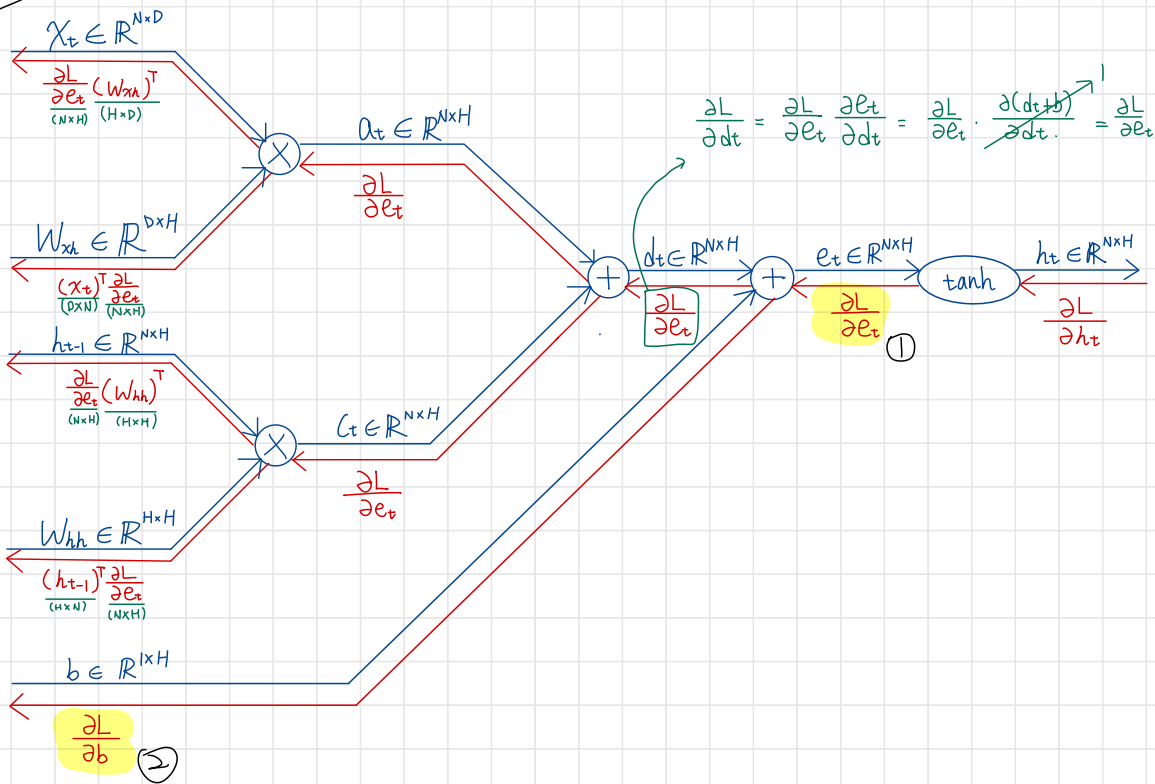


$$h_t = \tanh(W_{xh}x_t + W_{hh}h_{t-1} + b)$$



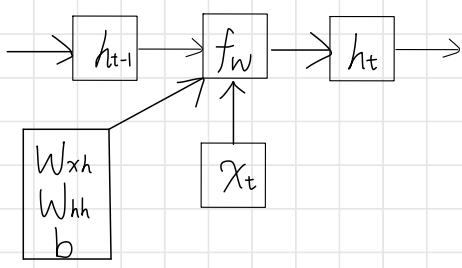
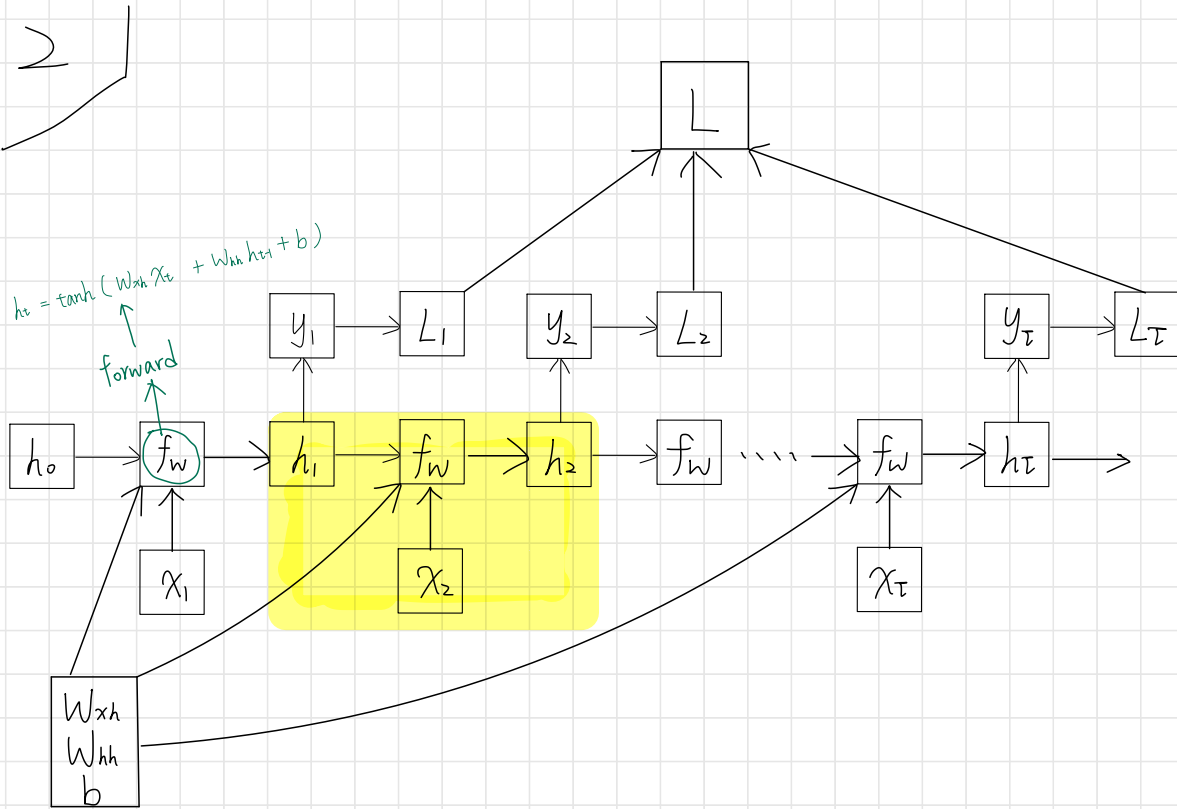
$$\textcircled{1} \quad f(x) = \tanh(x) \Rightarrow f'(x) = 1 - [f(x)]^2$$

$$h_t = \tanh(e_t) \Rightarrow \frac{\partial L}{\partial e_t} = \frac{\partial L}{\partial h_t} \frac{\partial h_t}{\partial e_t} = \frac{\partial L}{\partial h_t} [1 - h_t^2]$$

②

$$\begin{bmatrix} [e_t]_1 \\ [e_t]_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} [d_t]_1 \\ [d_t]_2 \\ \vdots \end{bmatrix} + \begin{bmatrix} b \\ b \\ \vdots \end{bmatrix}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial e_t} \frac{\partial e_t}{\partial b} = \sum_i^N \left[\frac{\partial L}{\partial e_t} \right]_i$$



3

