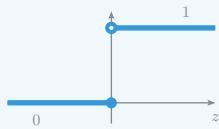


Funzioni di Attivazione

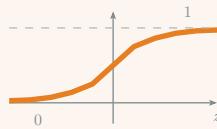
Introducono la **non linearità** che rende le reti neurali potenti

Step (gradino)



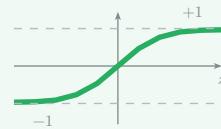
$$\sigma(z) = \begin{cases} 1 & \text{se } z > 0 \\ 0 & \text{altrimenti} \end{cases}$$

Sigmoid



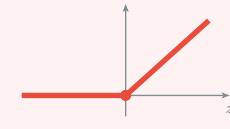
$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Tanh



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

ReLU



$$\text{ReLU}(z) = \max(0, z)$$

Perché servono? Senza funzioni di attivazione, la composizione di più strati lineari/affini collassa in un'unica trasformazione affine $\mathbf{A}\mathbf{x} + \mathbf{c}$. La non linearità rende la rete capace di rappresentare funzioni molto più complesse e, in condizioni opportune, di approssimare **qualsiasi funzione (approssimatore universale)**.