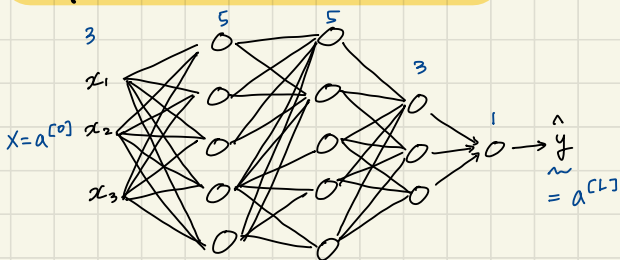


Deep Neural Network Notation



$L = \# \text{ layer, i.e. } L = 4.$

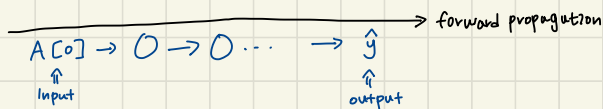
$n^{[L]} = \# \text{ units in Layer } L, \text{ i.e. } n^{[0]} = 3, n^{[1]} = 5, n^{[2]} = 5, n^{[3]} = 3, n^{[4]} = 1$

$a^{[L]} = \text{activation in layer } L, \text{ i.e. } a^{[L]} = g^{[L]}(z^{[L]})$

Forward Propagation for Layer L

$$\begin{aligned} z^{[L]} &= W^{[L]} \cdot a^{[L-1]} + b^{[L]} \\ a^{[L]} &= g^{[L]}(z^{[L]}) \end{aligned}$$

⋮ <vectorize>
 $z^{[L]} = W^{[L]} \cdot A^{[L-1]} + b^{[L]}$
 $A^{[L]} = g^{[L]}(z^{[L]})$



Backward Propagation for layer L

Input $da^{[L]}$

Output $da^{[L-1]}, dw^{[L]}, db^{[L]}$

$$dz^{[L]} = da^{[L]} * g^{[L]'}(z^{[L]})$$

$$dw^{[L]} = dz^{[L]} * a^{[L-1]}$$

$$db^{[L]} = dz^{[L]}$$

$$da^{[L-1]} = W^{[L]T} \cdot dz^{[L]}$$

$$dz^{[L-1]} = W^{[L+1]T} dz^{[L+1]} * g^{[L+1]'}(z^{[L+1]})$$

<vectorize>

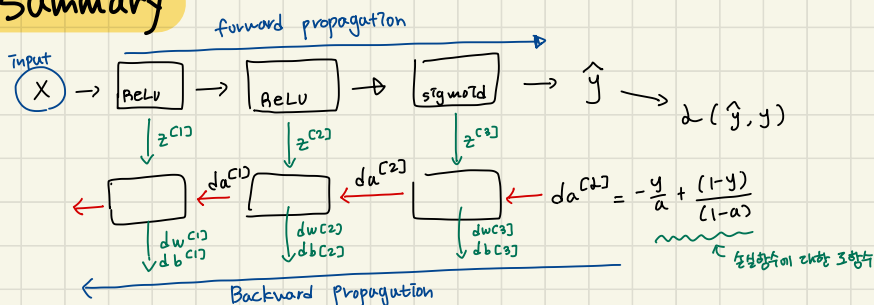
$$dz^{[L]} = da^{[L]} * g^{[L]'}(z^{[L]})$$

$$dw^{[L]} = \frac{1}{m} dz^{[L]} \cdot A^{[L-1]T}$$

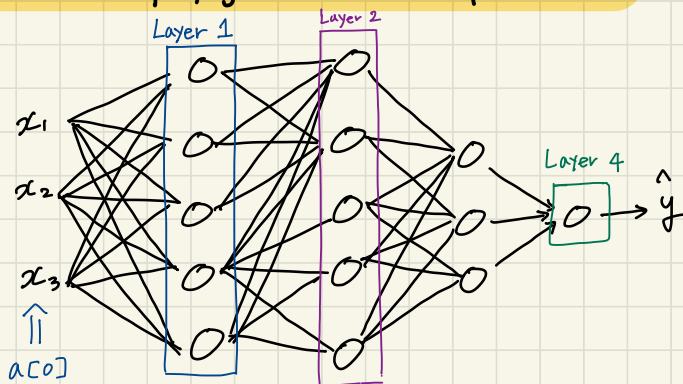
$$db^{[L]} = \frac{1}{m} \text{np.sum}(dz^{[L]}, \text{axis}=1, \text{keepdims}=\text{True})$$

$$da^{[L-1]} = W^{[L]T} * dz^{[L]}$$

Summary



Forward Propagation In a deep network



$$x : z^{C1} = W^{C1} \cdot \begin{bmatrix} a[0] \\ x \end{bmatrix} + b^{C1}$$

$$a^{C1} = g^{C1}(z^{C1})$$

~ 활성화 함수! 층마다 다를 수 있음.

$$z^{C2} = W^{C2} \cdot a^{C1} + b^{C2}$$

$$a^{C2} = g^{C2}(z^{C2})$$

$$\dots$$

$$z^{C4} = W^{C4} \cdot a^{C3} + b^{C4}$$

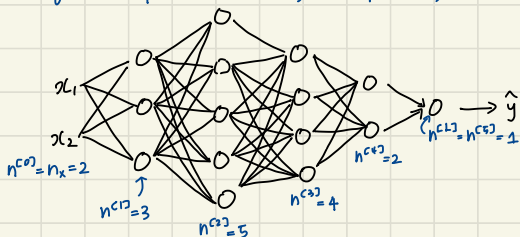
$$a^{C4} = g^{C4}(z^{C4}) = \hat{y}$$

$$\Rightarrow z^{C2} = W^{C2} \cdot a^{C1} + b^{C2}$$

$$a^{C2} = g^{C2}(z^{C2})$$

matrix dimension

- Parameter W & b .



$$\begin{aligned} z^{C1} &= W^{C1} \cdot x + b^{C1} \\ \begin{pmatrix} 3 \\ 1 \end{pmatrix} &= \begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= a^{C1} \end{aligned}$$

$(3,1) = (n^{C1}, 1)$

$(3,2) = (n^{C1}, n^{C2})$

$$W^{C1} : (n^{C1}, n^{C0})$$

$$W^{C2} : (n^{C2}, n^{C1})$$

$$\vdots$$

$$W^{C2} : (n^{C2}, n^{C1-1})$$

$$\downarrow$$

$$dW^{C2} : (n^{C2}, n^{C1-1})$$

$$b^{C1} : (n^{C1}, 1)$$

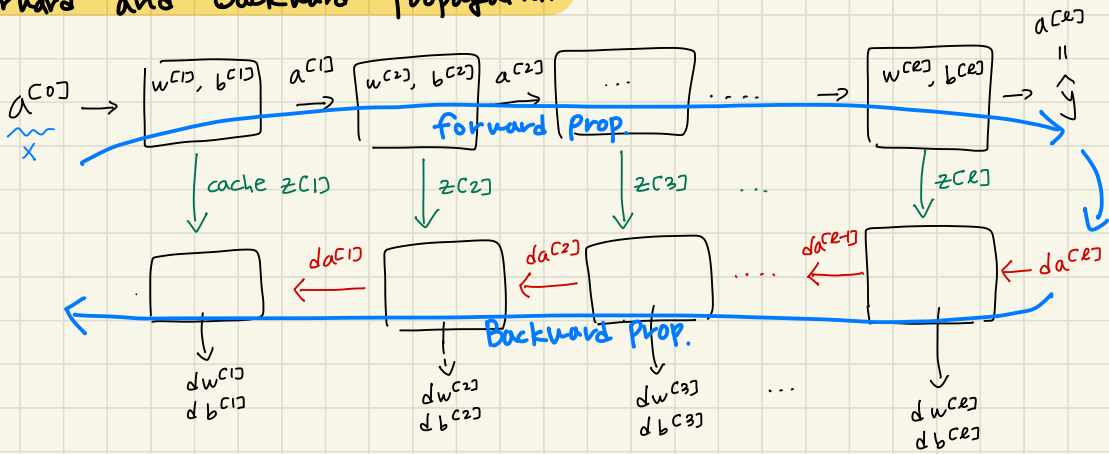
$$\vdots$$

$$b^{C2} : (n^{C2}, 1)$$

$$\downarrow$$

$$db^{C2} : (n^{C2}, 1)$$

Forward and Backward Propagation



⇒ 1번째 층에서 정방향 함수는 이전 층의 출력나 값인 $a^{[1]}$ 을 input으로 받고, $a^{[2]}$ 을 output으로 반환.
 이때, $z^{[2]}$, $w^{[2]}$, $b^{[2]}$ 값은 캐시에 저장

• 1번째 층에서 역방향 함수는 이전 층의 출력나 값인 $da^{[2]}$ 을 input으로 받고, $da^{[1]}$ 을 output으로 반환.
 이때, 이들을 계산하기 위해 정방향 함수 때 저장해두었던 캐시를 사용