Exercise LSTM

Q f_1 ? c_1 ? h_1 ?

Solution $\begin{bmatrix} i \\ f \\ o \\ g \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} W \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix}$

$$C_t = f \odot C_{t-1} + i \odot g$$

$$h_t = O \odot \tanh(C_t)$$

(i) $f_1 = \sigma \, W_f \begin{bmatrix} h_0 \\ x_1 \end{bmatrix}$

$= \sigma \, [W_{hf} \ W_{xf}] \begin{bmatrix} h_0 \\ x_1 \end{bmatrix}$

$f_1 = \sigma \begin{bmatrix} 0.1 & 0.2 & 0.5 & 0.6 \\ 0.3 & 0.4 & 0.7 & 0.8 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}$

$= \sigma \begin{bmatrix} 1.6 \\ 2.2 \end{bmatrix}$

$1 + 0.6 = 1.6$

$1.4 + 0.8 = 2.2$

$\sigma(1.6) = \dfrac{1}{1+e^{-1.6}} = 0.8320$

$\sigma(2.2) = \dfrac{1}{1+e^{-2.2}} = 0.9002.$

Comment：保留了大部分短时记忆和输入
stay the most parts of short-term
memory and input data

$f_1 = \begin{bmatrix} 0.8320 \\ 0.9002 \end{bmatrix}$

(ii) $C_t = f \odot C_{t-1} + i \odot g$

$C_1 = f_1 \odot C_0 + i_1 \odot g_1$

$$= \begin{bmatrix} 0.8320 \\ 0.9002 \end{bmatrix} \odot \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} + \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix} \odot \begin{bmatrix} 0.5 \\ 0.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.08320 \\ 0.18004 \end{bmatrix} + \begin{bmatrix} 0.15 \\ 0.24 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2332 \\ 0.4200 \end{bmatrix}$$

(iii) $h_t = o \odot \tanh C_t$

$h_1 = o_1 \odot \tanh C_1$

$$= \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \odot \tanh \begin{bmatrix} 0.2332 \\ 0.4200 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \odot \begin{bmatrix} 0.2291 \\ 0.3969 \end{bmatrix}$$

$$= \begin{bmatrix} 0.09164 \\ 0.23814 \end{bmatrix}$$