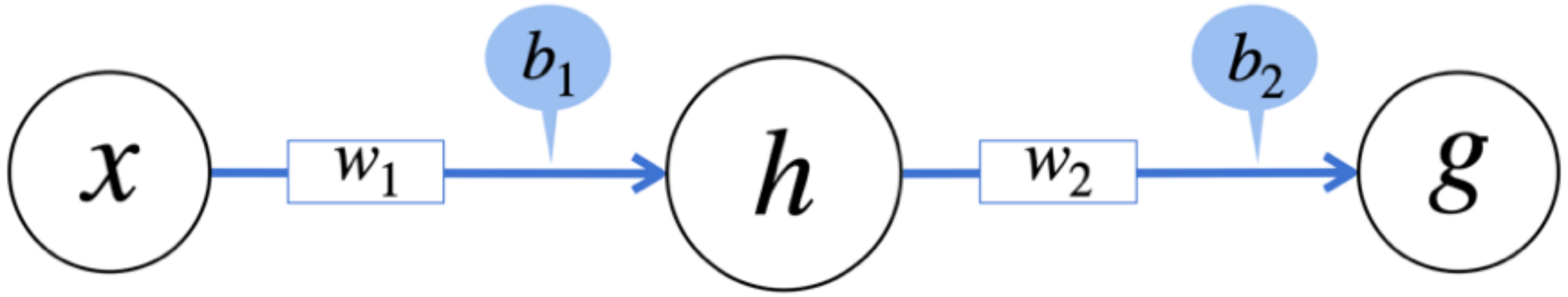
Backpropagation

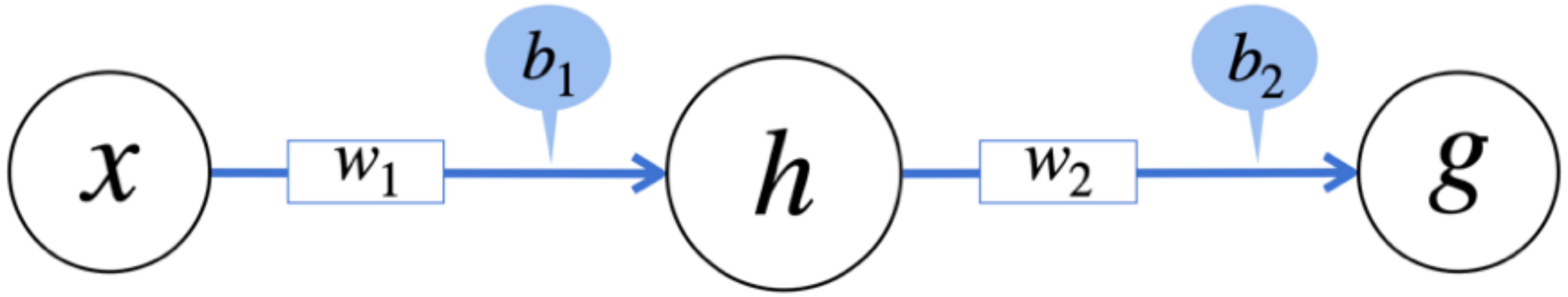


$$\frac{1}{N}\sum_{i}^{N}(y_i - g_i)^2 \implies \frac{1}{N}\sum_{i}^{N}\Big(\underbrace{y_i}_{\text{actual}} - \underbrace{(w_2 \cdot (w_1 \cdot x_i + b_1) + b_2}_{\text{predicted}}\Big)^2$$

$$\frac{\partial \text{Loss}}{\partial w_1} = \underbrace{\underbrace{\frac{\text{Loss}}{\partial g}}_{-2(y_i - g_i)} \cdot \underbrace{\frac{\partial g}{\partial h}}_{w_2} \cdot \underbrace{\frac{\partial h}{\partial w_1}}_{x}}_{\text{sum over all observations}}$$

```
this is the first part
of our gradient
```

Backpropagation



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

next part of gradient is with respect to our second parameter

$$\frac{\partial \text{Loss}}{\partial b_1} = \underbrace{\underbrace{\frac{\text{Loss}}{\partial g}}_{-2(y_i - g_i)} \cdot \underbrace{\frac{\partial g}{\partial h}}_{w_2} \cdot \underbrace{\frac{\partial h}{\partial b_1}}_{1}}_{\text{sum over all observations}}$$

and so forth for all parameters… this is backpropagation