khyeyoon

Broadcasting

numpy가 연산 중 다른 shape을 가진 배열을 처리하는 방법

$$\begin{bmatrix}1\\2\\3\\4\end{bmatrix} + 100 = \begin{bmatrix}101\\102\\103\\104\end{bmatrix}$$

$$100 \rightarrow \begin{bmatrix}100\\100\\100\\100\end{bmatrix}$$

$$\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix} + [100\ 200\ 300] = \begin{bmatrix}101&202&303\\104&205&306\end{bmatrix}$$

$$[100\ 200\ 300] \rightarrow \begin{bmatrix}100&200&300\\100&200&300\end{bmatrix}$$

$$\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix} + \begin{bmatrix}100\\200\end{bmatrix} = \begin{bmatrix}101&102&103\\204&205&206\end{bmatrix}$$

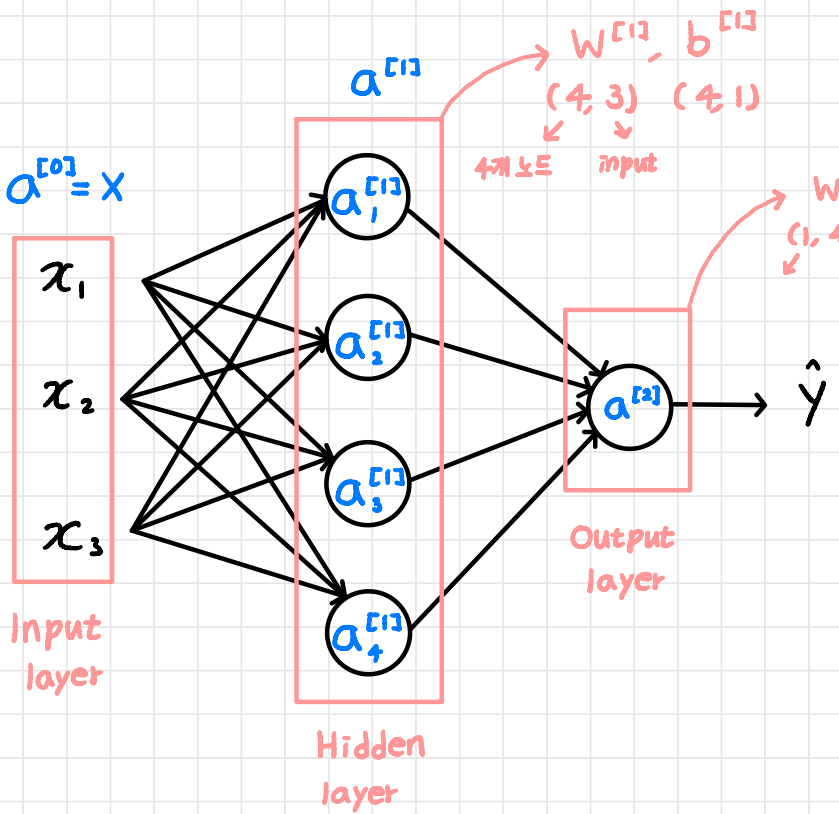
$$\begin{bmatrix}100\\200\end{bmatrix} \rightarrow \begin{bmatrix}100&100&100\\200&200&200\end{bmatrix}$$

numpy 연산 시 rank 1인 행렬은 사용하지 않기!
→ 오류가 발생할 수 있음

khyeyoon

Neural Network representation

"2 layer NN"



[표기법]

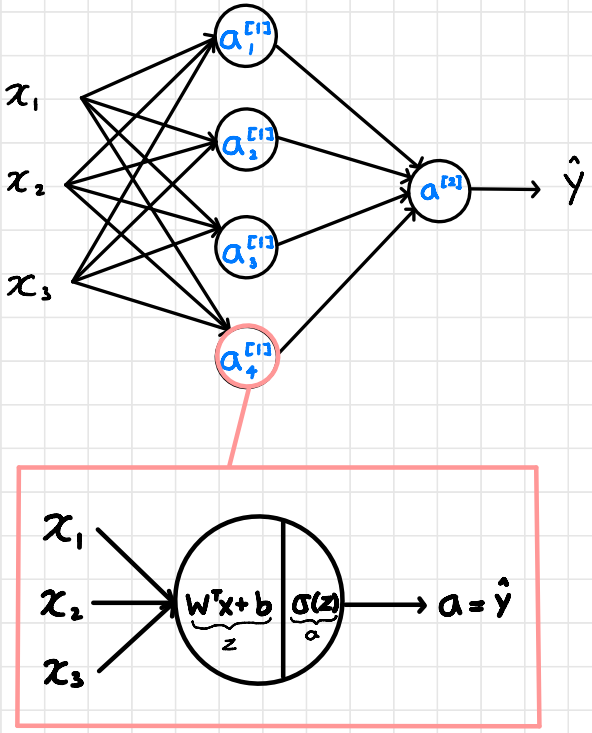
$a^{[l]}_2$

l → Layer

2 → node

Neural Network representation

Khyeyoon



Given input x:

$$z^{[1]} = W^{[1]} x + b^{[1]}$$
$(4,1)\quad (4,3)\ (3,1)\ (4,1)$

$$a^{[1]} = \sigma(z^{[1]})$$
$(4,1)\quad (4,1)$

$$z^{[2]} = W^{[2]} a^{[1]} + b^{[2]}$$
$(1,1)\quad (1,4)\ (4,1)\ (1,1)$

$$a^{[2]} = \sigma(z^{[2]})$$
$(1,1)\quad (1,1)$

khyeyoon

Vetorizing across multiple examples

for $i = 1$ to m:

$$z^{[1](i)} = W^{[1]} x^{(i)} + b^{[1]}$$

$$a^{[1](i)} = \sigma(z^{[1](i)})$$

$$z^{[2](i)} = W^{[2]} a^{[1](i)} + b^{[2]}$$

$$a^{[2](i)} = \sigma(z^{[2](i)})$$

$$Z^{[1]} = W^{[1]} X + b^{[1]}$$

$$A^{[1]} = \sigma(Z^{[1]})$$

$$Z^{[2]} = W^{[2]} A^{[1]} + b^{[2]}$$

$$A^{[2]} = \sigma(Z^{[2]})$$

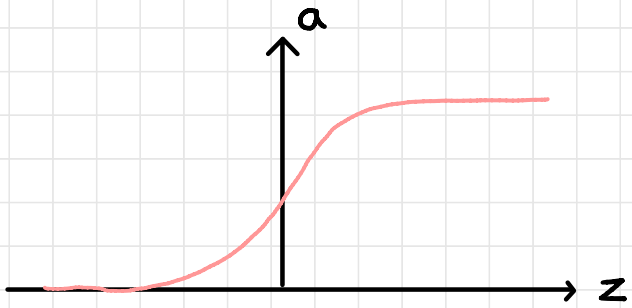
$$X = \begin{bmatrix} | & | & & | \\ x^{(1)} & x^{(2)} & \cdots & x^{(3)} \\ | & | & & | \end{bmatrix} \quad Z^{[1]} = \begin{bmatrix} | & | & & | \\ z^{[1](1)} & z^{[1](2)} & \cdots & z^{[1](m)} \\ | & | & & | \end{bmatrix} \quad A^{[1]} = \begin{bmatrix} | & | & & | \\ a^{[1](1)} & a^{[1](2)} & \cdots & a^{[1](m)} \\ | & | & & | \end{bmatrix}$$

hidden units

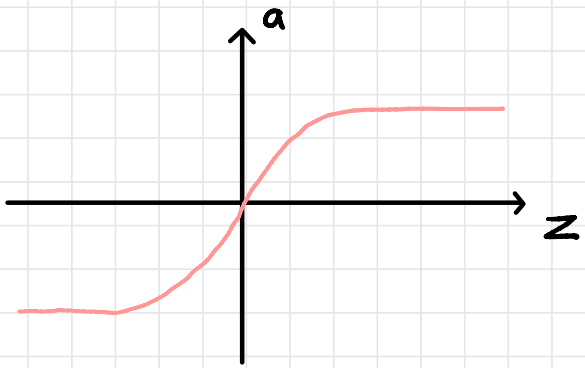
training examples

Khyeyoon

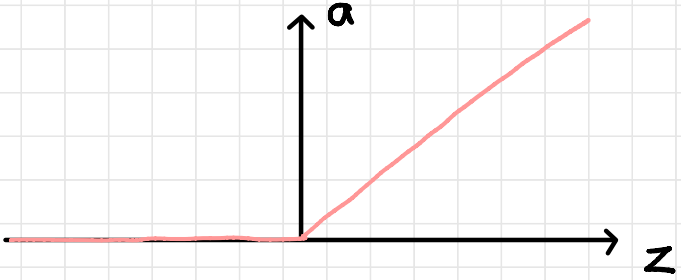
Pros and cons of activation functions



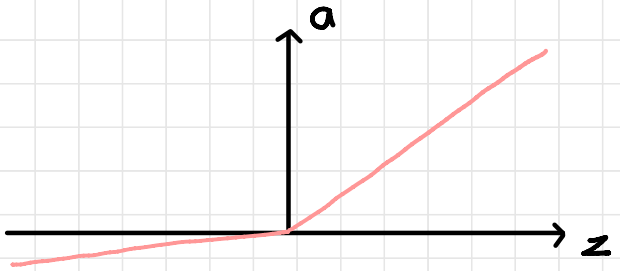
Sigmoid: $a = \frac{1}{1+e^{-z}}$



tanh: $a = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$



ReLU: $a = \max(0, z)$



Leaky ReLU: $a = \max(0.01z, z)$