TD3 - RNN

Ex 1 : Elman, 1990

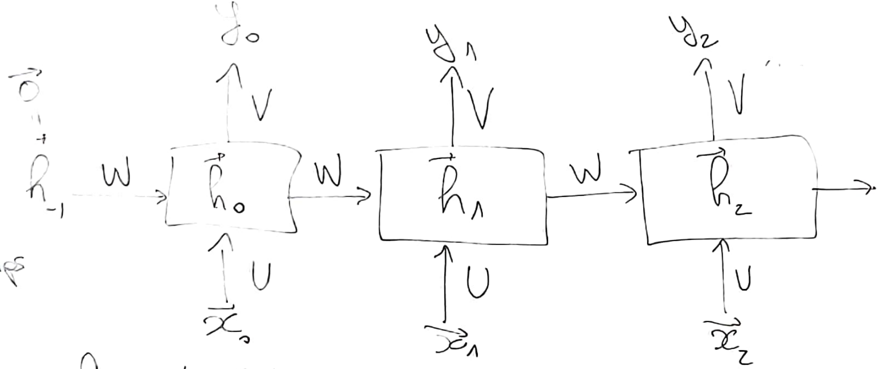
Régi par les équations suivantes :

état caché $h_t = \tanh(U x_t + W h_{t-1})$

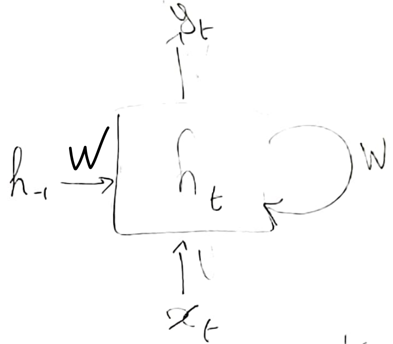
Sortie $y_t = \text{Softmax}(V h_t)$

avec U, V, W des matrices entraînables.

1) Dessiner une représentation "déroulée" de ce réseau pour une séquence de taille 3 $\{\vec{x}_0, \vec{x}_1, \vec{x}_2\}$

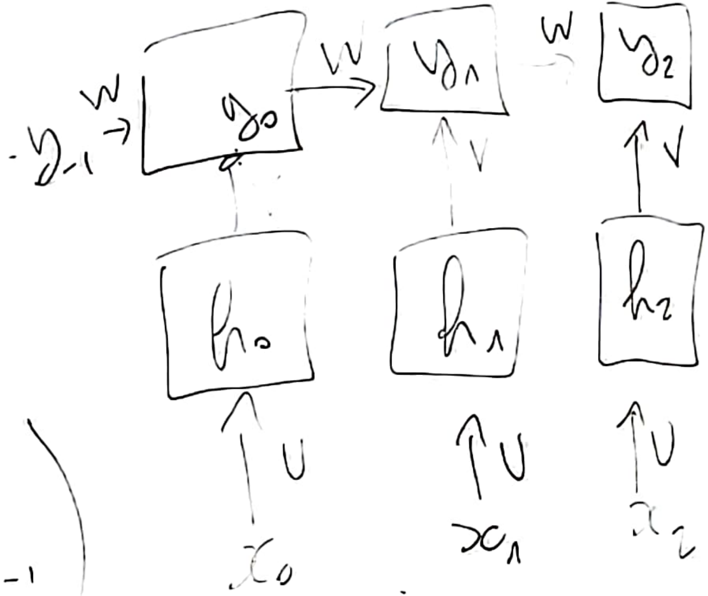


t. temps



$t = 0 \quad h_0 = \tanh(U x_0 + W h_{-1})$

non-déroulée.



$$h_t = \tanh(Ux_t)$$

$$y_t = \mathrm{softmax}(Vh_t + Wy_{t-1})$$

2) Donner l'expression de y_2 en fonction de x_0, x_1, x_2

On suppose que $h_{-1} = 0$.

$$y_2 = \mathrm{Softmax}\Big(V\Big(\tanh\big(U x_2 + W\big(\tanh\big(U x_1 + W \tanh(U \cdot x_0)\big)\big)\big)\Big)\Big)$$

$$\begin{tikzcd} \overset{w_1}{\overset{\shortparallel}{x_0}} \arrow[r] & \boxed{\text{dot}} \arrow[r, "w_2"] & \boxed{\tanh} \arrow[r] & \overset{w_3}{\overset{\shortparallel}{h_0}} \\ \overset{w_0}{\overset{\shortparallel}{U}} \arrow[ur] & & & \end{tikzcd}$$

$$\frac{\partial h_1}{\partial U} = ?$$

$$h_0 = \tanh(U x_0)$$

"forward"

$$w_0 = U; \quad w_1 = x_0$$

$$w_2 = w_0 w_1$$

$$w_3 = \tanh(w_2)$$

$$\frac{\partial h_0}{\partial U} = \boxed{(1 - h_0^2) x_0 = \text{"}\bar{U}\text{"}}$$

"reverse"

$$\bar{w}_0 = \bar{w}_2 \frac{\partial w_2}{\partial w_0} = (1 - h_0^2) \times w_1 = (1 - h_0^2) x_0$$

$$\bar{w}_2 = \frac{d w_3}{d w_2} = \bar{w}_3 \frac{d w_3}{d w_2} = 1 \times (1 - \tanh^2(w_2)) = 1 - w_3^2 = 1 - h_0^2$$

$$\bar{w}_3 = \frac{d w_3}{d w_3} = 1 \quad \text{"seed"}$$

$$\begin{tikzcd}
W_1 = x_0 \arrow[r] & \mathrm{dot} \arrow[r, "W_3"] & \mathrm{tanh} \arrow[r] & h_0 = W_5 \arrow[r] & \mathrm{dot} \arrow[d, "W_6"] & & W_8 = h_1 \\
W_0 = U \arrow[ur, "W_{03}"] \arrow[r, "W_{04}"] & \mathrm{dot} \arrow[rrr, "W_4"] & & & + \arrow[r, "W_7"] & \mathrm{tanh} \arrow[ur] & \\
W_2 = x_1 \arrow[ur] & & & & & &
\end{tikzcd}$$

W

$$W_0 = U;\ W_1 = x_0;\ W_2 = x_1$$

$$W_3 = W_0 W_1$$

$$W_4 = W_0 W_2$$

$$W_5 = \tanh(W_3)$$

$$W_6 = W\, W_5$$

$$W_7 = W_6 + W_4$$

$$W_8 = \tanh(W_7)$$

$$\bar{W}_{04} = \bar{W}_4 \times \frac{\partial W_4}{\partial W_{04}} = (1-h_1^2) \times W_2 = (1-h_1^2) \times x_1$$

$$\bar{W}_{03} = \bar{W}_3 \times \frac{\partial W_3}{\partial W_{03}} = (1-h_1^2)\, W\, (1-h_0^2) \times W_1 = (1-h_1^2)\, W\, (1-h_0^2)\, x_0$$

$$\bar{W}_4 = \bar{W}_7 \frac{\partial W_7}{\partial W_4} = \bar{W}_7 \times 1 = 1-h_1^2$$

$$\bar{W}_3 = \bar{W}_5 \times \frac{\partial W_5}{\partial W_3} = \bar{W}_5 \, (1-\underset{h_0^2}{\underset{\shortparallel}{W_5^2}}) = (1-h_1^2)\, W\, (1-h_0^2)$$

$$\bar{W}_5 = \bar{W}_6 \times \frac{\partial W_6}{\partial W_5} = (1-h_1^2)\, W$$

$$\bar{W}_6 = \bar{W}_7 \times 1 = \bar{W}_7 = 1-h_1^2$$

$$\bar{W}_7 = \bar{W}_8 \times \frac{\partial W_8}{\partial W_7} = 1 \times (1-W_8^2) = 1-h_1^2$$

$$\bar{W}_8 = 1$$

$$\bar{W}_0 = \bar{W}_{03} + \bar{W}_{04}$$

$$= (1-h_1^2)\, W\, (1-h_0^2)\, x_0 + (1-h_1^2)\, x_1$$

$$dV = (1-h_1^2)\left(x_1 + W(1-h_0^2)\, x_0\right)$$

$$\frac{\partial h_0}{\partial U} = (1 - h_0^2) x_0$$

$$\frac{\partial h_1}{\partial U} = (1 - h_1^2)\left(x_1 + W(1 - h_0^2) x_0\right)$$

$$\frac{\partial h_2}{\partial U} = (1 - h_2^2)\left(x_2 + W \frac{\partial h_1}{\partial U}\right)$$

$$= (1 - h_2^2)\left(x_2 + W\left[(1 - h_1^2)\left(x_1 + W (1 - h_0^2) x_0\right)\right]\right)$$

$$\frac{\partial h_0}{\partial U} = (1 - h_0^2)\, x_0$$

nn.RNN()

$$h_t = \tanh(x_t W_{ih}^T + b_{ih} + h_{t-1} W_{hh}^T + b_{hh})$$

U $\qquad$ W

∂U

$$= (1 - h_2^2)\left(x_2 + W\left[(1 - h_1^2)\left(x_1 + W (1 - h_0^2) x_0\right)\right]\right)$$