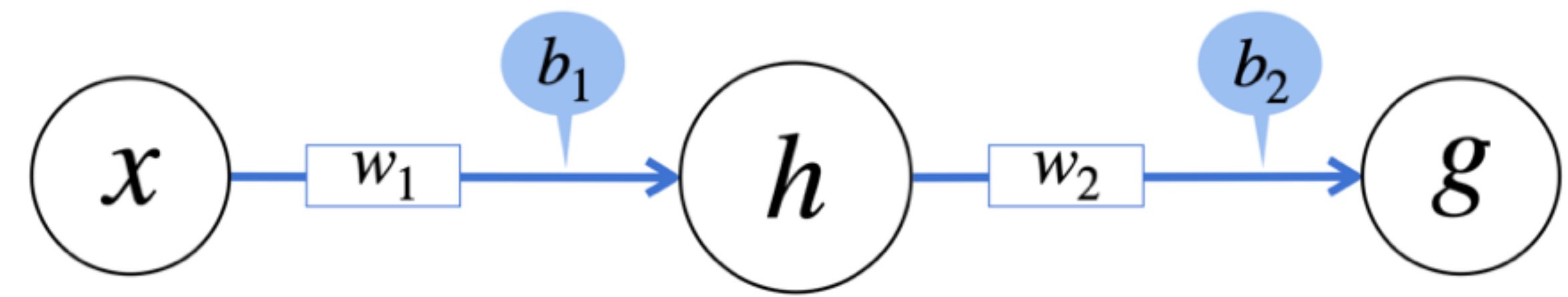


Backpropagation



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

$$\frac{\partial \text{Loss}}{\partial w_1} = \underbrace{\frac{\partial \text{Loss}}{\partial g}}_{-2(y_i - g_i)} \cdot \underbrace{\frac{\partial g}{\partial h}}_{w_2} \cdot \underbrace{\frac{\partial h}{\partial w_1}}_x$$

sum over all observations

this is the first part
of our gradient

Backpropagation

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

next part of gradient
is with respect to our
second parameter

$$\frac{\partial \text{Loss}}{\partial b_1} = \underbrace{\frac{\text{Loss}}{\partial g}}_{-2(y_i - g_i)} \cdot \underbrace{\frac{\partial g}{\partial h}}_{w_2} \cdot \underbrace{\frac{\partial h}{\partial b_1}}_1$$

sum over all observations

and so forth for all parameters...
this is backpropagation