Linear Layer의 Backpropagation

수업 목표

이번 수업의 핵심:

- 벡터 및 행렬을 통한 편미분
- Linear layer의 Backpropagation

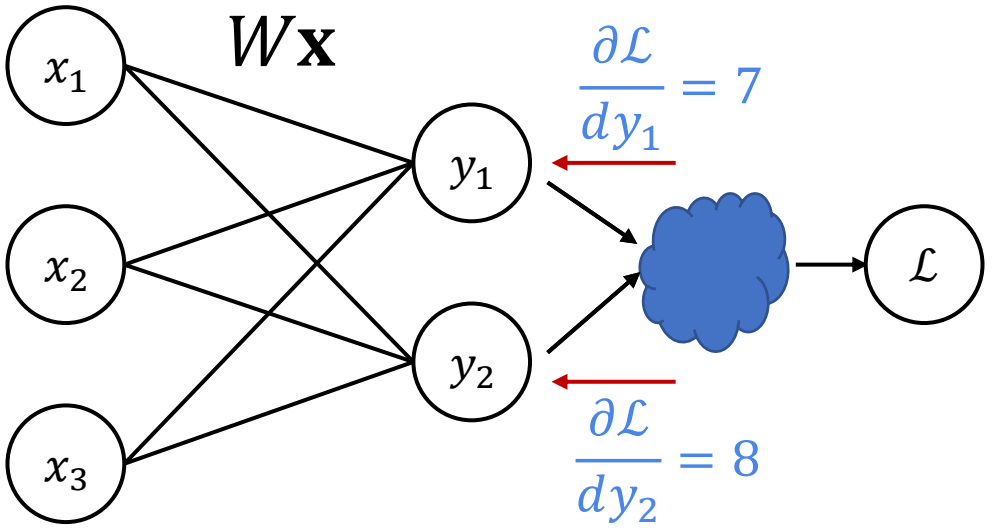
핵심 개념

- 벡터 및 행렬의 미분
- Linear layer backpropagation

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



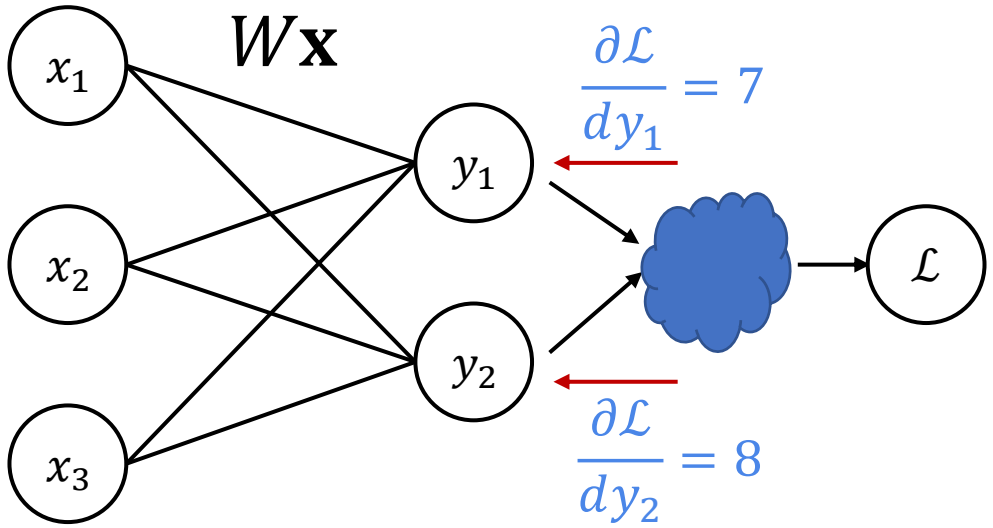
$\mathbf{x}$에 대한 Gradient:

$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} \frac{d\mathcal{L}}{dx_1} \\ \frac{d\mathcal{L}}{dx_2} \\ \frac{d\mathcal{L}}{dx_3} \end{bmatrix} \qquad \frac{d\mathcal{L}}{d\mathbf{y}} = \begin{bmatrix} \frac{d\mathcal{L}}{dy_1} \\ \frac{d\mathcal{L}}{dy_2} \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



$\mathbf{x}$에 대한 Gradient:

$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} \frac{d\mathcal{L}}{dx_1} \\ \frac{d\mathcal{L}}{dx_2} \\ \frac{d\mathcal{L}}{dx_3} \end{bmatrix}$$

$$\frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial x_1} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

$$\frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial x_2} + \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial x_2} = \begin{bmatrix} \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

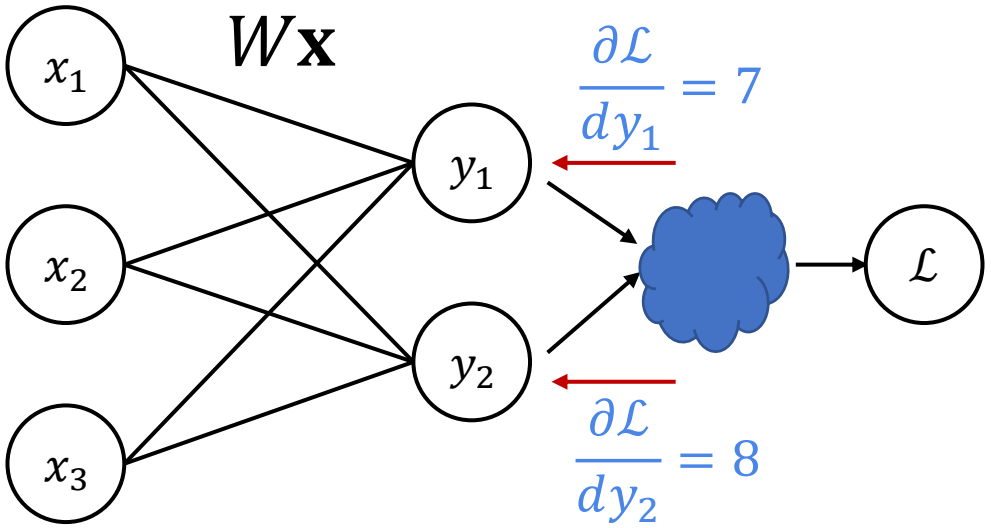
$$\frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial x_3} + \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial x_3} = \begin{bmatrix} \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \\ \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



$\mathbf{x}$에 대한 Gradient:

$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \\ \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

$$y_1 = 1 \times x_1 + 2x_2 + 3x_3$$

$$\frac{\partial y_1}{\partial x_1} = 1, \quad \frac{\partial y_1}{\partial x_2} = 2, \quad \frac{\partial y_1}{\partial x_3} = 3$$

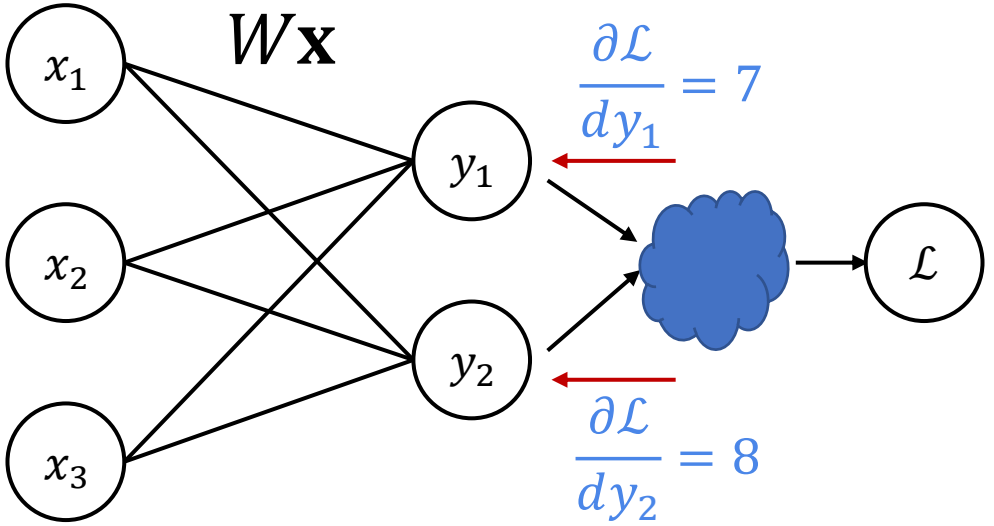
$$y_2 = 4x_1 + 5x_2 + 6x_3$$

$$\frac{\partial y_2}{\partial x_1} = 4, \quad \frac{\partial y_2}{\partial x_2} = 5, \quad \frac{\partial y_2}{\partial x_3} = 6$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



$\mathbf{x}$에 대한 Gradient:

$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \\ \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

$y_1 = 1 \times x_1 + 2x_2 + 3x_3$

$\frac{\partial y_1}{\partial x_1} = 1$, $\frac{\partial y_1}{\partial x_2} = 2$, $\frac{\partial y_1}{\partial x_3} = 3$

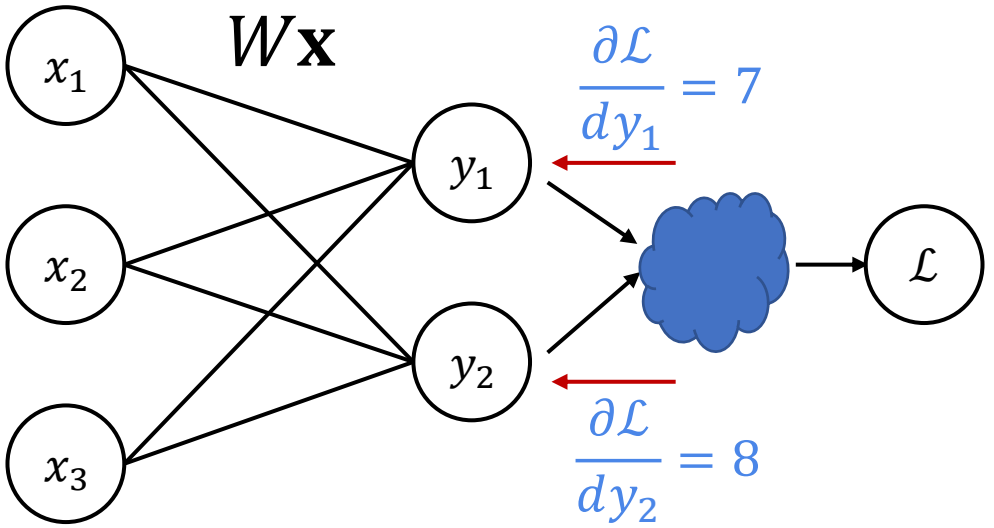
$y_2 = 4x_1 + 5x_2 + 6x_3$

$\frac{\partial y_2}{\partial x_1} = 4$, $\frac{\partial y_2}{\partial x_2} = 5$, $\frac{\partial y_2}{\partial x_3} = 6$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



$\mathbf{x}$에 대한 Gradient:

$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} 1 & \frac{\partial y_2}{\partial x_1} \\ 2 & \frac{\partial y_2}{\partial x_2} \\ 3 & \frac{\partial y_2}{\partial x_3} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

$$y_1 = 1 \times x_1 + 2x_2 + 3x_3$$

$$\frac{\partial y_1}{\partial x_1} = 1, \quad \frac{\partial y_1}{\partial x_2} = 2, \quad \frac{\partial y_1}{\partial x_3} = 3$$

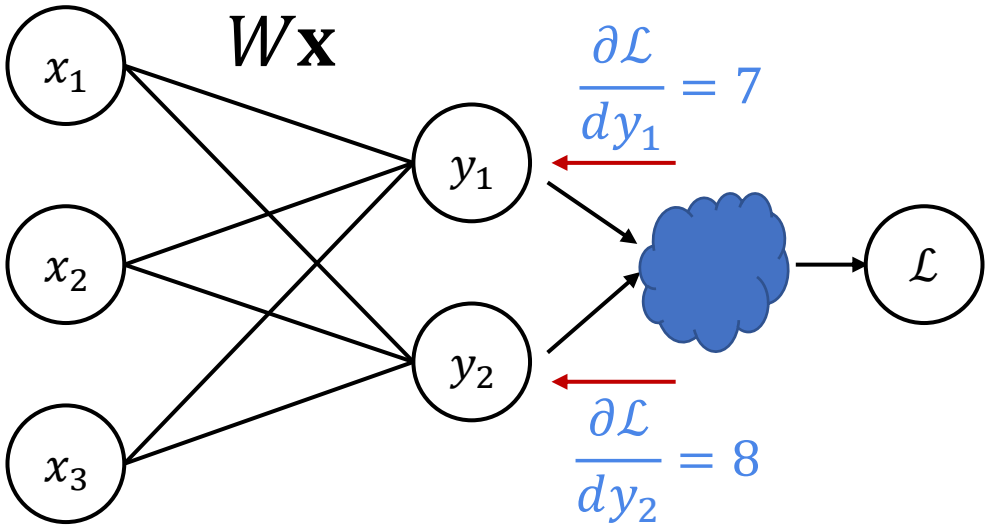
$$y_2 = 4x_1 + 5x_2 + 6x_3$$

$$\frac{\partial y_2}{\partial x_1} = 4, \quad \frac{\partial y_2}{\partial x_2} = 5, \quad \frac{\partial y_2}{\partial x_3} = 6$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

y_1	=	$w_{1,1}$ = 1	$w_{1,2}$ = 2	$w_{1,3}$ = 3	x_1 = 9
y_2		$w_{2,1}$ = 4	$w_{2,2}$ = 5	$w_{2,3}$ = 6	x_2 = 10
					x_3 = 11



$\mathbf{x}$에 대한 Gradient:

$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix}$$

$$y_1 = 1 \times x_1 + 2x_2 + 3x_3$$

$$\frac{\partial y_1}{\partial x_1} = 1, \quad \frac{\partial y_1}{\partial x_2} = 2, \quad \frac{\partial y_1}{\partial x_3} = 3$$

$$y_2 = 4x_1 + 5x_2 + 6x_3$$

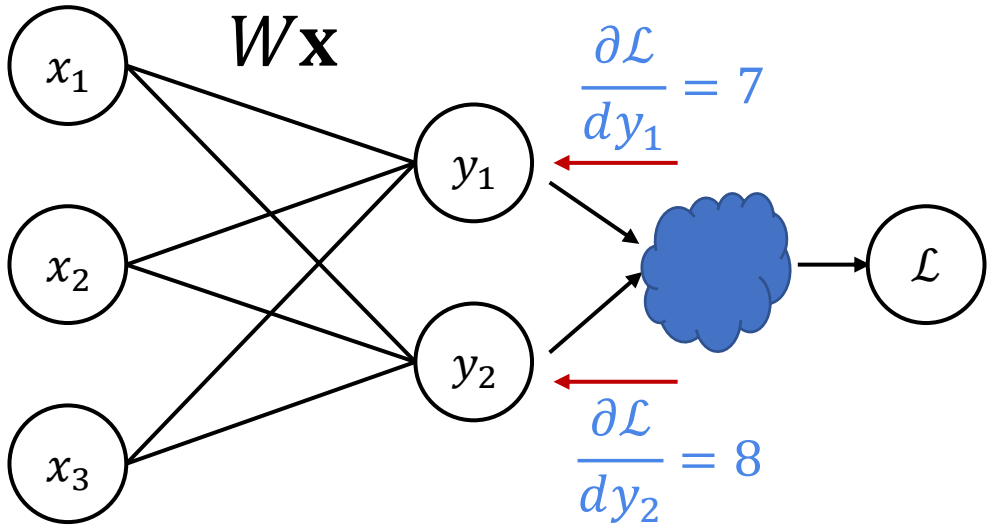
$$\frac{\partial y_2}{\partial x_1} = 4, \quad \frac{\partial y_2}{\partial x_2} = 5, \quad \frac{\partial y_2}{\partial x_3} = 6$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$

$\mathbf{x}$에 대한 Gradient:



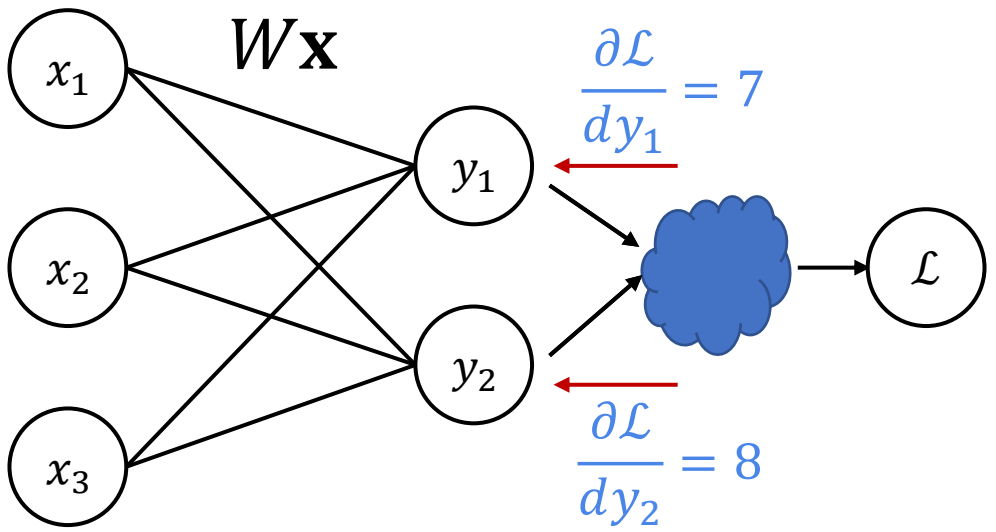
$$\frac{d\mathcal{L}}{d\mathbf{x}} = \begin{bmatrix} w_{1,1} & w_{2,1} \\ w_{1,2} & w_{2,2} \\ w_{1,3} & w_{2,3} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix} = W^\top \frac{d\mathcal{L}}{d\mathbf{y}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 39 \\ 54 \\ 69 \end{bmatrix}$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$

W에 대한 Gradient:

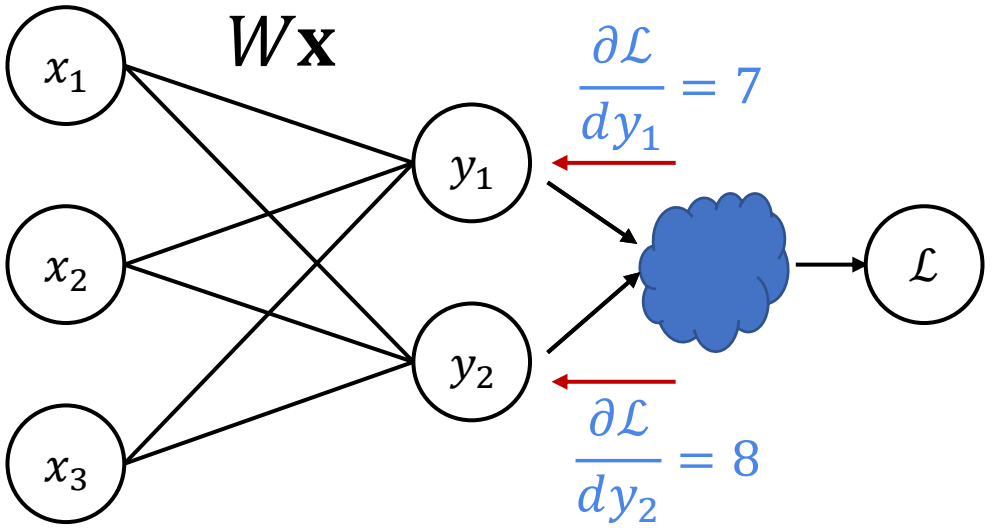


$$\frac{d\mathcal{L}}{dW} = \begin{bmatrix} \frac{d\mathcal{L}}{dw_{1,1}} & \frac{d\mathcal{L}}{dw_{1,2}} & \frac{d\mathcal{L}}{dw_{1,3}} \\ \frac{d\mathcal{L}}{dw_{2,1}} & \frac{d\mathcal{L}}{dw_{2,2}} & \frac{d\mathcal{L}}{dw_{2,3}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,1}} & \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,2}} & \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,3}} \\ \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,1}} & \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,2}} & \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,3}} \end{bmatrix}$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



W에 대한 Gradient:

$$\frac{d\mathcal{L}}{dW} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,1}} & \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,2}} & \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,3}} \\ \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,1}} & \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,2}} & \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,3}} \end{bmatrix}$$

$$y_1 = w_{1,1}x_1 + w_{1,2}x_2 + w_{1,3}x_3$$

$$\frac{\partial y_1}{\partial w_{1,1}} = x_1, \quad \frac{\partial y_1}{\partial w_{1,2}} = x_2, \quad \frac{\partial y_1}{\partial w_{1,3}} = x_3$$

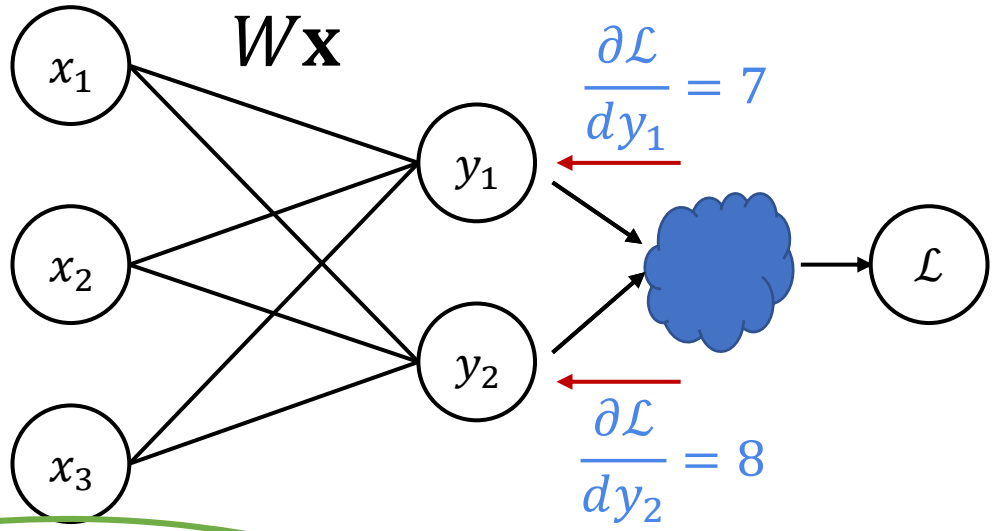
$$y_2 = w_{2,1}x_1 + w_{2,2}x_2 + w_{2,3}x_3$$

$$\frac{\partial y_2}{\partial w_{2,1}} = x_1, \quad \frac{\partial y_2}{\partial w_{2,2}} = x_2, \quad \frac{\partial y_2}{\partial w_{2,3}} = x_3$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$



W에 대한 Gradient:

$$\frac{d\mathcal{L}}{dW} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,1}} & \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,2}} & \frac{\partial \mathcal{L}}{\partial y_1}\frac{\partial y_1}{\partial w_{1,3}} \\ \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,1}} & \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,2}} & \frac{\partial \mathcal{L}}{\partial y_2}\frac{\partial y_2}{\partial w_{2,3}} \end{bmatrix}$$

$$y_1 = w_{1,1}x_1 + w_{1,2}x_2 + w_{1,3}x_3$$

$$\frac{\partial y_1}{\partial w_{1,1}} = x_1, \quad \frac{\partial y_1}{\partial w_{1,2}} = x_2, \quad \frac{\partial y_1}{\partial w_{1,3}} = x_3$$

$$y_2 = w_{2,1}x_1 + w_{2,2}x_2 + w_{2,3}x_3$$

$$\frac{\partial y_2}{\partial w_{2,1}} = x_1, \quad \frac{\partial y_2}{\partial w_{2,2}} = x_2, \quad \frac{\partial y_2}{\partial w_{2,3}} = x_3$$

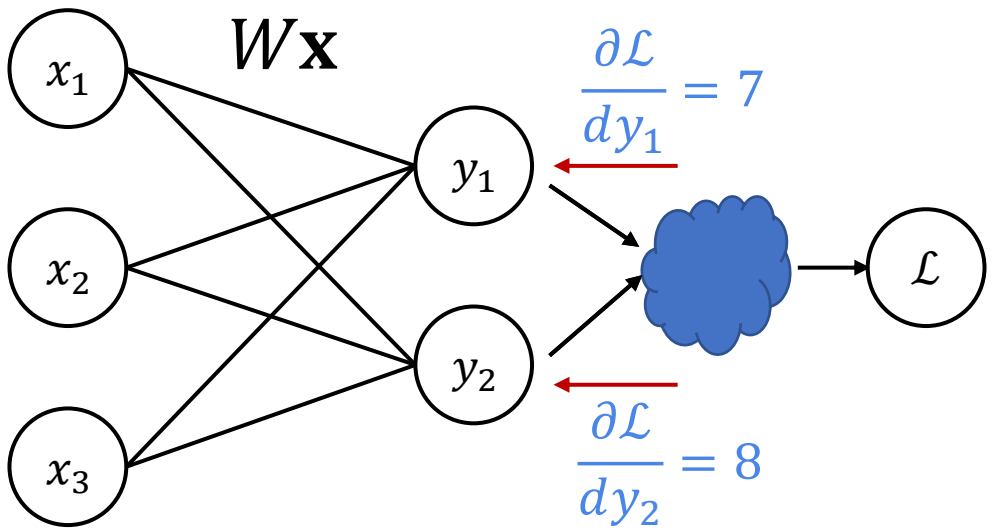
Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$

W에 대한 Gradient:

$$\frac{d\mathcal{L}}{dW} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} x_1 & \frac{\partial \mathcal{L}}{\partial y_1} x_2 & \frac{\partial \mathcal{L}}{\partial y_1} x_3 \\ \frac{\partial \mathcal{L}}{\partial y_2} x_1 & \frac{\partial \mathcal{L}}{\partial y_2} x_2 & \frac{\partial \mathcal{L}}{\partial y_2} x_3 \end{bmatrix}$$



$$y_1 = w_{1,1}x_1 + w_{1,2}x_2 + w_{1,3}x_3$$

$$\frac{\partial y_1}{\partial w_{1,1}} = x_1, \quad \frac{\partial y_1}{\partial w_{1,2}} = x_2, \quad \frac{\partial y_1}{\partial w_{1,3}} = x_3$$

$$y_2 = w_{2,1}x_1 + w_{2,2}x_2 + w_{2,3}x_3$$

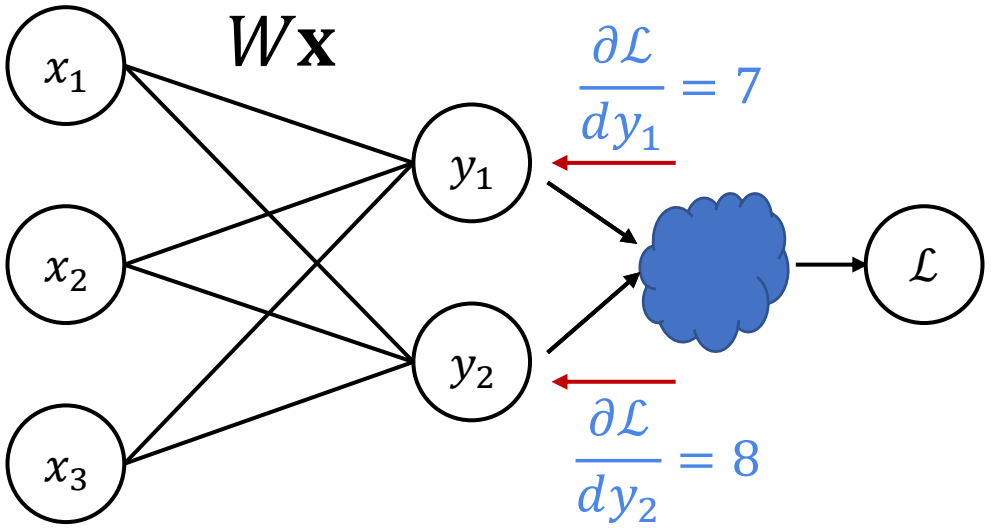
$$\frac{\partial y_2}{\partial w_{2,1}} = x_1, \quad \frac{\partial y_2}{\partial w_{2,2}} = x_2, \quad \frac{\partial y_2}{\partial w_{2,3}} = x_3$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$

W에 대한 Gradient:



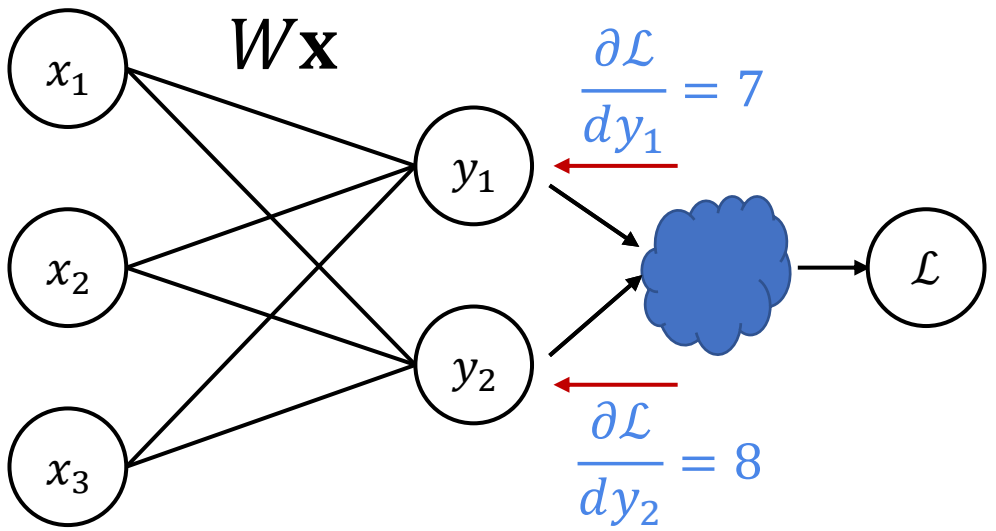
$$\frac{d\mathcal{L}}{dW} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} x_1 & \frac{\partial \mathcal{L}}{\partial y_1} x_2 & \frac{\partial \mathcal{L}}{\partial y_1} x_3 \\ \frac{\partial \mathcal{L}}{\partial y_2} x_1 & \frac{\partial \mathcal{L}}{\partial y_2} x_2 & \frac{\partial \mathcal{L}}{\partial y_2} x_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial y_1} \\ \frac{\partial \mathcal{L}}{\partial y_2} \end{bmatrix} [x_1 \quad x_2 \quad x_3] = \frac{d\mathcal{L}}{d\mathbf{y}} \mathbf{x}^\top = \begin{bmatrix} 7 \\ 8 \end{bmatrix} [9 \quad 10 \quad 11]$$

Linear Layer의 Backpropagation

예시: $\mathbf{y} = W\mathbf{x}$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{1,1} = 1 & w_{1,2} = 2 & w_{1,3} = 3 \\ w_{2,1} = 4 & w_{2,2} = 5 & w_{2,3} = 6 \end{bmatrix} \begin{bmatrix} x_1 = 9 \\ x_2 = 10 \\ x_3 = 11 \end{bmatrix}$$

$\mathbf{x}, W$에 대한 Gradient:



$$\frac{d\mathcal{L}}{d\mathbf{x}} = W^\top \frac{d\mathcal{L}}{d\mathbf{y}}, \qquad \frac{d\mathcal{L}}{dW} = \frac{d\mathcal{L}}{d\mathbf{y}} \mathbf{x}^\top$$

요약

- 행렬의 미분을 통한 Linear layer의 Backpropagation
- Linear layer의 Local gradient를 활용한 Backpropagation의 단순화