k)}{\sum \exp(u_k)}$$

식 5-30

$$\begin{aligned}
E &= -\sum_k \left(t_k \log(\exp(u_k)) - t_k \log \sum_k \exp(u_k)\right) \\
&= -\sum_k \left(t_k \log(\exp(u_k))\right) + \sum_k (t_k \log \sum_k \exp(u_k)) \\
&= -\sum \left(t_k \log(\exp(u_k))\right) + (\sum t_k)(\log \sum_k \exp(u_k))
\end{aligned}$$

식 5-31

$$E = -\sum_k t_k u_k + \log \sum_k \exp(u_k)$$

분류 문제에서 기울기: 출력층

$$
\begin{aligned}
\delta_k &= \frac{\partial E}{\partial u_k} \\
&= \frac{\partial}{\partial u_k}\left(-\sum_k t_k u_k + \log \sum_k \exp(u_k)\right) \\
&= -t_k + \frac{\exp(u_k)}{\sum_k \exp(u_k)} \\
&= -t_k + y_k \\
&= y_k - t_k
\end{aligned}
$$

분류 문제에서 기울기: 출력층

$$\delta_k = y_k - t_k$$

$$\partial w_{jk} = y_j \delta_k$$

$$\partial b_k = \delta_k$$

$$\partial y_i = \sum_{r=1}^{n} \delta_r w_{jr}$$

```python
# -- 출력층 --
class OutputLayer(BaseLayer):
    def __init__(self, n_upper, n):
        self.w = np.random.randn(n_upper, n) / np.sqrt(n_upper)  # 자비에르 초기화 기반의 초깃값
        self.b = np.zeros(n)

    def forward(self, x):
        self.x = x
        u = np.dot(x, self.w) + self.b
        self.y = np.exp(u)/np.sum(np.exp(u), axis=1, keepdims=True)  # 소프트맥스 함수

    def backward(self, t):
        delta = self.y - t

        self.grad_w = np.dot(self.x.T, delta)
        self.grad_b = np.sum(delta, axis=0)
        self.grad_x = np.dot(delta, self.w.T)
```

분류 문제에서 기울기: 은닉층

$$\delta_j = \frac{\partial E}{\partial u_j} = \partial y_j \frac{\partial y_j}{\partial u_j}$$

$$\partial w_{ij} = y_i \delta_j$$

$$\partial b_j = \delta_j$$

$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$

$$\delta_j = \partial y_j (1 - \mathrm{y_j}) \mathrm{y_j}$$

$$\partial w_{ij} = y_i \delta_j$$

$$\partial b_j = \delta_j$$

$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$

참고문헌

- [도서명] 실체가 손에 잡히는 딥러닝
- [저자] 아즈마 유키나가
- [역자] 최재원
- [출판사] 책만