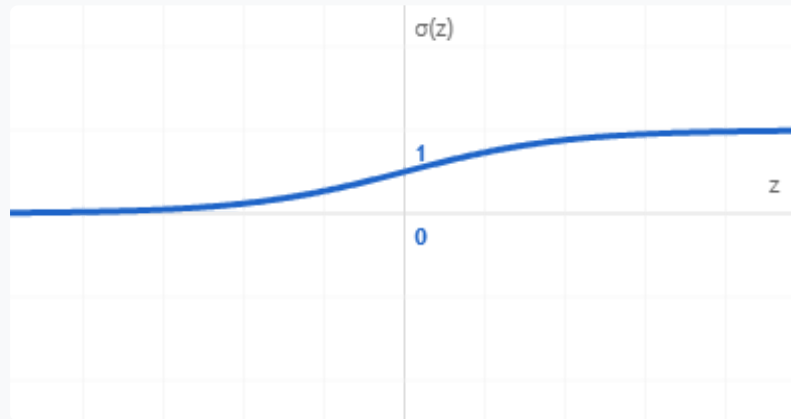


## Activation Functions: Sigmoid & Tanh

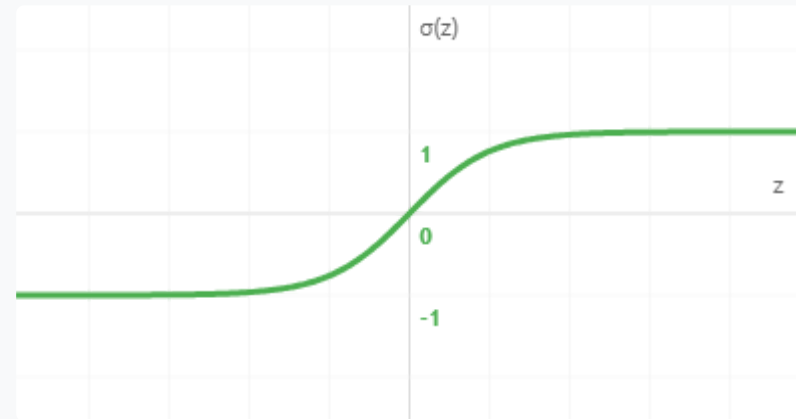
### Sigmoid ( $\sigma$ )



Formula

$$\sigma(z) = 1 / (1 + e^{-z})$$

### Tanh (tanh)



Formula

$$\tanh(z) = (e^z - e^{-z}) / (e^z + e^{-z})$$

### Key Properties

Range

(0, 1)

(-1, 1)

Zero-centered

✗ No

✓ Yes

Gradient

max = 0.25

max = 1.0

Use Case

Output layer

Hidden layers



### Common Issues



**Vanishing Gradient:** Gradients become very small for large  $|z|$ , slowing learning



**Tanh Advantage:** Zero-centered output helps optimization converge faster