Lecture4 142p – Backpropagation

박주찬

DICE Lab
School of Computer Science and Engineering
KOREATECH
green261535@gmail.com

Introduction

CS231n Lecture4 142pg에 나오는 computational graph에서 W_1, W_2에 대하여 back-porpagation을 증명한다.

Computation graph는 다음 **Fig. 1.** 과 같다.

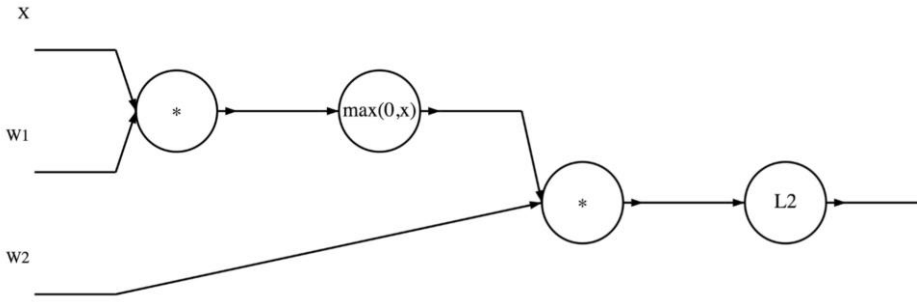

Fig. 1.

Fig. 2. Computational graph의 수식은 다음과 같다.

$z = X * W_1$

$h_1 = \mathrm{ReLU}(z)$

$y = h_1 * W_2$

$L = \|y\|^2$

X, z, W_1, W_2, h_1, y, L을 다음과 같이 정의 한다.

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, W_1 = \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix}, z = \begin{bmatrix} x_{11} * w_{11} + x_{12} * w_{12} \\ x_{21} * w_{11} + x_{22} * w_{12} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$h = \begin{bmatrix} \max(0, z_1) \\ \max(0, z_2) \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, W_2 = \begin{bmatrix} w_{21} & w_{22} \end{bmatrix}$$

$$y = \begin{bmatrix} \max(0, z_1) * w_{21} & \max(0, z_1) * w_{22} \\ \max(0, z_2) * w_{21} & \max(0, z_2) * w_{21} \end{bmatrix} = \begin{bmatrix} h_1 * w_{21} & h_1 * w_{22} \\ h_2 * w_{21} & h_2 * w_{21} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$

$L = \|y\|^2$

Derivation for $\frac{\partial L}{\partial W_2}$

$\frac{\partial \mathrm{L}}{\partial w_2}$를 구하려면 먼저 곱셈 연산에 대한 upstream과 local을 구해야 한다.

upstream은 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial \mathrm{y}} &= 2 \times \mathrm{y} \\ &= \begin{bmatrix} 2 \times \mathrm{y}_{11} & 2 \times \mathrm{y}_{12} \\ 2 \times \mathrm{y}_{21} & 2 \times \mathrm{y}_{22} \end{bmatrix}\end{aligned} \tag{1}$$

w_2의 각 원소에 대한 local은 다음과 같다.

$$\frac{\partial \mathrm{y}}{\partial w_{21}} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \quad \frac{\partial \mathrm{y}}{\partial w_{22}} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \tag{2}$$

upstream(1)과 local(2)을 곱해서 downstream을 구한다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{21}} &= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial \mathrm{L}}{\partial y_{ij}} * \frac{\partial y_{ij}}{\partial w_{21}} \\ &= 2 * y_{11} * h_1 + 2 * y_{21} * h_2\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{22}} &= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\partial \mathrm{L}}{\partial y_{ij}} * \frac{\partial y_{ij}}{\partial w_{22}} \\ &= 2 * y_{12} * h_1 + 2 * y_{22} * h_2\end{aligned} \tag{4}$$

$$\begin{aligned}&\frac{\partial \mathrm{L}}{\partial w_2} = \frac{\partial \mathrm{L}}{\partial L} * \frac{\partial \mathrm{L}}{\partial y} * \frac{\partial \mathrm{y}}{\partial w_2} \\ &= \begin{bmatrix} \frac{\partial \mathrm{L}}{\partial w_{21}} & \frac{\partial \mathrm{L}}{\partial w_{22}} \end{bmatrix} \\ &= \begin{bmatrix} 2 * y_{11} * h_1 + 2 * y_{21} * h_2 & 2 * y_{12} * h_1 + 2 * y_{22} * h_2 \end{bmatrix} \\ &= 2 * \begin{bmatrix} h_1 & h_2 \end{bmatrix} * \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \\ &= 2 * h^T * y\end{aligned} \tag{5}$$

$$\frac{\partial \mathrm{L}}{\partial w_2} = 2 * h^T * y \tag{6}$$

Derivation for $\frac{\partial L}{\partial W_1}$

$\frac{\partial \mathrm{L}}{\partial w_1}$를 구해주기 위해 먼저 $\frac{\partial \mathrm{L}}{\partial h}$를 구해준다. 방식은 upstream을 구하고 local을 구해서 서로 곱한 후 downstream을 구하는 방식으로 한다.

upstream은 (1)과 같다.

h_1, h_2에 대한local은 다음과 같다.

$$\begin{aligned}&\frac{\partial \mathrm{y}_{11}}{\partial \mathrm{h}_1} = \mathrm{w}_{21},\ \frac{\partial \mathrm{y}_{12}}{\partial \mathrm{h}_1} = \mathrm{w}_{22},\ \frac{\partial \mathrm{y}_{21}}{\partial \mathrm{h}_1} = 0,\ \frac{\partial \mathrm{y}_{22}}{\partial \mathrm{h}_1} = 0,\ \frac{\partial \mathrm{y}_{11}}{\partial \mathrm{h}_2} = 0,\ \frac{\partial \mathrm{y}_{12}}{\partial \mathrm{h}_2} = 0,\\ &\frac{\partial \mathrm{y}_{21}}{\partial \mathrm{h}_2} = \mathrm{w}_{21},\ \frac{\partial \mathrm{y}_{22}}{\partial \mathrm{h}_2} = \mathrm{w}_{22}\end{aligned} \tag{7}$$

upstream(1)과 local(7)을 곱해서 downstream을 구한다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial \mathrm{h}_1} &= \sum_{\mathrm{i}=1}^{2} \sum_{\mathrm{j}=1}^{2} \frac{\partial \mathrm{L}}{\partial \mathrm{y}_{\mathrm{ij}}} * \frac{\partial \mathrm{y}_{\mathrm{ij}}}{\partial \mathrm{h}_1} \\ &= 2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}\end{aligned} \tag{8}$$

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial \mathrm{h}_2} &= \sum_{\mathrm{i}=1}^{2} \sum_{\mathrm{j}=1}^{2} \frac{\partial \mathrm{L}}{\partial \mathrm{y}_{\mathrm{ij}}} * \frac{\partial \mathrm{y}_{\mathrm{ij}}}{\partial \mathrm{h}_2} \\ &= 2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial \mathrm{h}} &= \frac{\partial \mathrm{L}}{\partial \mathrm{L}} * \frac{\partial \mathrm{L}}{\partial \mathrm{y}} * \frac{\partial \mathrm{y}}{\partial \mathrm{h}} \\ &= \begin{bmatrix} \frac{\partial \mathrm{L}}{\partial \mathrm{h}_1} \\ \frac{\partial \mathrm{L}}{\partial \mathrm{h}_2} \end{bmatrix} \\ &= \begin{bmatrix} 2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22} \\ 2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22} \end{bmatrix}\end{aligned} \tag{10}$$

$$\frac{\partial \mathrm{L}}{\partial \mathrm{h}} = 2 * \begin{bmatrix} \mathrm{y}_{11} & \mathrm{y}_{12} \\ \mathrm{y}_{21} & \mathrm{y}_{22} \end{bmatrix} * \begin{bmatrix} \mathrm{w}_{21} \\ \mathrm{w}_{22} \end{bmatrix} = 2 * \mathrm{y} * \mathrm{w}_2^{\mathrm{T}} \tag{11}$$

$\frac{\partial \mathrm{L}}{\partial w_1}$를 구해주기 위해 $\frac{\partial h}{\partial z}$를 구해준다. 방식은 upstream을 구하고 local을 구해서 서로 곱한 후 downstream을 구하는 방식으로 한다

upstream은 (11)과 같다.

local은 다음과 같다.
ReLU(z)를 z에 대하여 미분을 하면 다음과 같이 두가지 경우가 나온다.

$$h_1(z) = \mathrm{ReLU(z)} = \max(0,\ z) = \begin{cases} \mathrm{z} & \mathrm{z} > 0 \\ 0 & \mathrm{z} \leq 0 \end{cases}$$

if) $z_1 > 0, z_2 > 0$일 때,$\frac{\partial h_1}{\partial z}, \frac{\partial h_2}{\partial z}$는 다음과 같다.

$$\frac{\partial h_1}{\partial z} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \frac{\partial h_2}{\partial z} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{12}$$

if) $z_1 > 0, z_2 \leq 0$일 때,$\frac{\partial h_1}{\partial z}, \frac{\partial h_2}{\partial z}$ 는 다음과 같다.

$$\frac{\partial h_1}{\partial z} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \frac{\partial h_2}{\partial z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{13}$$

if) $z_1 \leq 0, z_2 > 0$일 때,$\frac{\partial h_1}{\partial z}, \frac{\partial h_2}{\partial z}$ 는 다음과 같다.

$$\frac{\partial h_1}{\partial z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \frac{\partial h_2}{\partial z} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{14}$$

if) $z_1 \leq 0, z_2 \leq 0$일 때,$\frac{\partial h_1}{\partial z}, \frac{\partial h_2}{\partial z}$ 는 다음과 같다

$$\frac{\partial h_1}{\partial z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \frac{\partial h_2}{\partial z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{15}$$

위 조건들(12), (13), (14), (15)을 합해서 간단히 보면 다음과 같다

$$\frac{\partial h_1}{\partial z} = \mathrm{h}_1'(z) = \begin{bmatrix} \frac{\partial h_1}{\partial z_1} \\ \frac{\partial h_1}{\partial z_2} \end{bmatrix} = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & z_1 > 0, z_2 > 0 \\ \begin{bmatrix} 1 \\ 0 \end{bmatrix} & z_1 > 0, z_2 \le 0 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & z_1 \le 0, z_2 > 0 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & z_1 \le 0, z_2 \le 0 \end{cases}$$

$$\frac{\partial h_2}{\partial z} = \mathrm{h}_2'(z) = \begin{bmatrix} \frac{\partial h_2}{\partial z_1} \\ \frac{\partial h_2}{\partial z_2} \end{bmatrix} = \begin{cases} \begin{bmatrix} 0 \\ 1 \end{bmatrix} & z_1 > 0, z_2 > 0 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & z_1 > 0, z_2 \le 0 \\ \begin{bmatrix} 0 \\ 1 \end{bmatrix} & z_1 \le 0, z_2 > 0 \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & z_1 \le 0, z_2 \le 0 \end{cases} \tag{16}$$

$$\frac{\partial h_1}{\partial z} = \mathrm{h}_1'(z), \qquad \frac{\partial h_2}{\partial z} = \mathrm{h}_2'(z) \tag{17}$$

upstream(10)과 local(17)을 곱해서 downstream을 구한다.

$$\begin{aligned} \frac{\partial L}{\partial z} &= \frac{\partial L}{\partial h_1} * \frac{\partial h_1}{\partial z} + \frac{\partial L}{\partial h_2} * \frac{\partial h_2}{\partial z} \\ &= \left(2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}\right) * \mathrm{h}_1'(z), \\ &\qquad + \left(2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}\right) * \mathrm{h}_2'(z) \end{aligned} \tag{18}$$

위 downstream식을 z_1, z_2두가지에 대하여 모든 경우에 수를 확인해보면 아래와 같다.

if) $z_1 > 0, z_2 > 0$일 때, $\frac{\partial L}{\partial z}$는 다음과 같다.

$$\begin{aligned} \frac{\partial \mathrm{L}}{\partial \mathrm{z}} &= \begin{bmatrix} \left(2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}\right) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \left(2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}\right) \end{bmatrix} \\ &= \begin{bmatrix} \left(2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}\right) \\ \left(2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}\right) \end{bmatrix} \end{aligned} \tag{19}$$

if) $z_1 > 0, z_2 \le 0$일 때, $\frac{\partial L}{\partial z}$는 다음과 같다.

$$\begin{aligned} \frac{\partial \mathrm{L}}{\partial \mathrm{z}} &= \begin{bmatrix} \left(2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}\right) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \left(2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}\right) \\ 0 \end{bmatrix} \end{aligned} \tag{20}$$

if) $z_1 \leq 0, z_2 > 0$일 때, $\frac{\partial L}{\partial z}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial \mathrm{z}} &= \begin{bmatrix}0\\0\end{bmatrix} + \begin{bmatrix}0\\(2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22})\end{bmatrix} \\ &= \begin{bmatrix}0\\(2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22})\end{bmatrix}\end{aligned} \tag{21}$$

if) $z_1 \leq 0, z_2 \leq 0$일 때, $\frac{\partial L}{\partial z}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial \mathrm{z}} &= \begin{bmatrix}0\\0\end{bmatrix} + \begin{bmatrix}0\\0\end{bmatrix} \\ &= \begin{bmatrix}0\\0\end{bmatrix}\end{aligned} \tag{22}$$

$\frac{\partial \mathrm{L}}{\partial w_1}$를 구해주기 위해 $\frac{\partial L}{\partial z}$를 구해줬다. 이제 마지막으로 $\mathrm{z} = \mathrm{X} * \mathrm{W}_1$ 연산에 대한 upstream을 구하고 local을 구해서 서로 곱한 후 downstream을 구한다.

upstream은 (19), (20), (21), (22)와 같다.

local은 다음과 같다.

$$\frac{\partial z_1}{\partial w_{11}} = x_{11}, \quad \frac{\partial z_2}{\partial w_{11}} = x_{21}, \quad \frac{\partial z_1}{\partial w_{12}} = x_{12}, \quad \frac{\partial z_2}{\partial w_{12}} = x_{22} \tag{23}$$

upstream(19), (20), (21), (22)과 local(23)을 곱해서 downstream을 구한다.
먼저 W_1의 각 원소 별로 $\frac{\partial \mathrm{L}}{\partial w_{11}}$, $\frac{\partial \mathrm{L}}{\partial w_{12}}$를 구해준다.

if) $z_1 > 0, z_2 > 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{11}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{11}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{11}} \\ &= (2 * y_{11} * w_{21} + 2 * y_{12} * w_{22}) * x_{11} \\ &\qquad + (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * x_{21}\end{aligned} \tag{24}$$

if) $z_1 > 0, z_2 > 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{12}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{12}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{12}} \\ &= (2 * y_{11} * w_{21} + 2 * y_{12} * w_{22}) * x_{12} \\ &\quad + (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * x_{22}\end{aligned} \tag{25}$$

if) $z_1 > 0, z_2 \leq 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{11}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{11}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{11}} \\ &= (2 * y_{11} * w_{21} + 2 * y_{12} * w_{22}) * x_{11} + 0 * x_{21} \\ &= (2 * y_{11} * w_{21} + 2 * y_{12} * w_{22}) * x_{11}\end{aligned} \tag{26}$$

if) $z_1 > 0, z_2 \leq 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{12}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{12}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{12}} \\ &= (2 * y_{11} * w_{21} + 2 * y_{12} * w_{22}) * x_{12} + 0 * x_{22} \\ &= (2 * y_{11} * w_{21} + 2 * y_{12} * w_{22}) * x_{12}\end{aligned} \tag{27}$$

if) $z_1 \leq 0, z_2 > 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{11}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{11}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{11}} \\ &= 0 * x_{11} + (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * x_{21} \\ &= (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * x_{21}\end{aligned} \tag{28}$$

if) $z_1 \leq 0, z_2 > 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{12}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{12}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{12}} \\ &= 0 * x_{12} + (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * x_{22} \\ &= (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * x_{22}\end{aligned} \tag{29}$$

if) $z_1 \le 0, z_2 \le 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{11}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{11}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{11}} \\ &= 0 * x_{11} + 0 * x_{21} \\ &= 0\end{aligned} \tag{30}$$

if) $z_1 \le 0, z_2 \le 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_{12}}$ 는 다음과 같다.

$$\begin{aligned}\frac{\partial \mathrm{L}}{\partial w_{12}} &= \sum_{i=1}^{2} \frac{\partial \mathrm{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_{12}} \\ &= 0 * x_{12} + 0 * x_{22} \\ &= 0\end{aligned} \tag{31}$$

if) $z_1 > 0, z_2 > 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_1}$ 는 다음과 같다. (24), (25)사용

$$\begin{aligned}&\frac{\partial \mathrm{L}}{\partial \mathrm{w}_1} = \frac{\partial \mathrm{L}}{\partial \mathrm{L}} * \frac{\partial \mathrm{L}}{\partial \mathrm{y}} * \frac{\partial \mathrm{y}}{\partial \mathrm{h}} * \frac{\partial \mathrm{h}}{\partial \mathrm{z}} * \frac{\partial \mathrm{z}}{\partial \mathrm{w}_1} \\ &= \begin{bmatrix} \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{11}} \\ \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{12}} \end{bmatrix} \\ &= \begin{bmatrix} (2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}) * \mathrm{x}_{11} + (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) \\ (2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}) * \mathrm{x}_{12} + (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) \end{bmatrix} \\ &= 2 * \begin{bmatrix} \mathrm{x}_{11} & \mathrm{x}_{21} \\ \mathrm{x}_{12} & \mathrm{x}_{22} \end{bmatrix} * \begin{bmatrix} \mathrm{y}_{11} & \mathrm{y}_{12} \\ \mathrm{y}_{21} & \mathrm{y}_{22} \end{bmatrix} * \begin{bmatrix} \mathrm{w}_{21} \\ \mathrm{w}_{22} \end{bmatrix} \\ &= 2 * \mathrm{X}^{\mathrm{T}} * \mathrm{y} * \mathrm{W}_2^{\mathrm{T}}\end{aligned} \tag{32}$$

if) $z_1 > 0, z_2 \le 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_1}$ 는 다음과 같다. (26), (27)사용

$$\frac{\partial \mathrm{L}}{\partial \mathrm{w}_1} = \frac{\partial \mathrm{L}}{\partial \mathrm{L}} * \frac{\partial \mathrm{L}}{\partial \mathrm{y}} * \frac{\partial \mathrm{y}}{\partial \mathrm{h}} * \frac{\partial \mathrm{h}}{\partial \mathrm{z}} * \frac{\partial \mathrm{z}}{\partial \mathrm{w}_1}$$

$$= \begin{bmatrix} \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{11}} \\ \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{12}} \end{bmatrix}$$

$$= \begin{bmatrix} (2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}) * \mathrm{x}_{11} \\ (2 * \mathrm{y}_{11} * \mathrm{w}_{21} + 2 * \mathrm{y}_{12} * \mathrm{w}_{22}) * \mathrm{x}_{12} \end{bmatrix} \quad (33)$$

$$= 2 * \begin{bmatrix} \mathrm{x}_{11} \\ \mathrm{x}_{12} \end{bmatrix} * [\mathrm{y}_{11} \quad \mathrm{y}_{12}] * \begin{bmatrix} \mathrm{w}_{21} \\ \mathrm{w}_{22} \end{bmatrix}$$

$$= 2 * (X^T \text{의 첫번째} column) * (y\text{의 첫번째} row) * W_2^T$$

if) $z_1 \leq 0, z_2 > 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_1}$는 다음과 같다. (28), (29)사용

$$\frac{\partial \mathrm{L}}{\partial \mathrm{w}_1} = \frac{\partial \mathrm{L}}{\partial \mathrm{L}} * \frac{\partial \mathrm{L}}{\partial \mathrm{y}} * \frac{\partial \mathrm{y}}{\partial \mathrm{h}} * \frac{\partial \mathrm{h}}{\partial \mathrm{z}} * \frac{\partial \mathrm{z}}{\partial \mathrm{w}_1}$$

$$= \begin{bmatrix} \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{11}} \\ \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{12}} \end{bmatrix}$$

$$= \begin{bmatrix} (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * \mathrm{x}_{21} \\ (2 * \mathrm{y}_{21} * \mathrm{w}_{21} + 2 * \mathrm{y}_{22} * \mathrm{w}_{22}) * \mathrm{x}_{22} \end{bmatrix} \quad (34)$$

$$= 2 * \begin{bmatrix} \mathrm{x}_{21} \\ \mathrm{x}_{22} \end{bmatrix} * [\mathrm{y}_{21} \quad \mathrm{y}_{22}] * \begin{bmatrix} \mathrm{w}_{21} \\ \mathrm{w}_{22} \end{bmatrix}$$

$$= 2 * (X^T \text{의 두번째} column) * (y\text{의 두번째} row) * W_2^T$$

if) $z_1 \leq 0, z_2 \leq 0$일 때, $\frac{\partial \mathrm{L}}{\partial w_1}$는 다음과 같다. (30), (31)사용

$$\frac{\partial \mathrm{L}}{\partial \mathrm{w}_1} = \frac{\partial \mathrm{L}}{\partial \mathrm{L}} * \frac{\partial \mathrm{L}}{\partial \mathrm{y}} * \frac{\partial \mathrm{y}}{\partial \mathrm{h}} * \frac{\partial \mathrm{h}}{\partial \mathrm{z}} * \frac{\partial \mathrm{z}}{\partial \mathrm{w}_1}$$
$$= \begin{bmatrix} \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{11}} \\ \frac{\partial \mathrm{L}}{\partial \mathrm{w}_{12}} \end{bmatrix} \tag{35}$$
$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

References

1. http://cs231n.stanford.edu/ lecture 4