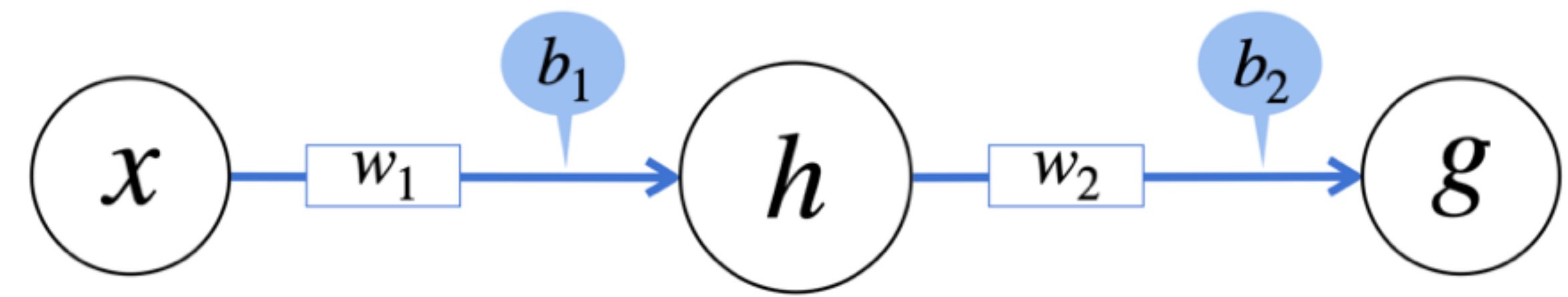


Backpropagation

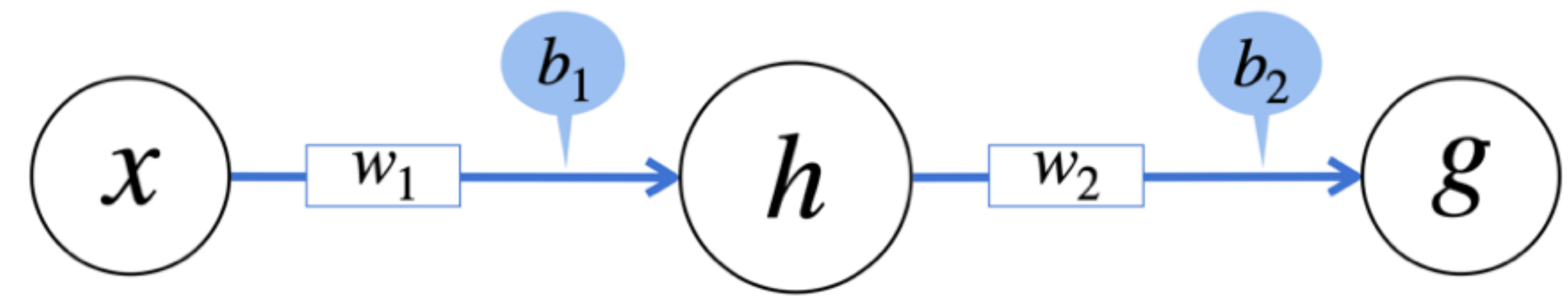


$$\frac{1}{N} \sum_i^N (y_i - g_i)^2 \implies \frac{1}{N} \sum_i^N \underbrace{(y_i)}_{\text{actual}} - \underbrace{(w_2 \cdot (w_1 \cdot x_i + b_1) + b_2)}_{\text{predicted}})^2$$

this is what the
gradient tells us!

How change w_1 to reduce our loss?
How change w_2 to reduce our loss?
How change b_1 to reduce our loss?
How change b_2 to reduce our loss?

Backpropagation



$$\frac{1}{N} \sum_i^N (y_i - g_i)^2 \implies \frac{1}{N} \sum_i^N \underbrace{(y_i}_{\text{actual}} - \underbrace{(w_2 \cdot (w_1 \cdot x_i + b_1) + b_2)}_{\text{predicted}})^2$$

$$\frac{\partial \text{Loss}}{\partial w_1} = \frac{\text{Loss}}{\partial g} \cdot \frac{\partial g}{\partial h} \cdot \frac{\partial h}{\partial w_1}$$

changing w_1 changes h , and changing h will change g ,
and changing g will change overall loss
 \implies we need the chain rule!