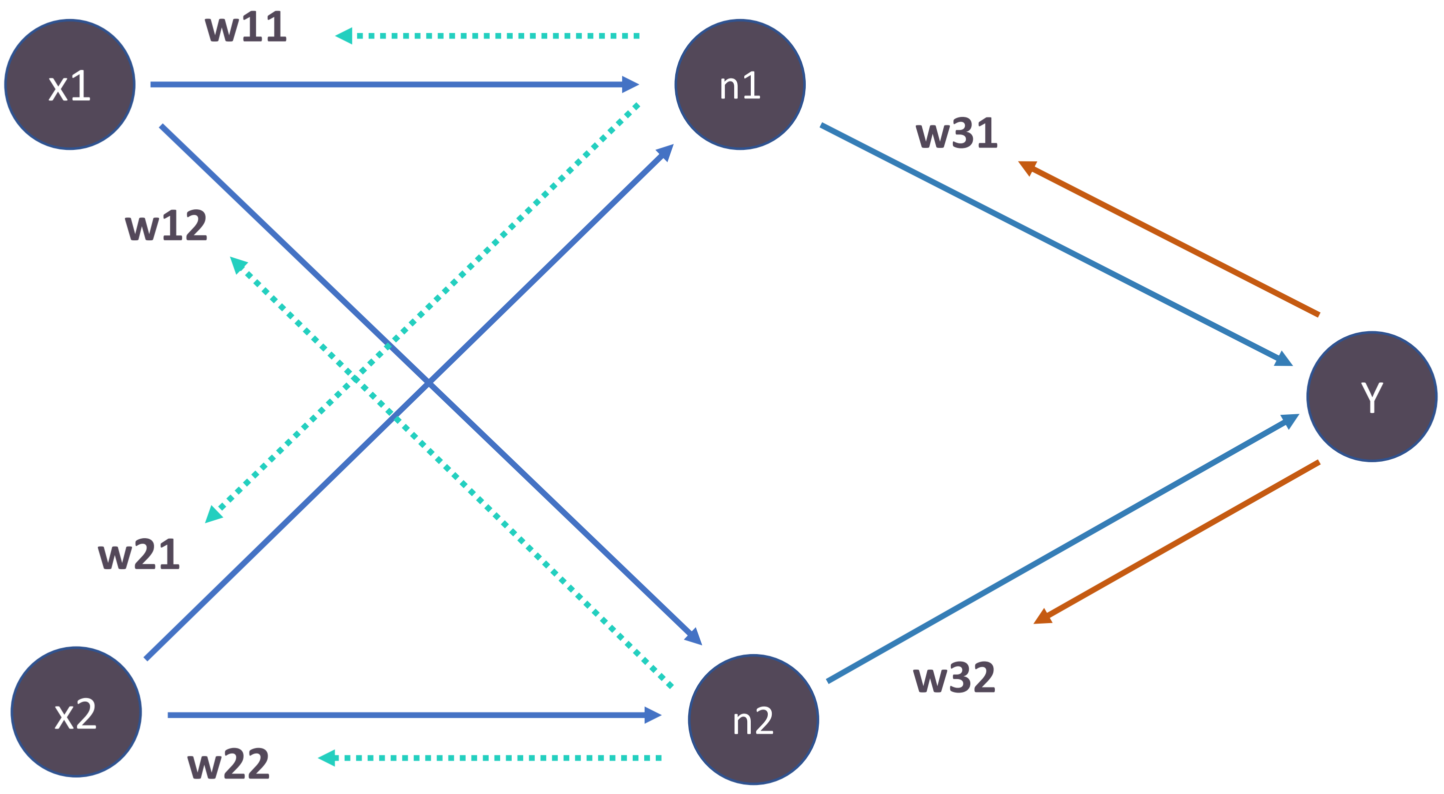
Deep Learning Study

오차 역전파

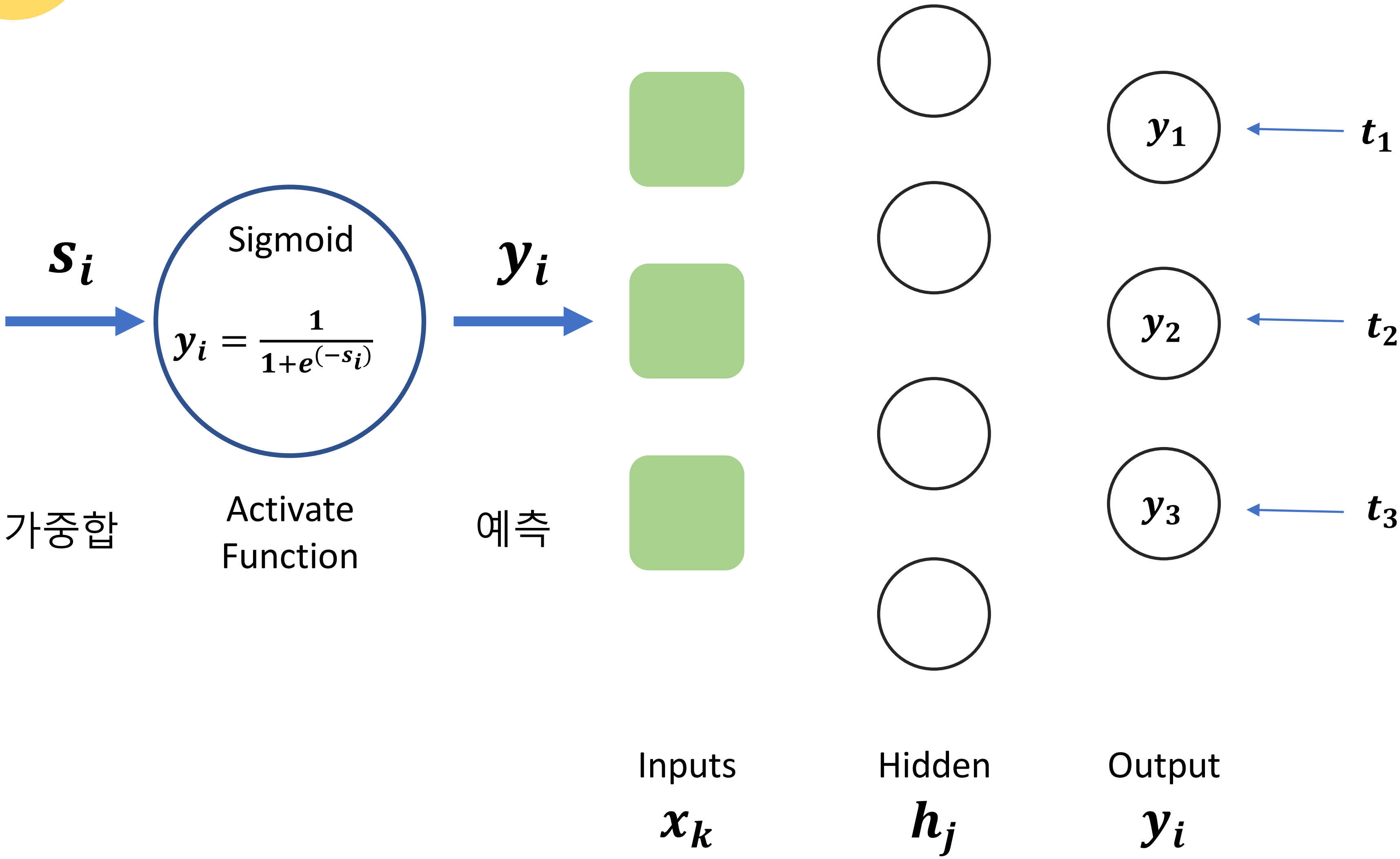
KAIG

INDEX

1 오차역전파



1 오차역전파



s_i
Sigmoid
$y_i = \frac{1}{1+e^{(-s_i)}}$
y_i
가중합
Activate Function
예측
y_1
y_2
y_3
t_1
t_2
t_3
Inputs
x_k
Hidden
h_j
Output
y_i

1 오차역전파

Error function, E

$$\mathbf{E} = -\sum_{i=1}^{nout}\{t_i \log(y_i) + (1 - t_i)\log(1 - y_i)\}$$

$where, y_i = \frac{1}{1+e^{(-s_i)}}, s_i = \sum_{j=1} w_{ji} h_j$

$ex)\ \ t_i = 1 \text{ and } x_i = 0 \sim E = \ \infty$

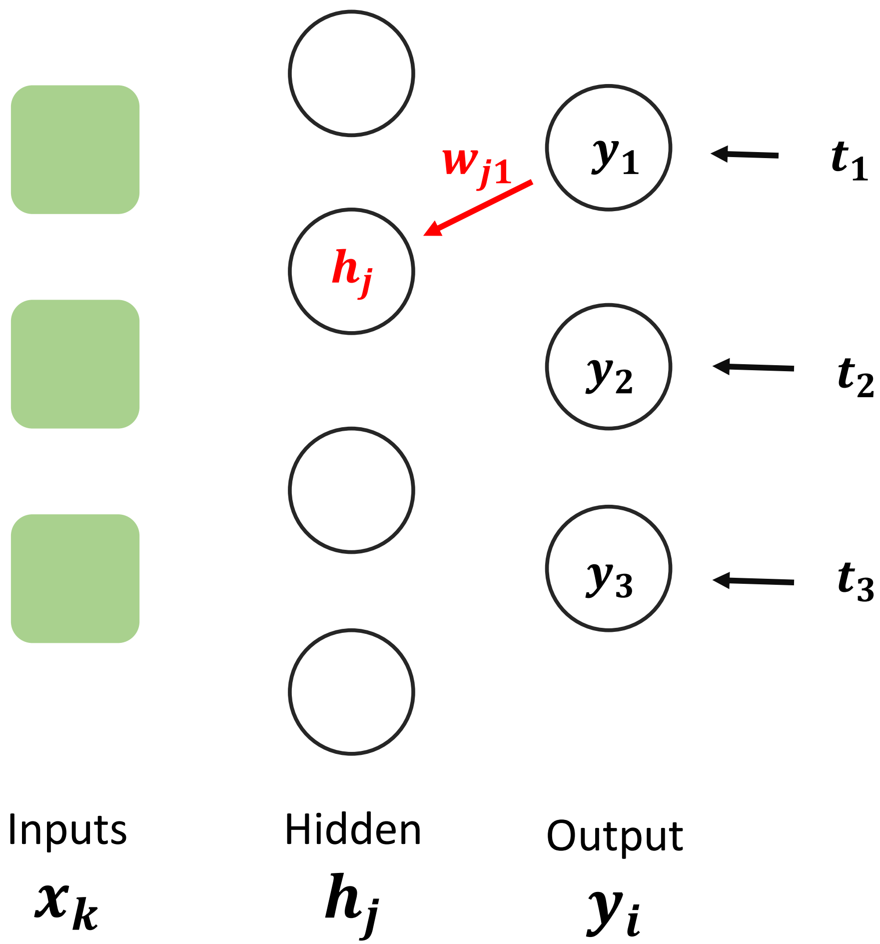
$ex)\ \ t_i = 1 \text{ and } x_i = 1 \sim E = 0$

t_i: 실제 클래스의 값 / x_i: 예상된 값

목표: 오차 E를 최소로 줄이는 것

1 오차역전파

Gradient ➔ f의 임의의 점 x 에서 Δx만큼 움직였을 때, 벡터로 표시된 f값이 변화하는 값



Input Data가 아닌,
Weight값을 조정함으로써
목표 t(target)값에 이르게 함

즉,
1.layer의 weight에 대한
  E의 gradient를 계산하고
  (E의 미분)

2. 변화량을 w의 반대 방향으로
   더해준다 (-1을 곱해줌)

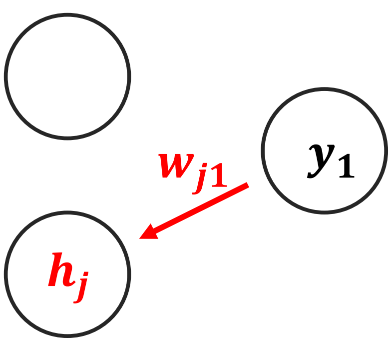
1 오차역전파

신경망 학습 순서를 정리해보자!

1. Input Layer를 지나며 weight를 곱하고
2. Activate function을 지난 다음
3. 마지막 output layer에서 오차 값을 계산하고
4. 업데이트를 위해 순차적으로 거꾸로 돌아간다.

➔ 따라서, 돌아갈 때는 Output -> Hidden 의 weight부터 구해준다.

체인 룰



$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_i}\frac{\partial y_i}{\partial s_i}\frac{\partial s_i}{\partial w_{ji}}$$

i번째 output인 y_i에 대한 오차값에 대해 w_{ji}가 바뀌어야 하는 정도

1 오차역전파

$$\frac{\partial E}{\partial w_{ji}} = \boxed{\frac{\partial E}{\partial y_i}} \frac{\partial y_i}{\partial s_i} \frac{\partial s_i}{\partial w_{ji}}$$

$y_1 \leftarrow t_1$ $E(x_1)$

$y_2 \leftarrow t_2$ $E(x_2)$

$y_3 \leftarrow t_3$ $E(x_3)$

$$\mathbf{E} = -\sum_{i=1}^{nout} \{t_i \log(y_i) + (1 - t_i) \log(1 - y_i)\}$$

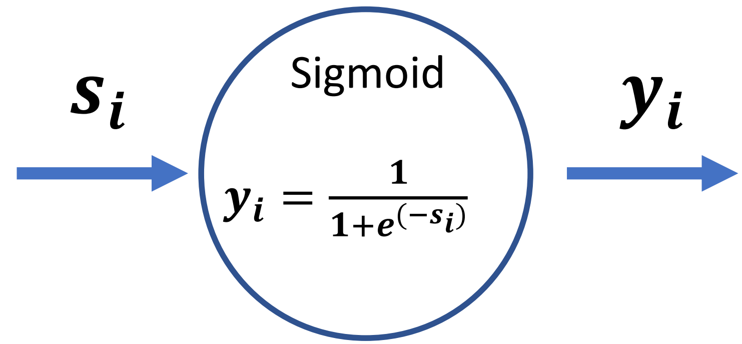
① $\mathrm{E} = \mathrm{E}(y_1) + E(y_2) \cdots$

② $\dfrac{\partial E}{\partial y_i} = \dfrac{\partial E(y_i)}{\partial y_i} = \dfrac{-t_1}{y_i} + \dfrac{1-t_i}{y_i}$

$$= \frac{y_i - t_i}{y_i(1 - y_i)} \cdots\cdots a$$

1 오차역전파

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_i}\boxed{\frac{\partial y_i}{\partial s_i}}\frac{\partial s_i}{\partial w_{ji}}$$



Activate
Function

$$\frac{d}{dx}y_i = \frac{d}{dx}\left[\frac{1}{1+e^{-x}}\right]$$

① $$= \frac{d}{dx}(1+e^{-x})^{-1}$$

$$y_i = \frac{1}{1+e^{(-s_i+b)}},\ s_i = \sum_{j=1} w_{ji}x_j$$

② $$= -(1+e^{-x})^{-2}(-e^{-x})$$

③ $$= \frac{(e^{-x})}{(1+e^{-x})^{-2}} = \frac{(e^{-x})}{1+e^{-x}}\cdot\frac{1}{1+e^{-x}} = \frac{(1+e^{-x})-1}{1+e^{-x}}\cdot\frac{1}{1+e^{-x}}$$

④ $$= \left(1+\frac{-1}{1+e^{-x}}\right)\cdot\frac{1}{1+e^{-x}}$$

$$= y_i\cdot(1-y_i) \quad \cdots\cdots b$$

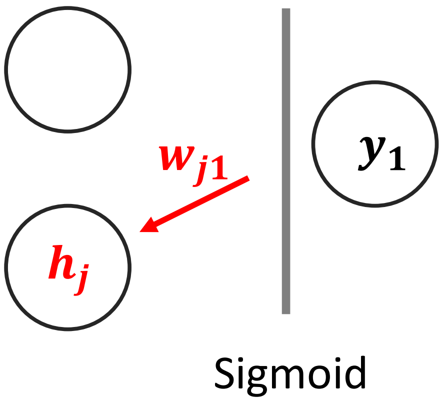
1 오차역전파

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_i}\frac{\partial y_i}{\partial s_i}\boxed{\frac{\partial s_i}{\partial w_{ji}}}$$

$$\boldsymbol{s_i = \sum_{j=1} w_{ji} h_j}$$

$$\boldsymbol{\frac{\partial s_i}{\partial w_{ji}} = \frac{\sum_{j=1} w_{ji} h_j}{\partial w_{ji}}}$$

$$\boldsymbol{= h_j \cdots\cdots c}$$



1 오차역전파

$$\frac{\partial E}{\partial w_{ji}} = \boxed{\frac{\partial E}{\partial y_i}\frac{\partial y_i}{\partial s_i}\frac{\partial s_i}{\partial w_{ji}}} = a \cdot b \cdot c$$

$$y_i = \frac{1}{1+e^{(-s_i)}}$$

$$s_i = \sum_{j=1} w_{ji} h_j$$

$$\frac{\partial E}{\partial y_i} = \frac{y_i - t_i}{y_i(1-y_i)} \cdots\cdots\cdots a$$

$$\frac{\partial y_i}{\partial s_i} = y_i \cdot (1 - y_i) \cdots\cdots b$$

$$\frac{\partial s_i}{\partial w_{ji}} = h_j \cdots\cdots\cdots\cdots\cdots\cdots c$$

$$\frac{\partial E}{\partial w_{ji}} = (y_i - t_i) h_j$$

출력층의 가중치 업데이트

1 오차역전파

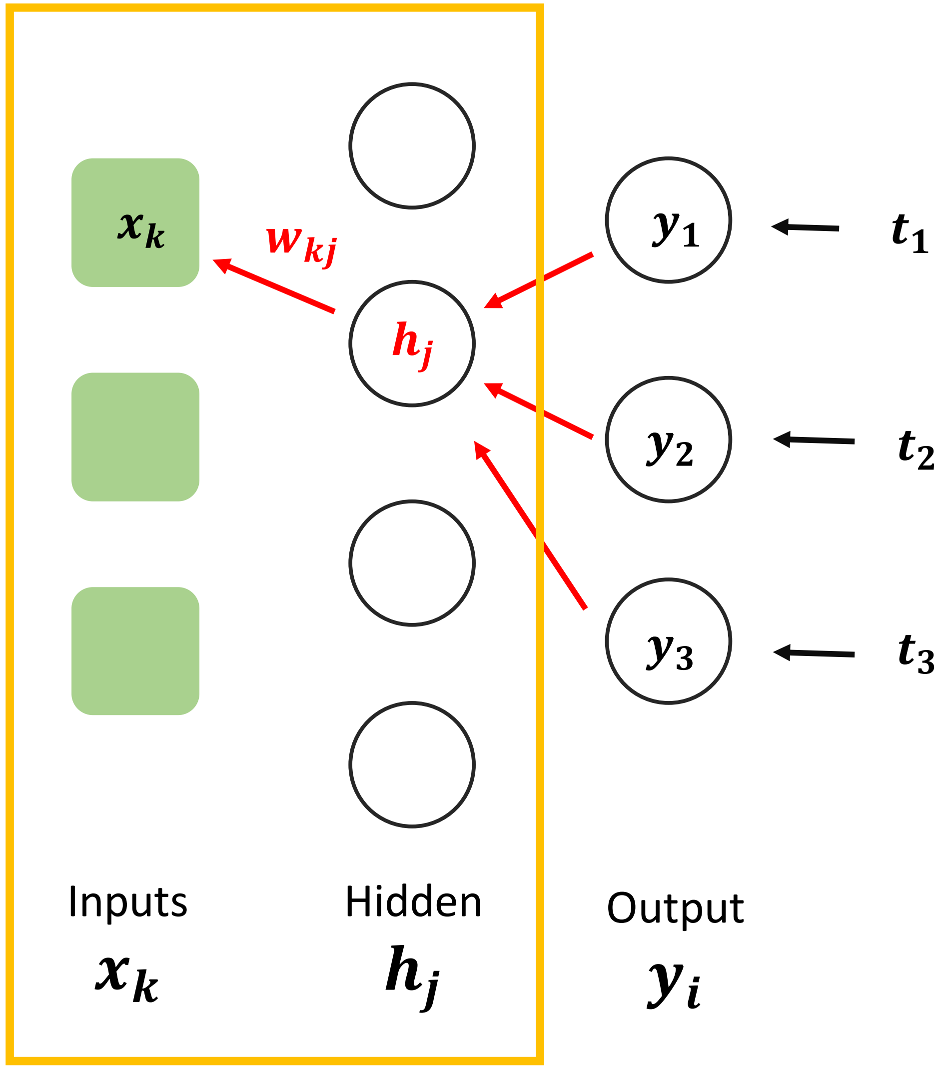
$$\frac{\partial E}{\partial w_{ji}} = (y_i - t_i)h_j$$

$$w_{ji}(T+1) = w_{ji}(T) - \boxed{(y_i - t_i)}h_j$$

δy(델타식y)

출력층의 가중치 업데이트

1 오차역전파



1 오차역전파

$$\frac{\partial E}{\partial w_{kj}} = \boxed{\frac{\partial E}{\partial s_j}\frac{\partial s_j}{\partial w_{kj}}}$$

$$\boxed{① \ \frac{\partial E}{\partial s_j}} = \sum_{i=1}^{nout} \frac{\partial E}{\partial s_i}\frac{\partial s_i}{\partial h_j}\frac{\partial h_j}{\partial s_j} = \sum_{i=1}^{nout} (y_i - t_i)(w_{ji})(h_j(1 - h_j))$$

$$\frac{\partial E}{\partial y_i} = \frac{y_i - t_i}{y_i(1 - y_i)} \quad \cdots\cdots\cdots a$$

$$\frac{\partial y_i}{\partial s_i} = y_i \cdot (1 - y_i) \quad \cdots\cdots b$$

$$(y_i - t_i)$$

은닉층의 가중치 업데이트

1 오차역전파

$$\frac{\partial E}{\partial w_{kj}} = \boxed{\frac{\partial E}{\partial s_j}\frac{\partial s_j}{\partial w_{kj}}}$$

① $$\frac{\partial E}{\partial s_j} = \sum_{i=1}^{nout} \boxed{\frac{\partial E}{\partial s_i}\frac{\partial s_i}{\partial h_j}} \frac{\partial h_j}{\partial s_j} = \sum_{i=1}^{nout}(y_i - t_i)(w_{ji})(h_j(1-h_j))$$

$$\frac{\partial E}{\partial h_j} = \frac{\partial E}{\partial y_i} \boxed{\frac{\partial y_i}{\partial s_i}\frac{\partial s_i}{\partial h_j}}$$

$$= \sum_{i=1} \frac{\partial E}{\partial y_j}(y_i(1-y_j)w_{ji}$$

$$\boldsymbol{\frac{\partial E}{\partial y_i} = \frac{y_i - t_i}{y_i(1-y_i)} \cdots\cdots\cdots a}$$

$$\frac{\partial y_i}{\partial s_i} = y_i \cdot (1-y_i) \quad \cdots\cdots b$$

은닉층의 가중치 업데이트

1 오차역전파

$$\frac{\partial E}{\partial w_{kj}} = \boxed{\frac{\partial E}{\partial s_j}\frac{\partial s_j}{\partial w_{kj}}}$$

① $\frac{\partial E}{\partial s_j} = \sum_{i=1}^{nout}\frac{\partial E}{\partial s_i}\frac{\partial s_i}{\partial h_j}\frac{\partial h_j}{\partial s_j} = \sum_{i=1}^{nout}(y_i - t_i)(w_{ji})(h_j(1-h_j))$

② $\frac{\partial s_j}{\partial w_{kj}} = w_k(\because s_j = \sum_{k=1} x_k w_{kj})$

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial s_j}\frac{\partial s_j}{\partial w_{kj}} = \sum_{i=1}^{nout}(y_i - t_i)(w_{ji})(h_j(1-h_j))(x_k)$$

은닉층의 가중치 업데이트

1 오차역전파

General Multiplayer Network

각 layer의 $\frac{\partial E}{\partial w_{ij}}$ 를 구하고 싶을 때,

1. 먼저 $\frac{\partial E}{\partial s_j}$ 를 계산한다.

2. $\frac{\partial s_j}{\partial w_{kj}} = x_k$ 를 1번에 곱해준다.

➔ 즉, 델타식을 먼저 구하고 나면 output의 델타식으로 다시 변환 가능하다.

$$\boxed{\begin{aligned} &\boldsymbol{new\ Output}: \boldsymbol{w_{31}(T+1) = w_{31}(T) - \delta y \cdot h_1} \\ &\boldsymbol{new\ Hidden}: \boldsymbol{w_{11}(T+1) = w_{11}(T) - \delta h \cdot x_1} \end{aligned}}$$