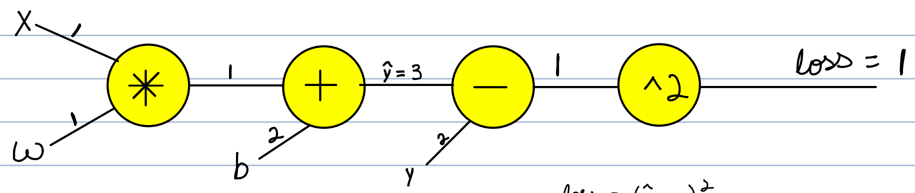
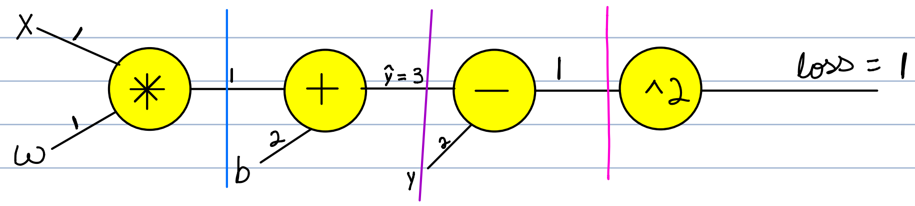
Forward pass



$loss = (\hat{y} - y)^2$

Back prop

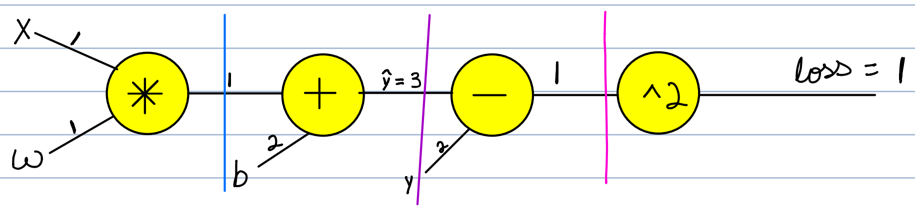
① Get local gradient w.r.t input



$\frac{\partial xw}{\partial x} = 1$ $\quad \frac{\partial (xw+b)}{\partial xw} = 1$ $\quad \frac{\partial (\hat{y} - y)}{\partial \hat{y}} = 1$ $\quad \frac{\partial (\hat{y}-y)^2}{\partial (\hat{y}-y)} = 2$

Back prop

② Get derivative of loss w.r.t each variable



$\frac{\partial xw}{\partial x} = 1$ $\quad \frac{\partial (xw+b)}{\partial xw} = 1$ $\quad \frac{\partial (\hat{y} - y)}{\partial \hat{y}} = 1$ $\quad \frac{\partial (\hat{y}-y)^2}{\partial (\hat{y}-y)} = 2(\hat{y}-y)$

$$\frac{\partial loss}{\partial (\hat{y}-y)^2} = 2(\hat{y}-y)^2 \approx 2(1) \approx 2$$

this is the loss so ≈ 1

$$\frac{\partial loss}{\partial (\hat{y}-y)} = \frac{\partial loss}{\partial (\hat{y}-y)^2} \cdot \frac{\partial (\hat{y}-y)^2}{\partial (\hat{y}-y)}$$

$$2 \cdot 2(1) \approx 4$$

$$\frac{\partial\, loss}{\partial\, xw+b} = \frac{\partial\, loss}{\partial(\hat{y}-y)^2} \cdot \frac{\partial(\hat{y}-y)^2}{\partial(\hat{y}-y)} \cdot \boxed{\frac{\partial(\hat{y}-y)}{\partial\, \hat{y}} \approx \frac{\partial(\hat{y}-y)}{\partial\, xw+b}}$$

$$2 \cdot 2(1) \cdot 1 \cdot 1 \approx 4$$

$$\frac{\partial\, loss}{\partial\, xw+b} = \frac{\partial\, loss}{\partial(\hat{y}-y)^2} \cdot \frac{\partial(\hat{y}-y)^2}{\partial(\hat{y}-y)} \cdot \frac{\partial(\hat{y}-y)}{\partial\, \hat{y}} \cdot \frac{\partial(\hat{y})}{\partial\, xw+b} \cdot \frac{\partial\, xw}{\partial x}$$

$$2(1) \cdot 2 \cdot 1 \cdot 1 \cdot 1 \approx 4$$