OUSSAMA KHALOUI
TIMOTHÉE BLANCHY

TP 2

Question 1 :

on a $\vec{w} = \arg\min_{\vec{w}} E_D(\vec{w})$

Les meilleurs paramètres sont ceux qui correspondent à $\nabla_{\vec{w}} E_D(\vec{w}) = 0$

$$\text{On a } \nabla_{\vec{w}} E_D(\vec{w}) = \nabla_{\vec{w}}\left(\left(T - W^T \Phi\right)^2 + \lambda W^T W\right)$$

$$= \nabla_{\vec{w}}\left(\left(T - W^T \Phi\right)^2\right) + \lambda \nabla_{\vec{w}}\left(W^T W\right)$$

$$= 2(T - W^T\phi)\nabla_{\vec{w}}(T - W^T\Phi) + \lambda \nabla_{\vec{w}}(W^T W)$$

on sait que T est indépendant de $\vec{w}$ donc $\nabla_{\vec{w}}(T) = 0$

$$\nabla_{\vec{w}} E_D(\vec{w}) = -2(T - W^T\phi)\nabla_{\vec{w}}(W^T\Phi) + \lambda \nabla_{\vec{w}}(\|W\|^2)$$

(on a $\nabla_{\vec{w}}(W^T\phi) = \phi^T$)

$$\nabla_{\vec{w}} E_D(\vec{w}) = 2(W^T\phi - T)\phi^T + 2\lambda W^T$$

si on force $\nabla_{\vec{w}} E_D(\vec{w}) = 0$

on a $W^T\phi.\phi^T - T\phi^T + \lambda W^T = 0$

$$W^T(\phi.\phi^T + \lambda I) = T\phi^T$$

$$W^T = T\phi^T(\phi\phi^T + \lambda I)^{-1}$$

Question 2 :

Soit E la fonction de perte

Supposons que E est une entropie croisée

$$E(\vec{w}) = \sum_{i=1}^{n} t_i \ln(y_i) + (1-t_i)\ln(1-y_i)$$

$$\vec{\nabla} E(\vec{w}) = \sum_{i=1}^{n} t_i \vec{\nabla}\left(\ln(y_i)\right) + (1-t_i)\vec{\nabla}\left(\ln(1-y_i)\right)$$

$$= \sum_{i=1}^{n} t_i \frac{\vec{\nabla}\left(\sigma(w^T\phi_i)\right)}{y_i} + (1-t_i)\frac{\vec{\nabla}\left(1-\sigma(w^T\phi_i)\right)}{1-y_i}$$

on a $\vec{\nabla}\left(\sigma(w^T\phi_i)\right) = \vec{\nabla}\left(\frac{1}{1+\exp(-w^T\phi_i)}\right)$

$$= -\frac{\vec{\nabla}\left(1+\exp(-w^T\phi_i)\right)}{\left(1+\exp(-w^T\phi_i)\right)^2}$$

$$= +\frac{\vec{\nabla}(w^T\phi_i)\exp(-w^T\phi_i)}{\left(1+\exp(-w^T\phi_i)\right)^2}$$

$$= \phi_i \cdot \sigma(w^T\phi_i)\left(1-\sigma(w^T\phi_i)\right)$$

$$= \phi_i \cdot y_i(1-y_i)$$

$$\vec{\nabla} E(\vec{w}) = \sum_{i=1}^{n} t_i \frac{\phi_i \cdot \cancel{y_i}(1-y_i)}{\cancel{y_i}} - (1-t_i)\frac{\phi_i \cdot y_i\cancel{(1-y_i)}}{\cancel{1-y_i}}$$

$$= \sum_{i=1}^{n} t_i \phi_i (1-y_i) - (1-t_i)\phi_i y_i$$

$$= \sum_{i=1}^{n} \phi_i\left(t_i - \cancel{t_i y_i} - y_i + \cancel{t_i y_i}\right)$$

$$\vec{\nabla} E(\vec{w}) = \sum_{i=1}^{n} \phi_i (t_i - y_i)$$

Question 3 : Définissons d'abord la fonction d'entropie :

$$E = -p_1 \log p_1 - p_2 \log p_2 - p_3 \log p_3$$

Nous avons 2 contraintes à respecter :

$$\begin{cases} p_1 + p_2 + p_3 = 1 \\ p_1 = 2 p_2 \end{cases}$$

La fonction lagrangienne est définie comme suit :

$$L = E + \lambda(p_1 + p_2 + p_3 - 1) + \mu(p_1 - 2p_2)$$

Les probabilités qui maximisent l'entropie vérifient :

$$\begin{cases} \dfrac{\partial L}{\partial p_1} = -\log p_1 - 1 + \lambda + \mu = 0 & (1) \\ \dfrac{\partial L}{\partial p_2} = -\log p_2 - 1 + \lambda - 2\mu = 0 & (2) \\ \dfrac{\partial L}{\partial p_3} = -\log p_3 - 1 + \lambda = 0 & (3) \\ \dfrac{\partial L}{\partial \lambda} = p_1 + p_2 + p_3 - 1 = 0 & (4) \\ \dfrac{\partial L}{\partial \mu} = p_1 - 2 p_2 = 0 & (5) \end{cases}$$

$$(1)-(2) \Rightarrow -\log p_1 + \log p_2 + \mu + 2\mu = 0$$

$$\Rightarrow -\log(2p_2) + \log p_2 + 3\mu = 0 \quad \text{(en utilisant (5))}$$

$$\Rightarrow -\log 2 - \cancel{\log p_2} + \cancel{\log p_2} + 3\mu = 0$$

$$\Rightarrow \mu = \frac{1}{3}$$

Les équations deviennent :

$$\begin{cases} \dfrac{\partial L}{\partial p_1} = -\log p_1 - \dfrac{2}{3} + \lambda = 0 & (1) \\ \dfrac{\partial L}{\partial p_2} = -\log p_2 - \dfrac{5}{3} + \lambda = 0 & (2) \\ \dfrac{\partial L}{\partial p_3} = -\log p_3 - 1 + \lambda = 0 & (3) \\ \dfrac{\partial L}{\partial \lambda} = p_1 + p_2 + p_3 - 1 = 0 & (4) \\ \dfrac{\partial L}{\partial} = p_1 - 2p_2 = 0 & (5) \end{cases}$$

$$(3) \Rightarrow -\log p_3 - 1 + \lambda = 0$$

$$\Rightarrow -\log(1 - p_2 - p_1) - 1 + \lambda = 0$$

$$\Rightarrow -\log(1 - 3p_2) - 1 + \lambda = 0$$

$$\Rightarrow \log(1 - 3p_2) = \lambda - 1$$

$$\Rightarrow \quad 1 - 3p_2 = 2^{\lambda - 1}$$

$$\Rightarrow \quad 3p_2 = 1 - 2^{\lambda - 1}$$

$$\Rightarrow \quad p_2 = \frac{1 - 2^{\lambda - 1}}{3} \quad (*)$$

$$(2) \quad \Rightarrow \quad \log p_2 = \lambda - \frac{5}{3}$$

$$\Rightarrow \quad p_2 = 2^{\lambda - \frac{5}{3}} \quad (**)$$

donc $2^{\lambda - \frac{5}{3}} = \dfrac{1 - 2^{\lambda - 1}}{3}$ (d'après $(*)$ et $(**)$)

d'où $3 \cdot \dfrac{2^\lambda}{2^{5/3}} + \dfrac{2^\lambda}{2} = 1$

$$2^\lambda \left(\frac{3}{2^{5/3}} + \frac{1}{2} \right) = 1$$

$$2^\lambda = \frac{1}{\frac{3}{2^{5/3}} + \frac{1}{2}}$$

$$p_1 = 2^{-\frac{2}{3} + \lambda}$$

$$= 2^{-\frac{2}{3}} \cdot 2^\lambda$$

$$= \frac{2^{-\frac{2}{3}}}{\frac{3}{2^{5/3}} + \frac{1}{2}}$$

$$= \frac{1}{\frac{3}{2} + \frac{1}{2^{1/3}}}$$

$$p_1 = \frac{2}{3 + 2^{2/3}}$$

$$p_2 = \frac{1}{3 + 2^{2/3}} \qquad (\text{car } p_1 = 2p_2)$$

$$p_3 = 1 - 3p_2 = 1 - \frac{3}{3 + 2^{2/3}}$$

$$p_3 = \frac{2^{2/3}}{3 + 2^{2/3}}$$