⟨ Activation Function ⟩ – alternative to the sigmoid activation

ex) demand prediction

price

shipping cost

marketing

material

affordability $\vec{a}_1^{[1]}$

awareness $\vec{a}_2^{[1]}$

perceived quality $\vec{a}_3^{[1]}$

– until now, we have used sigmoid function as activation function for binary classification

– degree to which possible buyers are aware of t-shirt may not be binary (little bit aware, somewhat aware, extremely aware...)
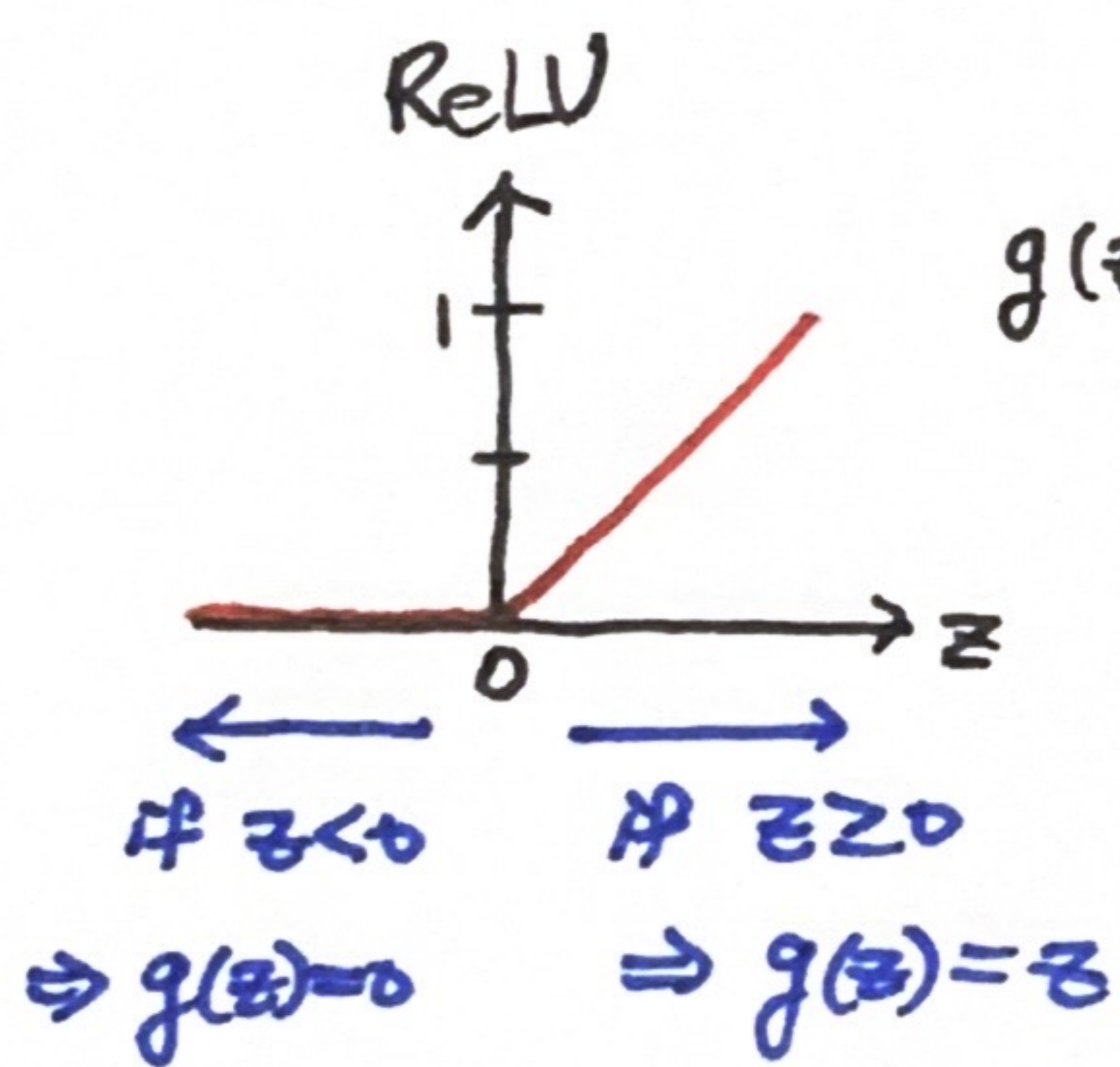
– maybe awareness should be any non-negative number

$$a_2^{[1]} = g\left(\underbrace{\vec{w}_2^{[1]} \cdot \vec{x} + b_2^{[1]}}_{z} \right)$$

Sigmoid

$$g(z) = \frac{1}{1+e^{-z}} \qquad (0 < g(z) < 1)$$

* common choice of activation function in NN : ReLU

ReLU

$$g(z) = \max(0, z) \qquad \Rightarrow g(z) \text{ could be zero or non-negative value}$$

if $z < 0$ if $z \geq 0$
$\Rightarrow g(z) = 0$ $\Rightarrow g(z) = z$

* Linear activation function

$$g(z) = z$$