

(with SL)

To make predictions from data, we need:

- data $x=2$ $y=20$

- model $\hat{y} = ax + b$

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

} get best model

$\theta = \{a, b\}$

$\min_{\theta} L \rightarrow GD$

repeat:

$$\theta_{t+1} = \theta_t - \eta \frac{\partial L}{\partial \theta}$$

↓

$$\begin{aligned} m &= ax \\ \hat{y} &= m + b \\ s &= \hat{y} - y \\ L &= s^2 \end{aligned}$$

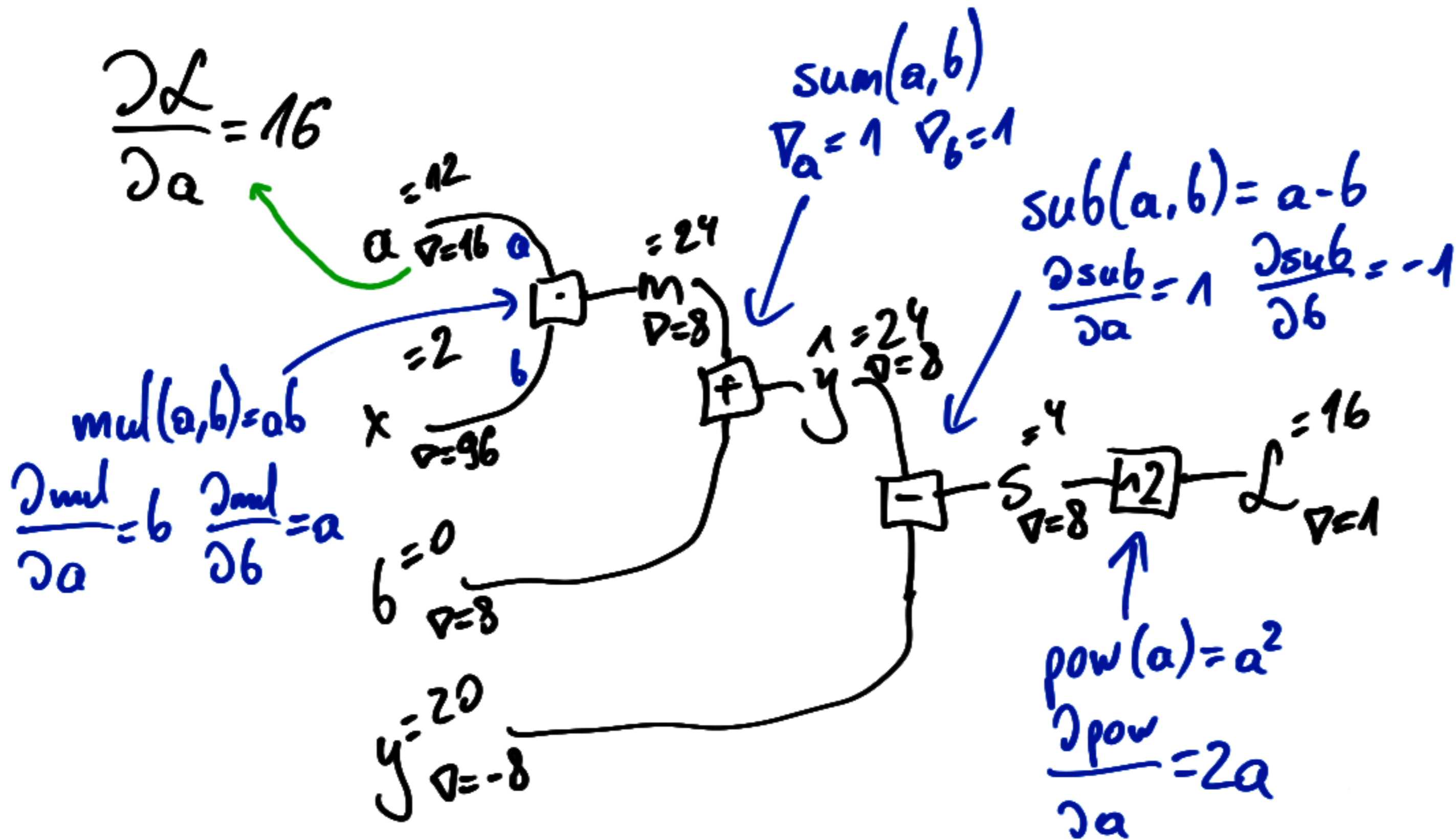
$$\frac{\partial L}{\partial a} = \frac{\partial L}{\partial s} \frac{\partial s}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial m} \frac{\partial m}{\partial a}$$

↓ ↓ ↓ ↓

2s 1 1 x

$$\rightarrow = 2(\hat{y} - y)x = 2(ax + b - y)x$$

$$\frac{\partial L}{\partial a} = 16$$



New function in lib needs:

- $\text{conv}(k, x) = k * x$
- $\frac{\partial \text{conv}}{\partial k} = \dots$
- $\frac{\partial \text{conv}}{\partial x} = \dots$