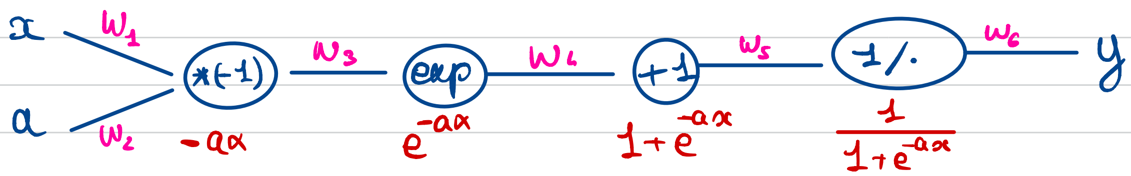
Apprentissage automatique 1 - TD n°1

Exercice 1 : fonction logistique-sigmoide

1. Dessiner le graphe de calcul de la fonction sigmoide : $y = f(x, a) = 1/(1 + e^{-ax})$
2. Calculer sa dérivée par rapport à x, appelée "dx", en utilisant le graphe et le mode de différentiation automatique "forward"
3. Même chose mais en utilisant le mode "reverse"
4. Sans faire de nouveau calculs, donner "da"
5. Si l'on considère une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, dans quel cas le mode forward est plus efficace que le mode reverse ?

1)



2) passe forward :

$$w_1 = x, \quad w_2 = a$$
$$w_3 = -(w_1 * w_2)$$
$$w_4 = e^{w_3}$$
$$w_5 = 1 + w_4$$
$$w_6 = \frac{1}{w_5}$$

$$\dot{w}_1 = \frac{\partial w_1}{\partial x} = 1 \; ; \quad \dot{w}_2 = 0$$
$$\dot{w}_3 = -a$$
$$\dot{w}_4 = \dot{w}_3 \times \exp(w_3) = -a\, e^{-ax}$$
$$\dot{w}_5 = -a\, e^{-ax}$$
$$\dot{w}_6 = \frac{-\dot{w}_5}{w_5^2} = \frac{a\, e^{-ax}}{(1+e^{-ax})^2} \quad \text{"dx"}$$

3) mode "reverse" = backward :

$$\bar{w}_1 = \frac{dy}{\partial w_1} = \frac{dy}{\partial w_3} \times \frac{\partial w_3}{\partial w_1} = \frac{-\exp(w_3)}{w_5^2} \times (-w_2) = \frac{a\, e^{-ax}}{(1+e^{-ax})^2} \quad \text{"dx"}$$
$$\bar{w}_3 = \frac{-1}{w_5^2} \times \exp(w_3)$$
$$\bar{w}_4 = \frac{dy}{\partial w_4} = \frac{dy}{\partial w_5} \times \frac{\partial w_5}{\partial w_4} = \frac{-1}{w_5^2} \times w_4$$
$$\bar{w}_5 = \frac{dy}{\partial w_5} = -1/w_5^2 = \frac{-1}{(1+e^{-ax})^2}$$
$$\bar{w}_6 = 1$$

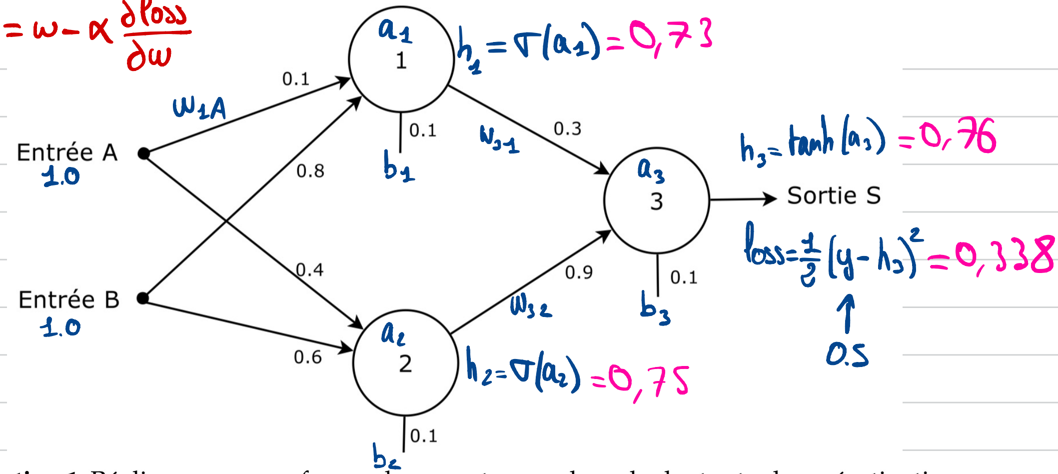
4) $\text{"da"} = \dfrac{x\, e^{-xa}}{(1+e^{-xa})^2}$ car on peut inverser x et a puis la fonction est symétrique

5) Lorsque $n > m$ on fait la backward.

Lorsque $m > n$ on fait la forward.

actualisation des poids ← $w = w - \alpha \frac{\partial loss}{\partial w}$

Exercice 2 : Perceptron multicouche (MLP) pour la régression



Question 1. Réaliser une passe forward pour cet exemple : calculer toutes les préactivations et activations. Quelle est la valeur $\hat{y}$ prédite par le réseau ? Calculer la valeur de la fonction de coût associée.

$\frac{\partial loss}{\partial\, poids} =$ **Question 2.** Réaliser une passe backward pour cet exemple : calculer tous les gradients des poids, ainsi que les variations de valeurs associées : ΔW.

Question 3. Actualiser les poids. Dessiner le réseau en reportant les valeurs des poids que vous avez trouvées. $= w = w - \alpha \frac{\partial loss}{\partial w}$

Question 4. Refaire une passe forward. La prédiction est-elle meilleure cette fois-ci ?

oui on trouve 0,011 pour la loss

1) passe forward : $\vec{AB} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}$; $\vec{W} = \begin{bmatrix} 0.1 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}$; $\vec{W_3} = \begin{bmatrix} 0.3 \\ 0.9 \end{bmatrix}$; $\vec{b} = \begin{pmatrix} 0,1 \\ 0,1 \end{pmatrix}$

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \vec{AB} \times \vec{W} + \vec{b} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix} \times \begin{bmatrix} 0.1 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} = \begin{pmatrix} 0.1 + 0.8 + 0.1 \\ 0.4 + 0.6 + 0.1 \end{pmatrix} = \begin{bmatrix} 1.0 \\ 1.1 \end{bmatrix}$$

$$a_3 = 0,3 \times h_1 + 0,9 \times h_2 + b_3$$

$$= 0,3 \times \frac{1}{1 + e^{-a_1}} + 0,9 \times \frac{1}{1 + e^{-a_2}} + 0,1$$

$$= 0,3 \times 0,73 + 0,9 \times 0,75 + 0,1 = 0,99$$

$$\hat{y} = h_3 = \tanh(a_3) = 0,76$$

$$\text{coût} = loss = \frac{1}{2}(0,5 - h_3)^2 = 0,0338$$

2) 3) Actualisation des poids : $w = w - \alpha \frac{\partial loss}{\partial w}$

$$\frac{\partial loss}{\partial w_{31}} = \frac{\partial loss}{\partial h_3} \times \frac{\partial h_3}{\partial a_3} \times \frac{\partial a_3}{\partial w_{31}} = (-0,5 + h_3) \times \left(1 - \left(\frac{e^{2a_3} - 1}{e^{2a_3} + 1}\right)^2\right) \times h_1 \qquad 2)$$

$$\Rightarrow w_{31} = w_{31} - (-0,5 + h_3) \times (1 - \tanh(a_3)^2) \times h_1 = 0,219 \qquad 3)$$

$$\frac{\partial loss}{\partial w_{32}} = \frac{\partial loss}{\partial h_3} \times \frac{\partial h_3}{\partial a_3} \times \frac{\partial a_3}{\partial w_{32}} = (-0,5 + h_3) \times (1 - \tanh(a_3)^2) \times h_2 \qquad 2)$$

$$\Rightarrow w_{32} = w_{32} - (-0,5 + h_3) \times (1 - h_3^2) \times h_2 = 0,82 \qquad 3)$$

$$\frac{d\,loss}{d w_{1A}} = \overbrace{\frac{d\,loss}{d h_3} \times \frac{d h_3}{d a_3}}^{\text{déjà fait}} \times \frac{d a_3}{d h_1} \times \frac{d h_1}{d a_1} \times \frac{d a_1}{d w_{1A}} = (-0{,}5 + h_3) \times (1 - h_3^2) \times W_{31} \times h_1 (1 - h_1)$$

$$\Rightarrow w_{2A} = w_{1A} - \frac{d\,loss}{d w_{1A}} = 0{,}1 - 0{,}0065 = 0{,}0935$$

(. . .)

dernier loss : 0,011