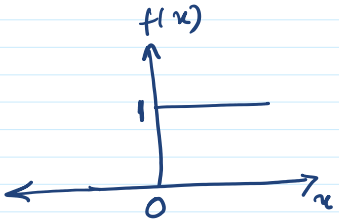
(to introduce non-linearity)



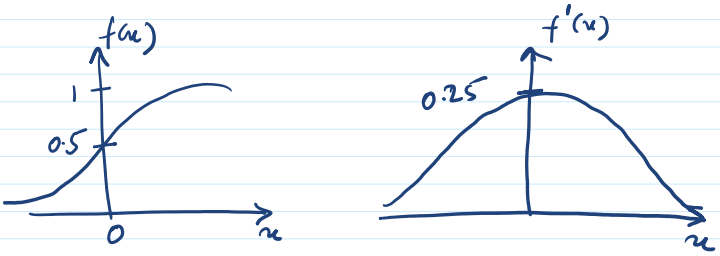
We have to take $\frac{\partial f}{\partial x}$.

$f'(x)$ at $x = 0$?

"Not differentiable"

- Sigmoid

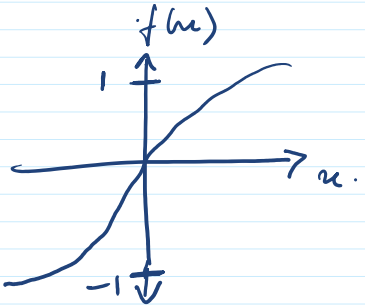
$$f(x) = \frac{1}{1+e^{-x}}$$



$$f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = f(x)\left(1 - f(x)\right)$$

- Tanh

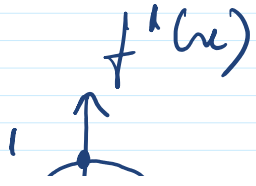
$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

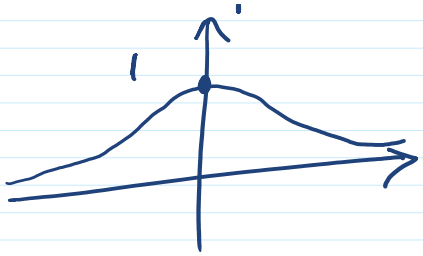


$$f'(x) = \frac{(e^{x}+e^{-x}) \times (e^{x}+e^{-x}) - (e^{x}-e^{-x})(e^{x}-e^{-x})}{(e^{x}+e^{-x})^2}$$

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

$$= 1 - \left(\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}\right)^2$$



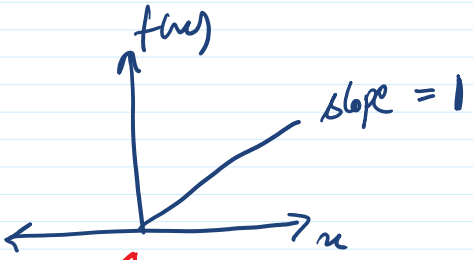


$= 1 - \left(\frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}\right)$

$= 1 - (\tanh(x))^2$

- ReLU

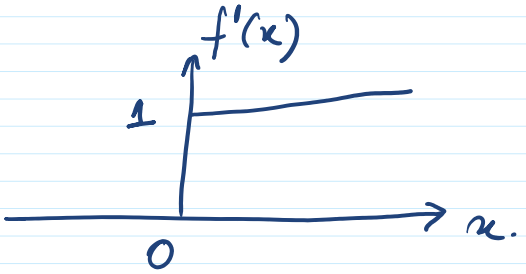
Rectified Linear Unit



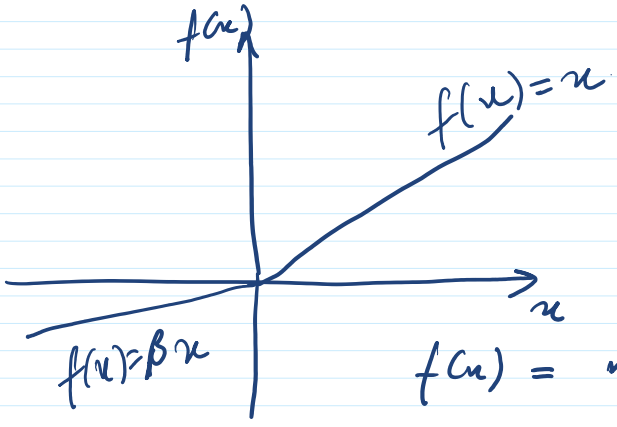
$f(x) = \max(x, 0)$

undefined

Doesn't allow -ve values to go forward.



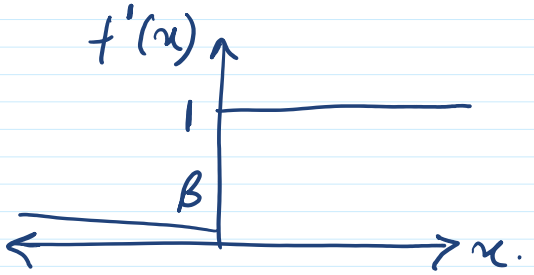
- Leaky ReLU



$f(x) = \max(\beta x, x)$

for negative x, allows a small gradient

instead of being completely zero.



Can we use any Non-linear function as activation function?

⇒ Monotonic Function

Squashing v/s non-squashing function