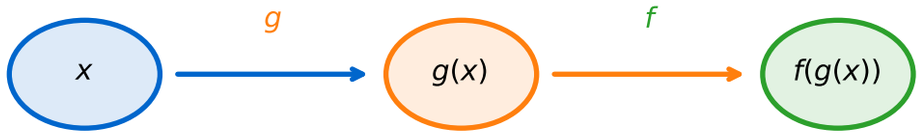
Chain Rule: Foundation of Backpropagation

Chain Rule: The Key to Backprop



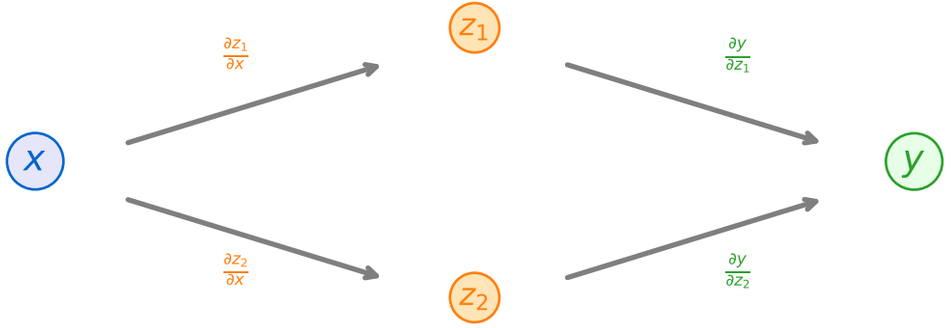
Chain Rule:

$$\frac{d}{dx} f(g(x)) = \frac{df}{dg} \cdot \frac{dg}{dx}$$

Intuition: Rate of change multiplies through the chain

Example:
$f(u) = u^2,\ \ g(x) = 3x + 1$
$\frac{df}{du} = 2u,\ \ \frac{dg}{dx} = 3$
$\frac{d}{dx} f(g(x)) = 2(3x+1) \cdot 3 = 6(3x+1)$

Multi-Variable Chain Rule



Total Derivative (sum over all paths):

$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial z_1}\frac{\partial z_1}{\partial x} + \frac{\partial y}{\partial z_2}\frac{\partial z_2}{\partial x}$$

In neural networks: gradients flow back through ALL paths