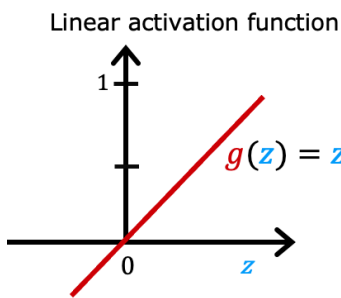


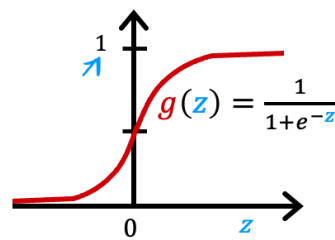
Examples of Activation Functions

"No activation function"

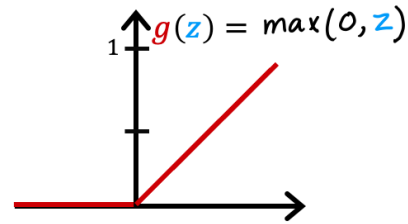


$$a_2^{[1]} = g(\overbrace{w_2^{[1]} \cdot \vec{x} + b_2^{[1]}}^z)$$

Sigmoid



ReLU Rectified Linear Unit



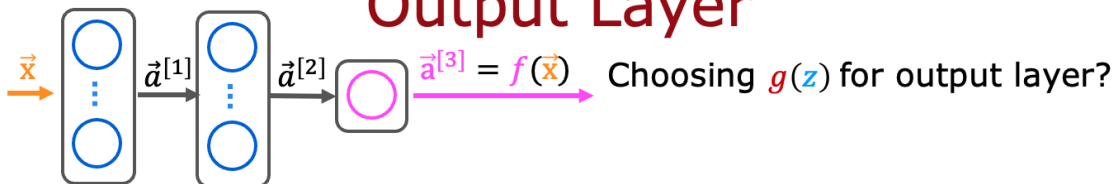
Which of the following activation functions is the most common choice for the hidden layers of a neural network?

- ☒ ReLU (rectified linear unit)
- ☐ Sigmoid
- ☐ Linear
- ☐ Most hidden layers do not use any activation function

✓ Correct

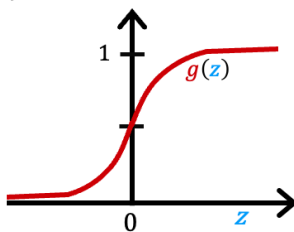
Yes! A ReLU is most often used because it is faster to train compared to the sigmoid. This is because the ReLU is only flat on one side (the left side) whereas the sigmoid goes flat (horizontal, slope approaching zero) on both sides of the curve.

Output Layer



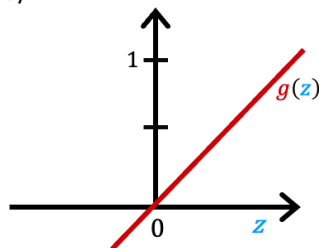
Binary classification

Sigmoid
 $y=0/1$



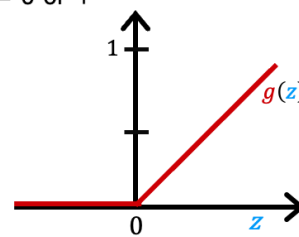
Regression

Linear activation function
 $y = +/-$



Regression

ReLU
 $Y = 0 \text{ or } +$



2. For the task of predicting housing prices, which activation functions could you choose for the output layer? Choose the 2 options that apply.

- ☒ ReLU
- ✓ Correct
Yes! ReLU outputs values 0 or greater, and housing prices are positive values.
- ☐ Sigmoid
- ☒ linear
- ✓ Correct
Yes! A linear activation function can be used for a regression task where the output can be both negative and positive, but it's also possible to use it for a task where the output is 0 or greater (like with house prices).

3. True/False? A neural network with many layers but no activation function (in the hidden layers) is not effective; that's why we should instead use the linear activation function in every hidden layer.

- ☒ False
- ☐ True
- ✓ Correct

Yes! A neural network with many layers but no activation function is not effective. A linear activation is the same as "no activation function".