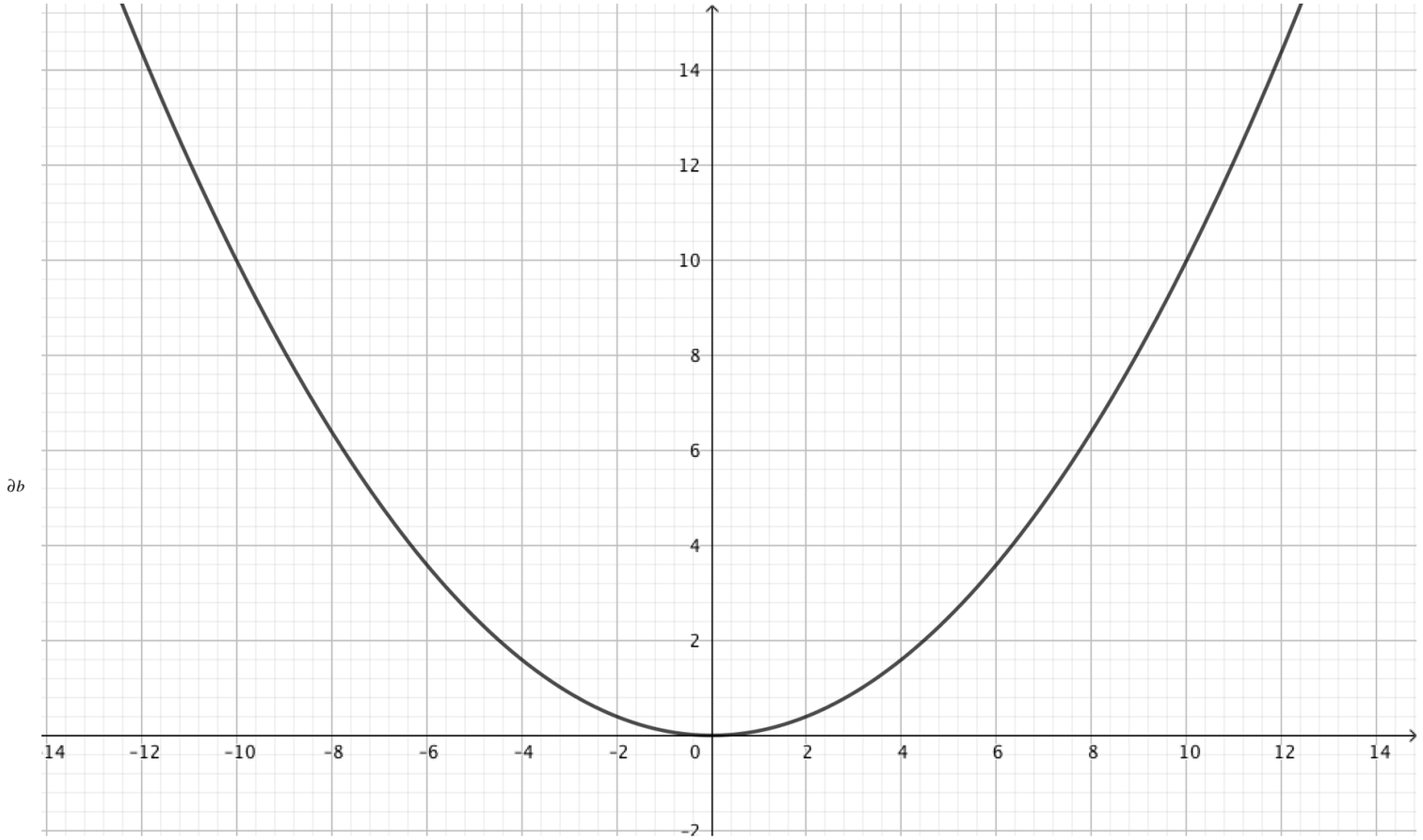
Back-propagration (Backprop)

Contents

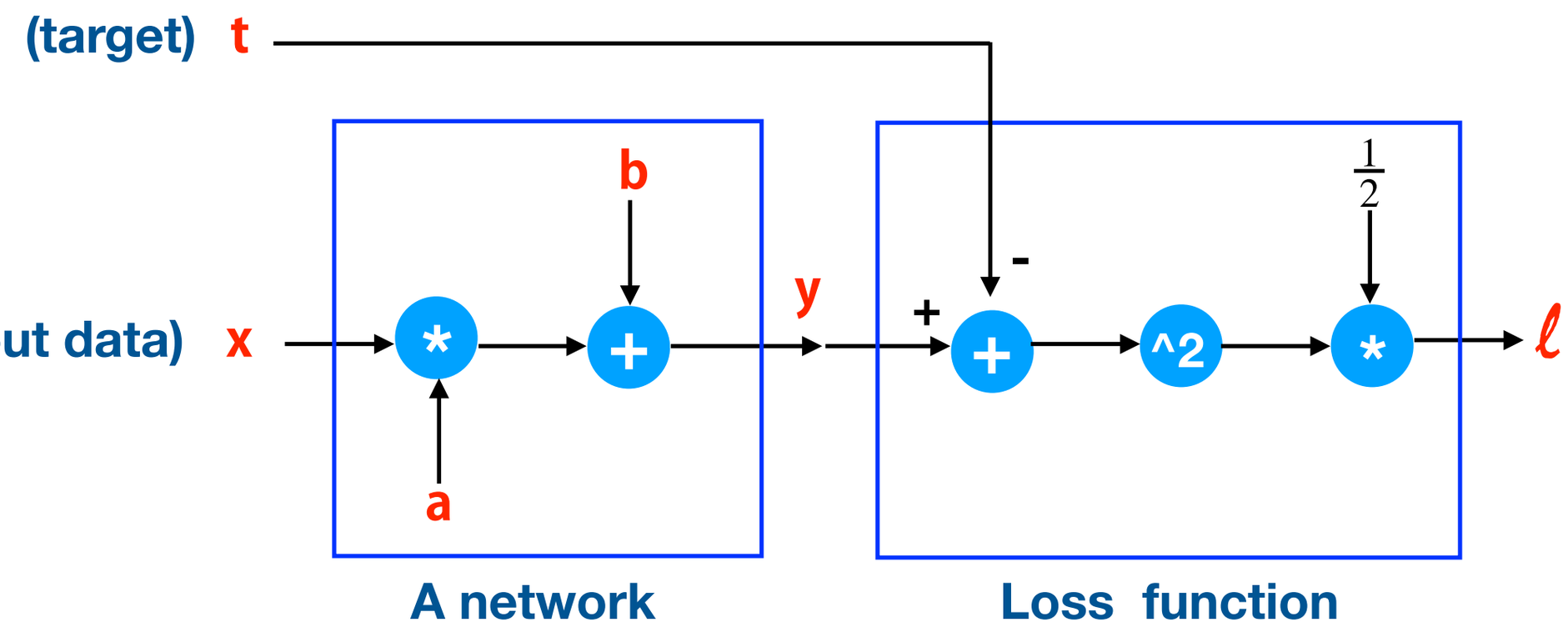
- Gradient Descent
- Back-propagation

3

Gradient Descent

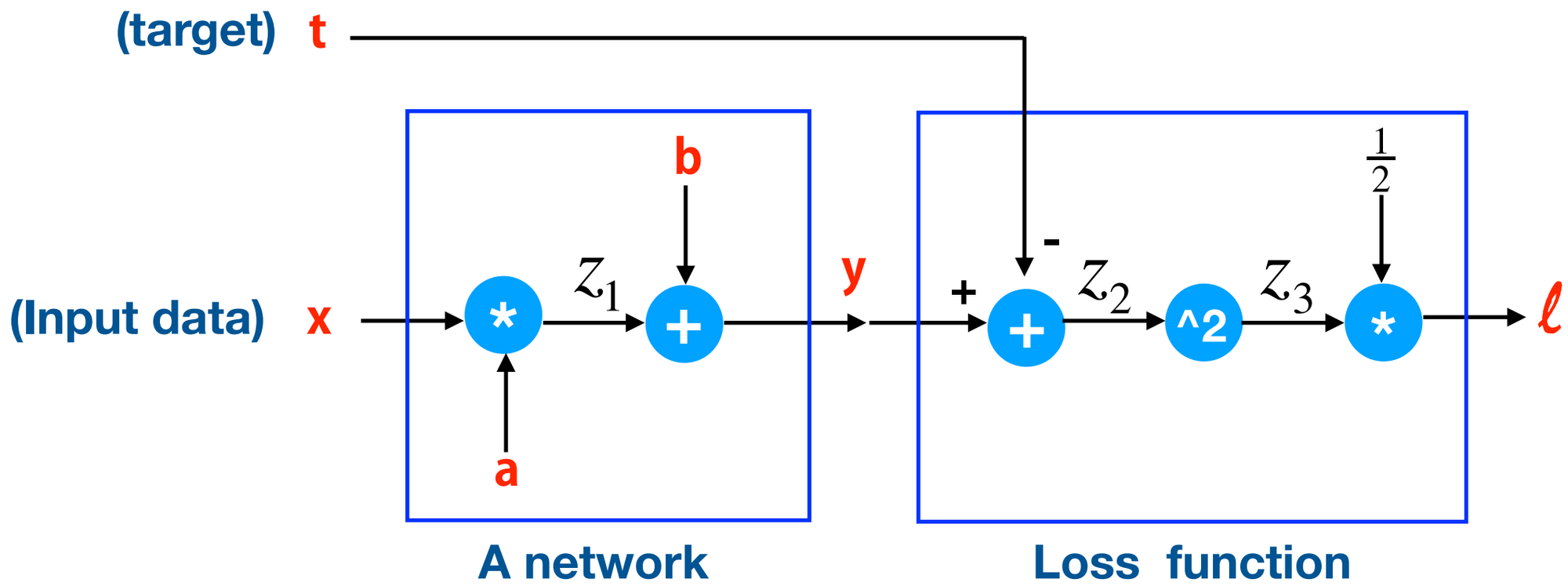


4 Simple network



∂b

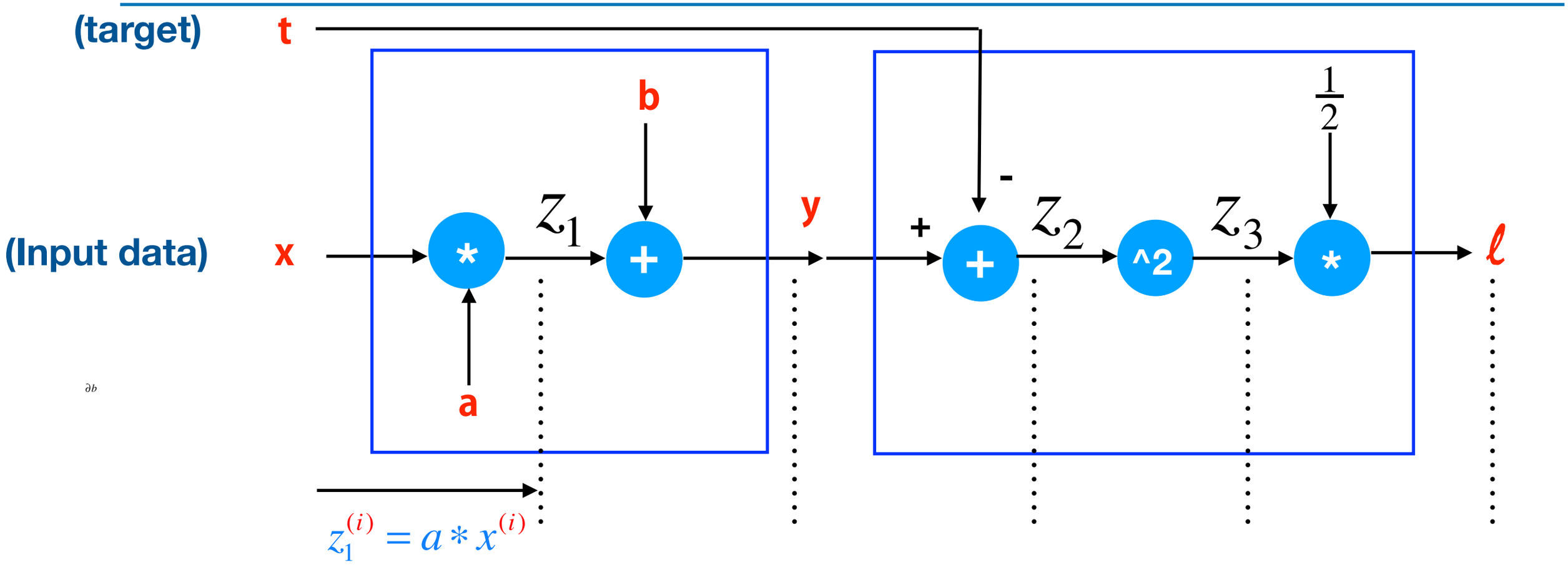
5 Simple network



∂b

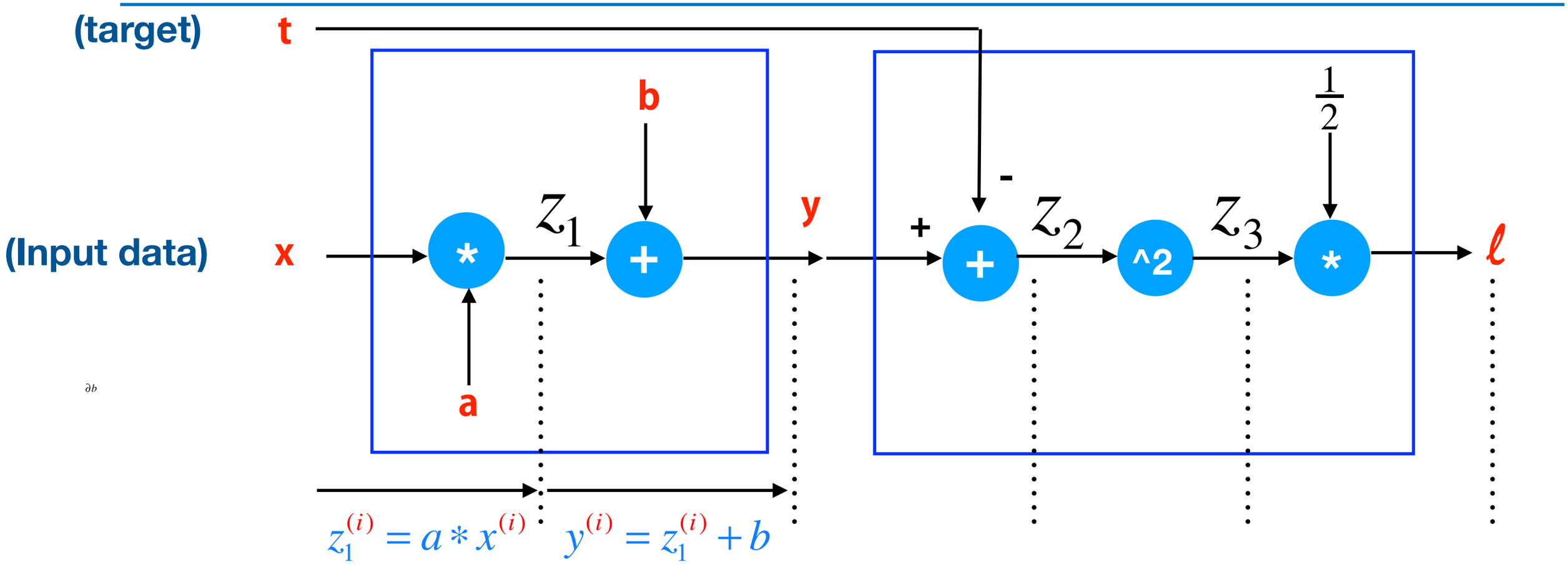
6

Back-propagation
Forward pass

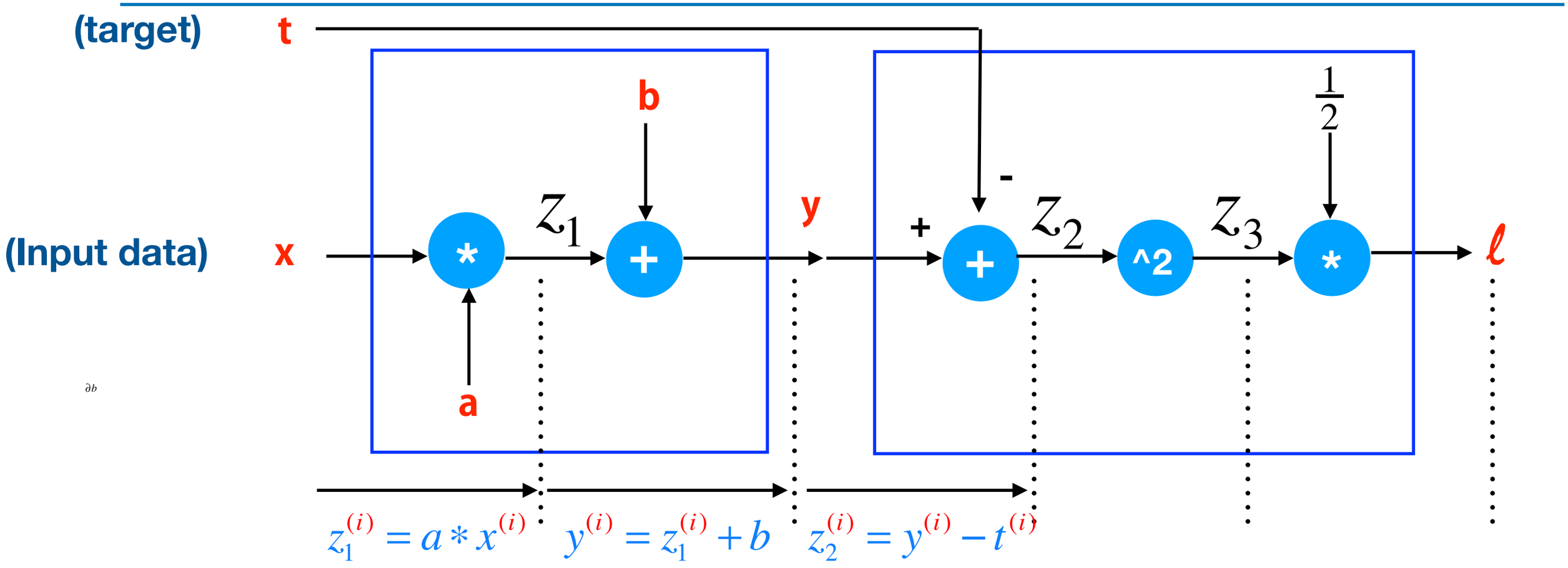


7

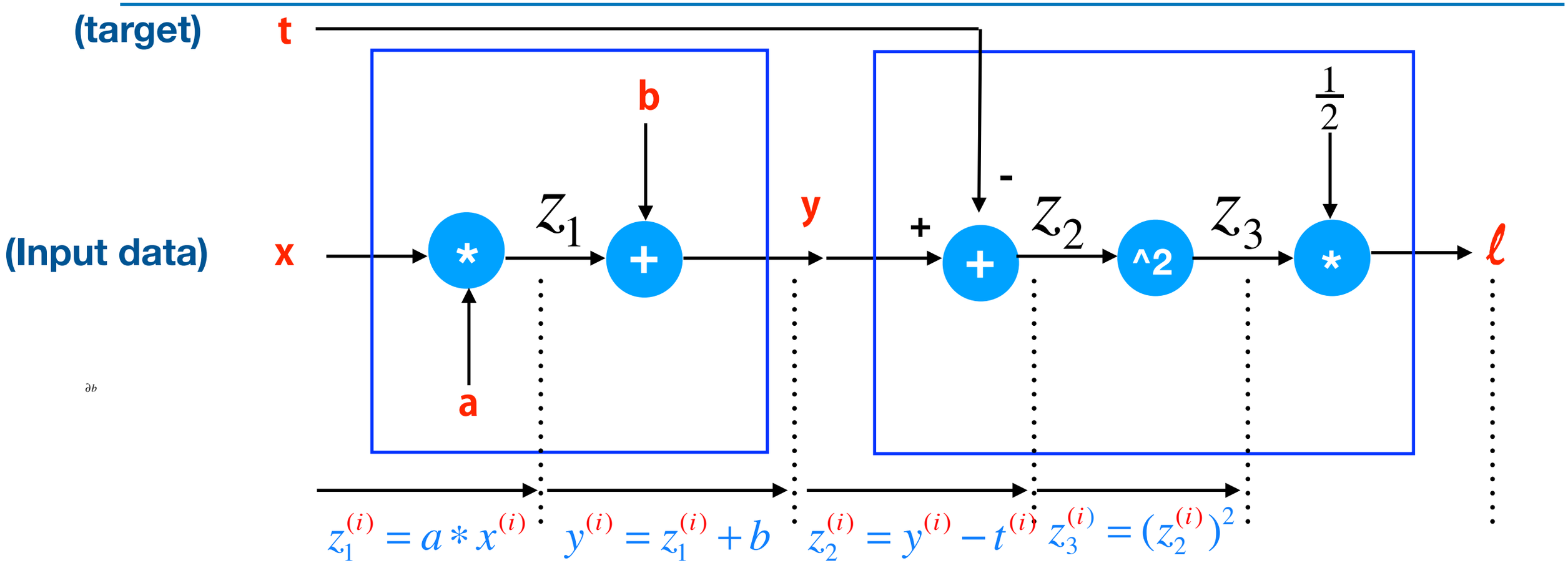
Back-propagation
Forward pass



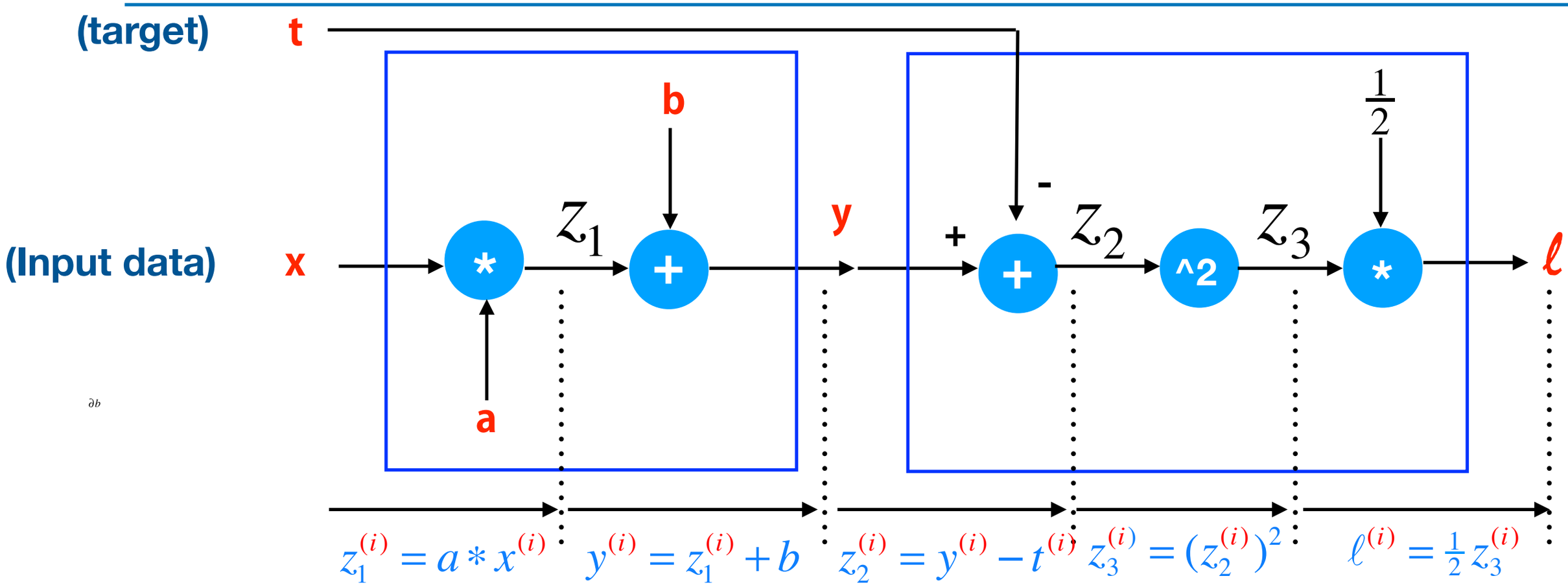
Back-propagation
Loss computation



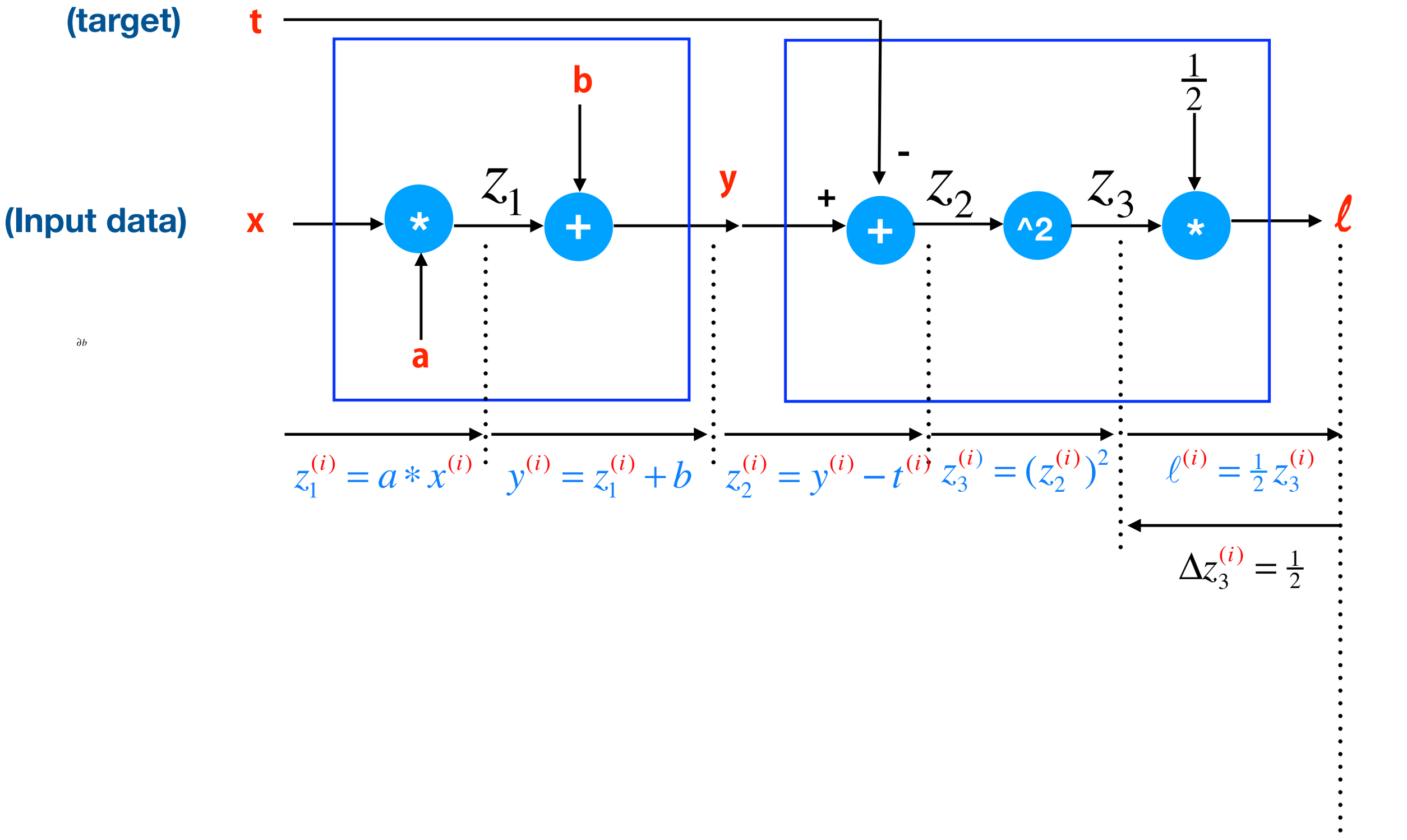
Back-propagation
Loss computation



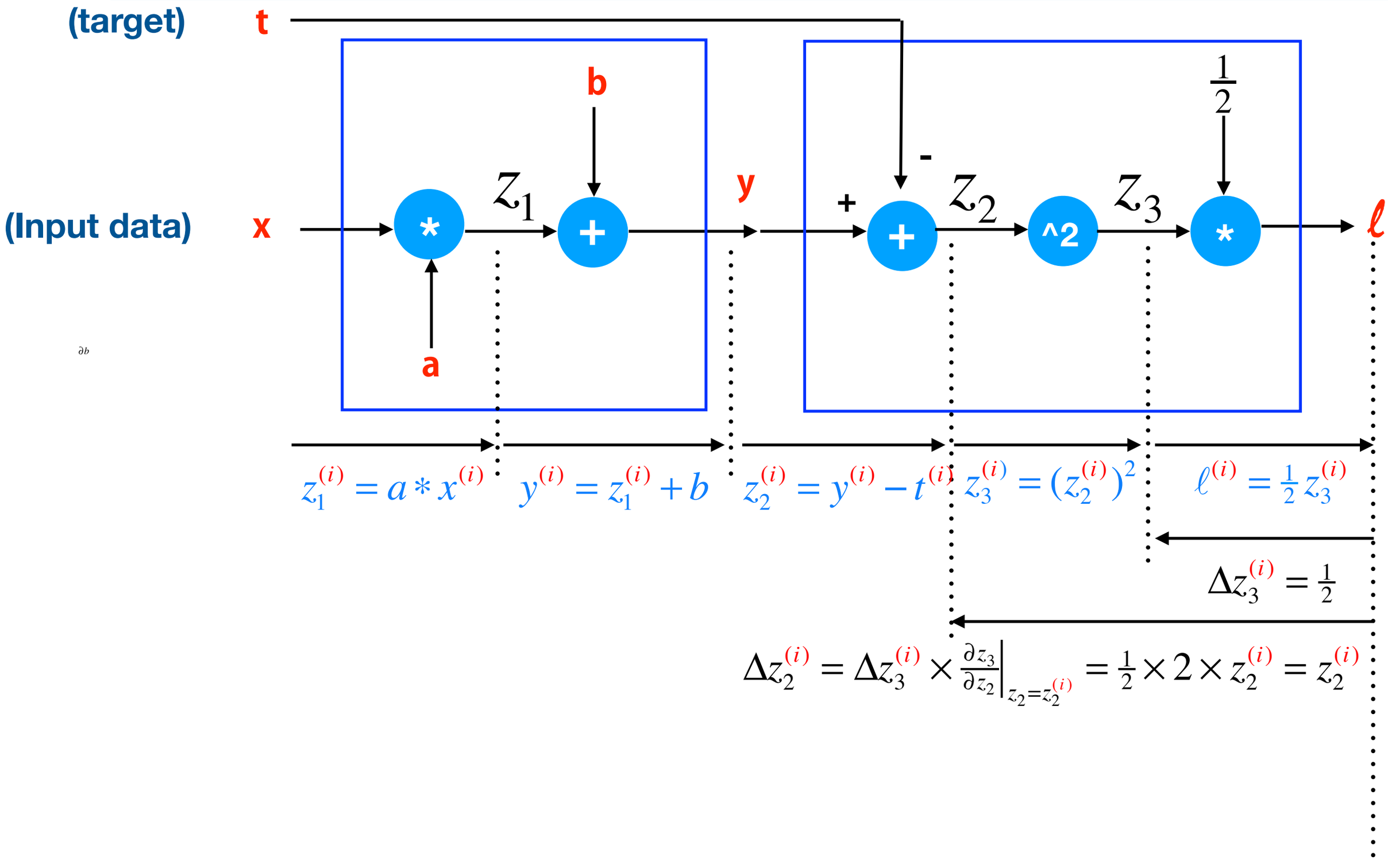
Back-propagation
Loss computation



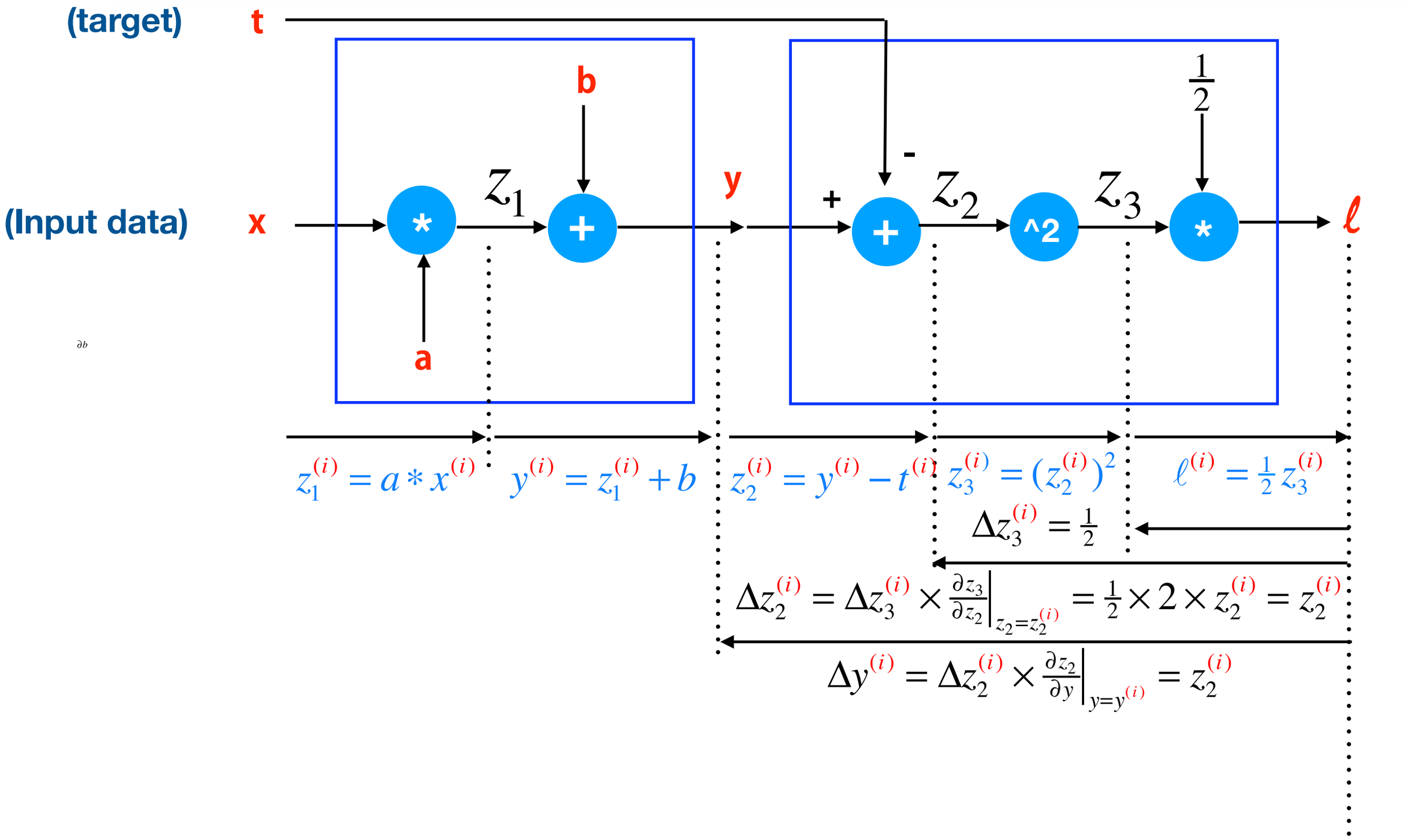
Back-propagation
Backward pass



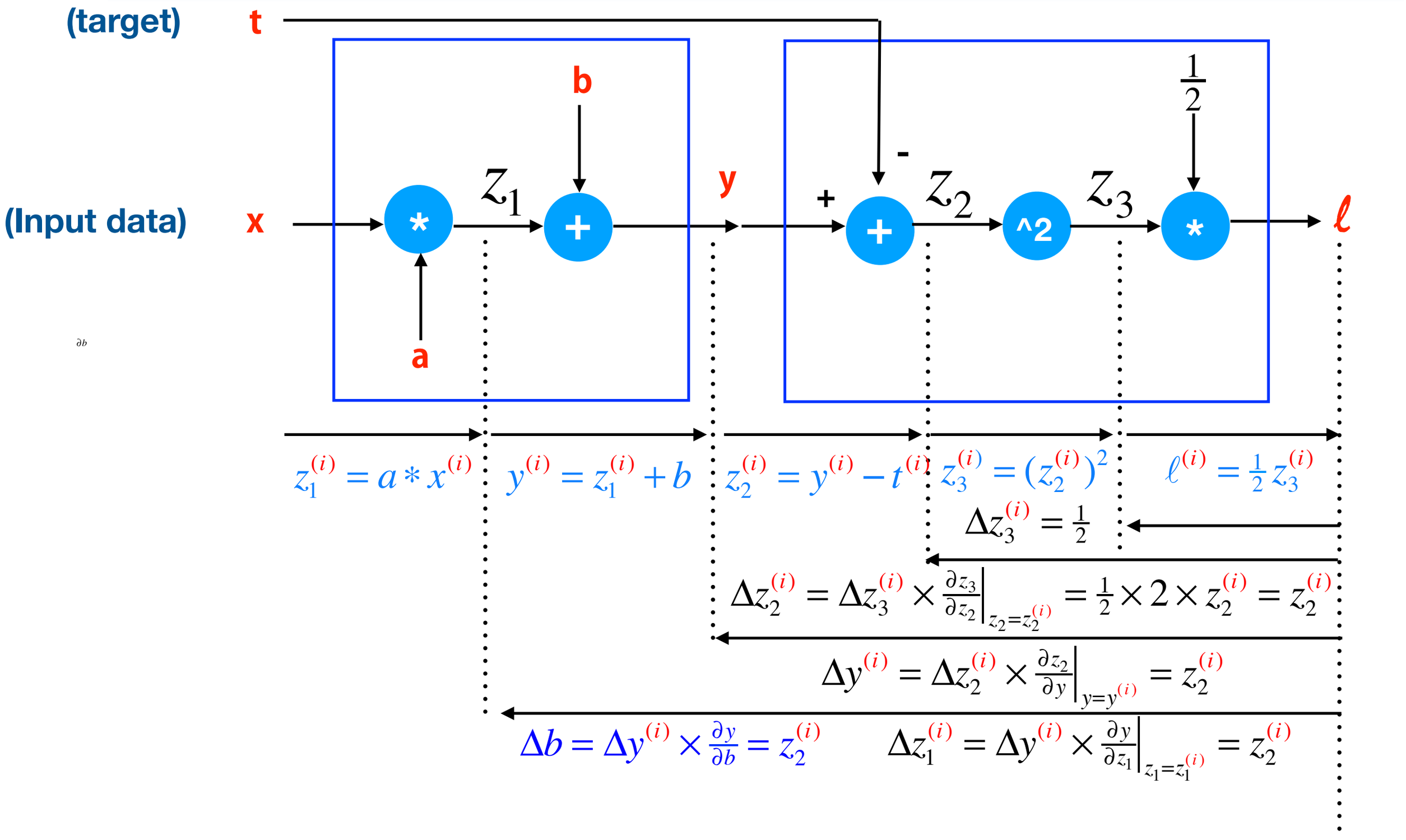
Back-propagation
Backward pass



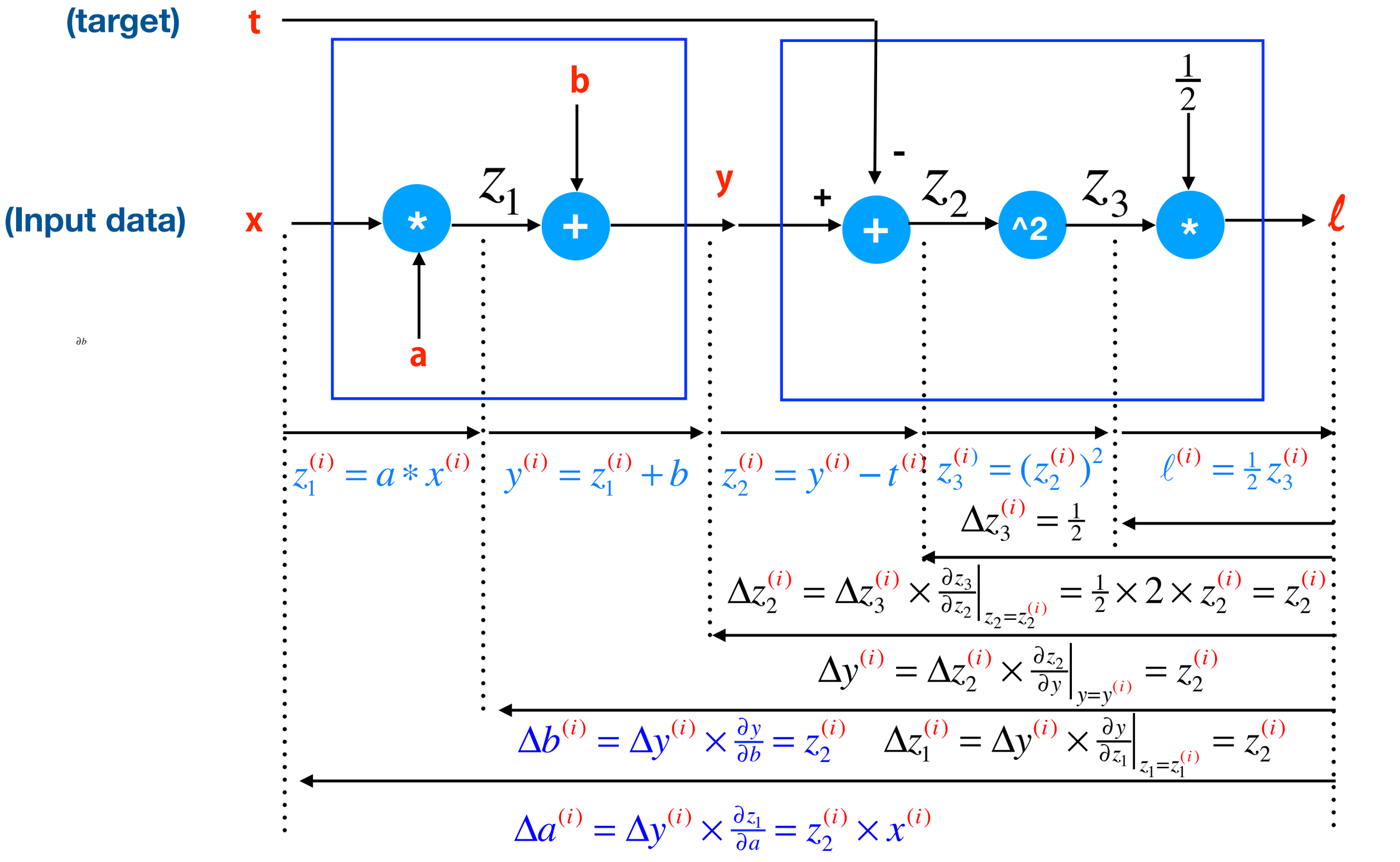
Back-propagation
Backward pass



Back-propagation
Backward pass



Back-propagation
Backward pass



Back-propagation
Backward pass

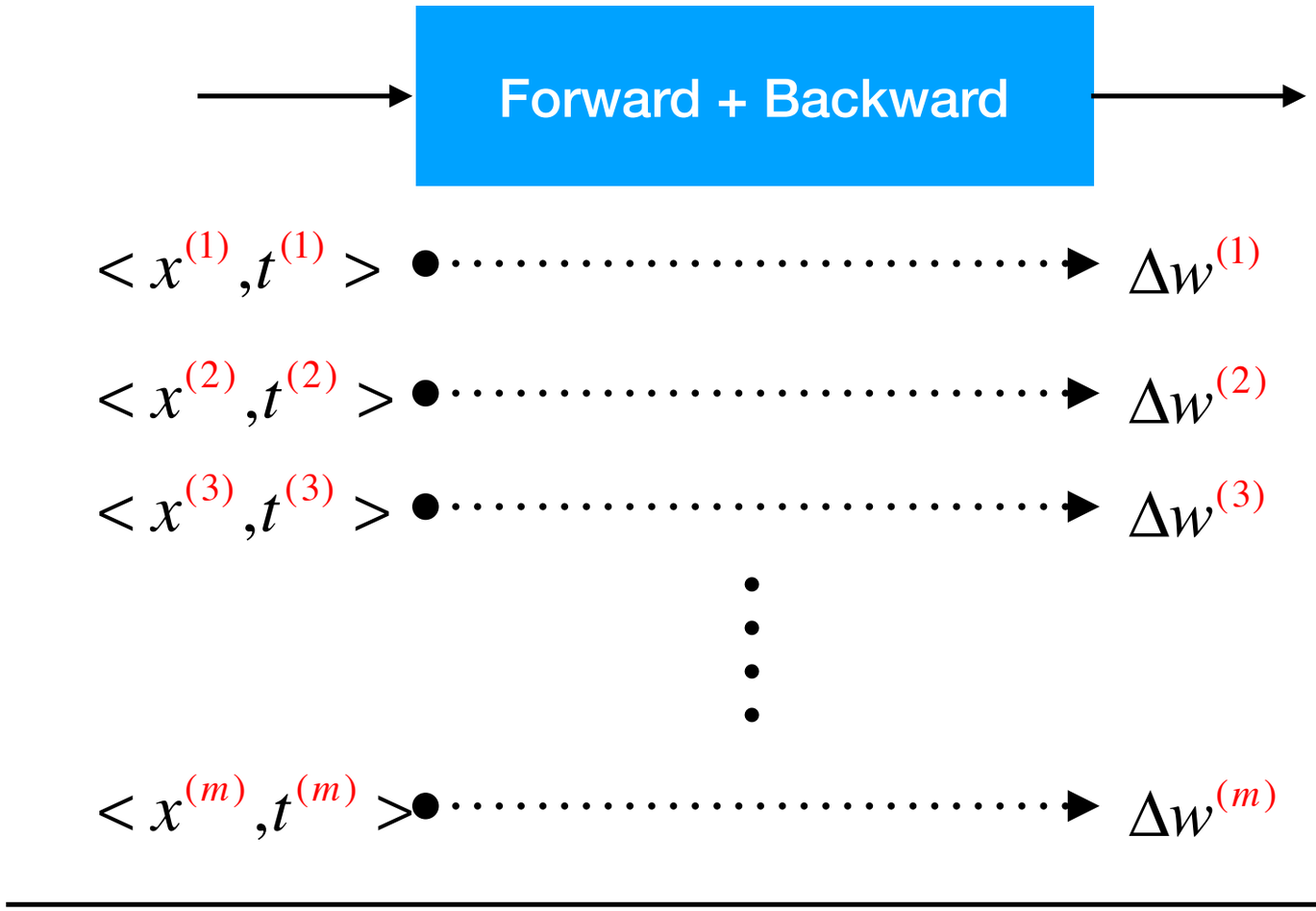
$$w = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\Delta w^{(i)} = \begin{bmatrix} \Delta a^{(i)} \\ \Delta b^{(i)} \end{bmatrix}$$

∂b

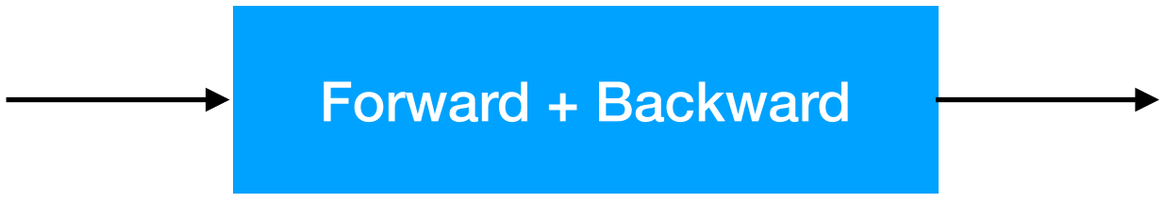
α : learning rate

Back-propagation
Backward pass



$$\Delta w = \frac{1}{m} \sum_{i=1}^{m} \Delta w^{(i)}$$

Back-propagation
Weight updating



∂b

$$w = w - \alpha \times \Delta w$$