TP 2 :

Régression Logistique

Q0. $e^t > 0 \quad \forall t \quad \Rightarrow \quad e^t/(1+e^t) \in \,]0,1[$

$\lim_{t \to -\infty} \sigma(t) = 0/1 = 0$

Si $\gamma = 0$: $\quad F(w) = \sum_{a \in A} f_a(w)$

Minimiser F revient à minimiser la différence entre y_a et $\sigma(\langle w, x_a \rangle)$ i.e. faire en sorte que $\sigma(\langle w, x_a \rangle) \simeq y_a$.

1. Étude théorique, $\gamma > 0$.

1.1. Existence et Unicité

Q1. Puisque $\sigma(\cdot) \in (0,1)$ $\quad \begin{cases} \sigma(\langle w, x_a \rangle) \in (0,1) \\ 1 - \sigma(\langle w, x_a \rangle) \in (0,1) \end{cases}$

En prenant le négatif du log, $f_a \geq 0$.

$F(w) \geq \underbrace{F_A(w)}_{\geq 0} + \frac{\gamma}{2}\|w\|^2 \geq \frac{\gamma}{2}\|w\|^2$

Q2. [Existence et Unicité]

1. $g(\langle \lambda w_1 + (1-\lambda) w_2 , x \rangle) = g(\lambda \langle w_1 , x\rangle + (1-\lambda)\langle w_2 , x\rangle)$

$\leq \lambda g(\langle w_1 , x\rangle) + (1-\lambda)\, g(\langle w_2 , x\rangle)$

2. $\sigma'(t) = \dfrac{e^t(e^t+1) - e^{2t}}{(e^t+1)^2} = \dfrac{e^t}{(e^t+1)^2}$

$= \dfrac{e^t}{1+e^t}\left(1 - \dfrac{e^t}{e^t+1}\right)$

$= \sigma(t)\,(1-\sigma(t))$

$g'(t) = \dfrac{\sigma'(t)}{1-\sigma(t)} = \sigma(t)$

$g''(t) = \sigma'(t) \geq 0$ (since $\sigma(t) \in (0,1)$.)

Donc g est convexe et f_a est convexe

3. On a $\forall w \in \mathbb{R}^d$, $F(w) = \underbrace{\sum f_a(w)}_{\text{convexe (somme de fcts convexes)}} + \dfrac{\gamma}{2}\|w\|^2$

donc F est fortement convexe

$$D^2F(w) = \underbrace{\sum_{a\in A} D^2 f_a(w)}_{\geq 0} + \gamma I_d \succcurlyeq \gamma I_d.$$

(since f_a is convex and twice diff).

4. F admet au plus un minimiseur. Comme f admet un minimiseur alors F admet un unique minimiseur.

1.2- Calcul du gradient et de la hessienne.

Q3. if $a \in A_0$: $\nabla f_a(w) = \frac{\sigma(\langle w, x_a\rangle)(1-\cancel{\sigma(\langle w,x_a\rangle)})}{\cancel{1-\sigma(\langle w,x_a\rangle)}} x_a$

if $a \in A_1$: $\nabla f_a(w) = -(1-\sigma(w\cdot x_a))\, x_a$

$$= (\sigma(\langle w, x_a\rangle) - 1)\, x_a$$

ie $$\boxed{\nabla f_a(w) = (\sigma(w\cdot x_a) - y_a)\, x_a}$$

$$\nabla F(w) = \sum_{a\in A} (\sigma(\langle w | x_a\rangle) - y_a)\, x_a + \gamma w$$

Q4. 1. $x = [d, 1] \Rightarrow xx^T \in M_d(\mathbb{R})$

$$(xx^T)^T = (x^T)^T x^T = xx^T$$

$$(xx^T)y = \lambda y \Rightarrow y^T(xx^T)y = \lambda \|y\|^2$$

$$\Rightarrow \frac{\|x^T y\|^2}{\|y\|^2} = \lambda \geqslant 0$$

$$xx^T = x_i x_j$$

$$\langle Av | v \rangle = v^T x x^T v = \langle x, v \rangle^2.$$

2. $\dfrac{\partial G_{i}}{\partial w_j}(w) =$