2) $f(w) = \frac{1}{2n} \|y - Xw\|_2^2 + \frac{\lambda}{2} \|w\|_2^2 = \frac{1}{n}\sum_{i=1}^{n} f_i(w)$ où $f_i(w) = \frac{1}{2}\left[(y_i - x_i^T w)^2 + \lambda \|w\|^2\right]$

$\nabla f_i(w) = (y_i - x_i^T w)\, x_i + \lambda w$.

$$\nabla f(w) := \frac{1}{n}\sum_{i=1}^{n} \underbrace{x_i}_{\text{vect.}} \underbrace{(y_i - x_i^T w)}_{\text{scalaire.}} + \lambda w = \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T + \lambda I\right]}\, w + \text{cste } \%\, w.$$

$$\|\nabla f(w) - \nabla f(w')\| \leq \underbrace{\left|\left|\left| \overbrace{\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T + \lambda I}^{M} \right|\right|\right|}_{L} \|w - w'\|$$

si M : symmétrique réelle, $M = U^T D U$, $U^T U = I$.

$$\boxed{L = |||M|||} = \sup_u \frac{\|Mu\|}{\|u\|} = \sqrt{\sup_u \frac{u^T M^T M u}{u^T u}} \underset{\uparrow\, M \text{ symmétrique}}{=} \sup \{|\lambda| ;\ \lambda \in \mathrm{Spec}(M)\} = \sup_u \frac{|u^T M u|}{u^T u}.$$

si $x_i \in \mathbb{R}^d$, $|||\underbrace{x_i x_i^T}_{M_i}||| = \sup \{\lambda ;\ \lambda \in \mathrm{Spec}(M_i)\}$

$$\boxed{L = \left|\left|\left| \frac{1}{n}\sum x_i x_i^T + \lambda I \right|\right|\right|}$$
$$= \left|\left|\left| \frac{1}{n}\sum x_i x_i^T \right|\right|\right| + \lambda.$$

si $v \in \{x_i\}^\perp$, $(x_i x_i^T) v = 0$ donc : 0 vp. d'ordre $d-1$.

$(x_i x_i^T) x_i = \|x_i\|^2 x_i \Rightarrow \|x_i\|^2$ vp d'ordre 1.

Donc, $\boxed{|||x_i x_i^T||| = \|x_i\|^2.} \Rightarrow |||x_i x_i^T + \lambda I||| = \|x_i\|^2 + \lambda$.

$$L_{max} = \max[L_{\nabla f_1}, \ldots, L_{\nabla f_n}] = \max\left[\underbrace{|||x_1 x_1^T + \lambda I|||}, \ldots, |||x_n x_n^T + \lambda I|||\right]$$
$$(\|x_1\|^2 + \lambda, \ldots, \|x_n\|^2 + \lambda).$$

$(L_1, \ldots, L_d)$:

$$\frac{\partial f}{\partial w_j} = \frac{1}{n}\sum_{i=1}^{n} -x_i[j]\,(y_i - x_i^T w) + \lambda\, w[j].$$

$$\frac{\partial^2 f}{\partial w_j^2} = \boxed{\frac{1}{n}\sum_{i=1}^{n} x_i^2[j] + \lambda := L_j}$$

Régression Logistique :

$$l(y_i|x_i) = \begin{cases} \dfrac{e^{x_i^T w}}{1 + e^{x_i^T w}} = \dfrac{1}{1 + e^{-x_i^T w}} & \text{si } y_i = 1 \\ \dfrac{1}{1 + e^{x_i^T w}} & \text{si } y_i = -1 \end{cases}$$
$$= \frac{1}{1 + e^{-y_i x_i^T w}}.$$

Loss : $f(w) := -\frac{1}{n}\sum_{i=1}^{n} \log l(y_i|x_i) + \frac{\lambda}{2}\|w\|^2$.

$$f(w) := \frac{1}{n}\sum_{i=1}^{n} \underbrace{\log\left(1 + e^{-y_i x_i^T w}\right)}_{\varphi(y_i x_i^T w).} + \frac{\lambda}{2}\|w\|^2.$$

où $\varphi(t) = \log(1 + e^{-t})$.

$\varphi'(t) = \dfrac{-e^{-t}}{1 + e^{-t}} = \dfrac{-1}{1 + e^t}$

$\varphi''(t) = \dfrac{e^t}{(1 + e^t)^2}$

$$\nabla f(w) = \frac{1}{n}\sum_{i=1}^{n} \varphi'(y_i x_i^T w)\, y_i x_i + \lambda w.$$

Smc: $\nabla f(w) = \frac{1}{n}\sum_{i=1}^{n} \frac{-1}{1+e^{y_i x_i^T w}}\, y_i x_i + \lambda w.$; $X = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix}$

$$= \frac{1}{n}[x_1,\dots,x_n]\left\{ \begin{bmatrix} -y_1 \\ \vdots \\ -y_n \end{bmatrix} \underset{\substack{\uparrow \\ \text{multip. terme} \\ \text{à terme.}}}{\odot} \left[-\frac{1}{1+e^{y \odot Xw}} \right] \right\} + \lambda w.$$

$$\nabla^2 f(w) = \frac{1}{n}\sum_{i=1}^{n} \underbrace{\frac{e^{y_i x_i^T w}}{(1+e^{y_i x_i^T w})^2}}_{\varphi(y_i x_i^T w)} \underset{1}{\cancel{y_i^2}}\, x_i x_i^T + \lambda I.$$

$$\frac{u}{(1+u)^2} = \overbrace{\left(\frac{u}{1+u}\right)}^{p} \overbrace{\left(\frac{1}{1+u}\right)}^{1-p} \le 1/4.$$

$$|||\nabla^2 f(w)||| = \sup_{\|u\|=1} u^T \nabla^2 f\, u = \sup_{\|u\|=1} \frac{1}{n}\sum_{i=1}^{n} \varphi(y_i x_i^T w)\, u^T x_i x_i^T u + \lambda \underbrace{u^T u}_{=1}.$$

$$\le \frac{1}{4} \sup_{\|u\|=1} u^T \left[\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \right] u + \lambda.$$

$$= \frac{1}{4}\, \left|\left|\left| \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \right|\right|\right| + \lambda.$$

$$L = \left|\left|\left| \frac{1}{4}\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \right) \right|\right|\right| + \lambda.$$

$$L_{max} = \max\left(\frac{1}{4}\|x_1\|^2 + \lambda, \dots, \frac{1}{4}\|x_n\|^2 + \lambda \right)$$

$$(L_1,\dots,L_d) = \left(\frac{1}{4}\frac{1}{n}\sum_{i=1}^{n} x_i^2[1] + \lambda, \dots, \frac{1}{4}\frac{1}{n}\sum_{i=1}^{n} x_i^2[d] + \lambda \right)$$