Éléments de corrections TD4

Roland.Bacher@univ-grenoble-alpes.fr

March 27, 2024

0.1 Exo 39

Q et T sont des polynômes (de degré 2, respectivement 3) en les coordonnées de leur argument. Ils sont donc C^∞.

$$\begin{aligned} d^2Q(A)(B,C) &= \frac{d}{ds}\frac{d}{dt}(A+sB+tC)|_{t=s=0} \\ &= \frac{d}{ds}\frac{d}{dt}(A^2+sAB+tAC+sBA+tBC+s^2B^2+stBC+stCB+t^2C^2)|_{t=s=0} \\ &= BC+CB \end{aligned}$$

et on a donc $d^2Q(B,C) = BC+CB$ en tout point.

Pour $dT(A)(B,C)$ on trouve par un calcul analogue la somme du produit des 6 permutations possible de A,B,C, à savoir

$$ABC+ACB+BAC+BCA+CAB+CBA\ .$$

0.2 Exo 40

$F = f\circ\gamma$ avec $\gamma : \mathbb{R} \longrightarrow \mathbb{S}^1$ la courbe paramétrée $\theta \longmapsto (\cos\theta, \sin\theta)$. Comme f et γ sont C^∞, leur composé $F = f\circ\gamma$ aussi.

On a

$$\begin{aligned} F'(\theta) &= \nabla f(\cos\theta,\sin\theta)\cdot\gamma'(\theta) \\ &= (\frac{\partial f}{\partial x},\frac{\partial f}{\partial y})(-\sin\theta,\cos\theta) \\ &= -\sin\theta\frac{\partial f}{\partial x}(\cos\theta,\sin\theta)+\cos\theta\frac{\partial f}{\partial y}(\cos\theta,\sin\theta) \end{aligned}$$

et

$$\begin{aligned}
F''(\theta) = & -\cos\theta\frac{\partial f}{\partial x}(\cos\theta,\sin\theta) - \sin\theta\frac{\partial f}{\partial y}(\cos\theta,\sin\theta) \\
& -\sin\theta\left(\frac{\partial^2 f(\cos\theta,\sin\theta)}{\partial x^2}, \frac{\partial^2 f(\cos\theta,\sin\theta)}{\partial x\partial y}\right)\cdot(-\sin\theta,\cos\theta) \\
& +\cos\theta\left(\frac{\partial^2 f(\cos\theta,\sin\theta)}{\partial y\partial x}, \frac{\partial^2 f(\cos\theta,\sin\theta)}{\partial^2 y}\right)\cdot(-\sin\theta,\cos\theta)
\end{aligned}$$

ce qui donne après simplifications en $\theta = 0$

$$\begin{aligned}
f''(0) &= -\frac{\partial f}{\partial x}(1,0) + \frac{\partial^2 f}{\partial y^2}(1,0) \\
&= -a + b
\end{aligned}$$

0.3 Exo 41

f est la multiplication (dans $\mathbb{C}$, identifié à $\mathbb{R}^2$) par $z = x_1 + x_2 i$. Pour F on a $F(z) = z^2$. Les deux applications sont C^∞.

Plus précisément, on a

$$\begin{aligned}
F(x_1,x_2) &= (x_1^2 - x_2^2, 2x_1x_2), \\
df(x)h &= f(h), \\
dF(x)h &= f(x)h + f(h)x = 2f(x)h \\
d^2 f(x) &= 0 \in \mathcal{L}(\mathbb{R}^2),
\end{aligned}$$

Calcul pour $D^2F(x)(h,k)$ avec $x,h,k \in \mathbb{C}$:

$$\begin{aligned}
d^2(x)(h,k) &= \frac{d}{ds}\frac{d}{dt}(x+sh+tk)^2\big|_{s=t=0} \\
&= \frac{d}{ds}2(x+sh+tk)k\big|_{s=t=0} \\
&= 2hk
\end{aligned}$$

Généralisation: 2.

$$\begin{aligned}
dF(x)(h) &= \frac{d}{ds}\left(f(x+sh)(x+sh)\right)\big|_{x=0} \\
&= \frac{d}{ds}\left(f(x) + df(x)sh + o(s))(x+sh)\right)\big|_{s=0} \\
&= f(x)h + (df(x)h)x
\end{aligned}$$

car

$$\begin{aligned}\frac{d}{ds}f(x)x = 0,\\ \frac{d}{ds}s^2(df(x)h)h\big|_{s=0} = 0,\\ \frac{d}{ds}o(s)\big|_{s=0} = 0\end{aligned}$$

où $o(s)$ désigne un élément de $\mathcal{L}(E, \mathcal{L}(E))$ qui est un petit o de s.

Si f est linéaire on a $df(x)h = f(h)$ et on obtient

$$dF(x)h = f(h)x + f(x)h$$

ce qui est le résultat obtenu ci-dessus.

Autre façon : Définissons $\psi : \Omega \longrightarrow \mathcal{L}(E) \times E$ et $\phi : \mathcal{L}(E) \times E \longrightarrow E$ par $\psi(x) = (f(x), x)$ et $\phi(\alpha, y) = \alpha(y)$. On a $F = \phi \circ \psi$ et donc $dF = d\phi \circ \psi \cdot d\psi$ (où le produit $\cdot$ est le produit des applications linéaires).

On a

$$d\psi(x)h = (df(x)h, h)$$

et

$$d\phi(\alpha, y)(\beta, z) = \alpha(z) + \beta(y)$$

et on obtient donc

$$\begin{aligned}dF(x)h &= (d\phi \circ \psi)(x)) \cdot d\psi(x)h\\ &= d\phi \circ \psi)(x)(df(x)h, h)\\ &= d\phi(f(x), x)(df(x)h, h)\\ &= f(x)h + (df(x)h)x\end{aligned}$$

comme ci-dessus.

2. **Refaire dans le cas non linéaire!**

$$\begin{aligned}F(x + sh + tk) &= f(x + sh + tk)(x + sh + tk)\\ &= f(x)(x) + f(sh + tk)(x) + f(x)(sh + tk) + f(sh)(sh) + f(sh)(tk) + f(tk)(sh) + f(\end{aligned}$$

et on obtient donc

$$d^2F(x)(h, k) = f(h)k + f(k)h$$

en évaluant $\frac{d}{dt}\frac{d}{ds}F(x + sh + tk)$ en $s = t = 0$.

0.4 Exo 42

(1) Fixons $x \neq 0$ et t et posons $h(s) = g(sx, stx)$. On a

$$\begin{aligned} F(x,t) &= \frac{1}{x^2}h(1) \\ &= \frac{1}{x^2}\left(h(0) + h'(0) + \int_0^1 \frac{(1-s)^1}{1!}h''(s)ds\right) \end{aligned}$$

Comme

$$\begin{aligned} h(0) &= g(0,0) = 0, \\ h'(s) &= x\frac{\partial g}{\partial x_1}(sx, stx) + tx\frac{\partial g}{\partial x_2}(sx, stx), \\ h''(s) &= x^2\frac{\partial^2 g}{\partial x_1^2}(sx, stx) + 2tx^2\frac{\partial^2 g}{\partial x_1 \partial x_2}(sx, stx) + t^2x^2\frac{\partial^2 g}{\partial x_2^2}(sx, stx) \\ &= d^2g(sx, stx)((x, tx), (x, tx)) \\ &= x^2 d^2 g(sx, stx)((1,t),(1,t)) \end{aligned}$$

on obtient

$$\begin{aligned} F(x,t) &= \frac{1}{x^2}\left(0 + 0 + \int_0^1 (1-s)d^2g(sx, stx)((sx, stx), (sx, stx))ds\right) \\ &= \int_0^1 (1-s)d^2g(sx, stx)((1,t),(1,t))ds \end{aligned}$$

qui est continu en (s, x) car g est C^2.

(2) En prenant la limite $x \to 0$ ci-dessus (et en utilisant la continuité des dérivée d'ordre 2 de g) on obtient

$$\begin{aligned} F(0,t) &= \int_0^1 (1-s)d^2g(0,0)((1,t),(1,t))ds \\ &= \frac{1}{2}d^2g(0,0)((1,t),(1,t)) \\ &= \frac{1}{2}\left(\frac{\partial^2 g(0,0)}{\partial x_1^2} + 2t\frac{\partial^2 g(0,0)}{\partial x_1 \partial x_2} + t^2\frac{\partial^2 g(0,0)}{\partial x_2^2}\right). \end{aligned}$$

0.5 Exo 43

(1) On a

$$\begin{aligned} 0 < f(x+\lambda h) &= f(x) + df(x)\lambda h + \int_0^1 \frac{(1-t)^1}{1}d^2f(x+t\lambda h)\lambda^2(h,h)dt \\ &< f(x) + \lambda df(x)h + \lambda^2 M \parallel h \parallel^2 \int_0^1 (1-t)dt \\ &= f(x) + \lambda df(x)h + \lambda^2\frac{M}{2} \parallel h \parallel^2 . \end{aligned}$$

(2) Comme la dernière expression ci-dessus est un polynôme de degré 2 en λ qui ne prend que des valeurs strictements positives pour λ dans $\mathbb{R}$, on a

$$(df(x)h)^2 - 4f(x)\frac{M}{2} \parallel h \parallel^2 < 0 \ .$$

On a donc

$$\parallel df(x) \parallel = \sup_{h \neq 0} \frac{|df(x)h|}{\parallel h \parallel} \leq \sqrt{2f(x)M} \ .$$

0.6 Exo 44 (améliorable, je pense)

φ est une composition de fonctions C^1, donc φ est C^1.

Comme $df(x)$ est une isométrie pour tout x, la fonction

$$\varphi(x) = \langle df(x)h, df(x)k \rangle = \langle h, k \rangle$$

est constante et sa différentielle est nulle.

On a aussi

$$\begin{aligned}
&\varphi(x+l) - \varphi(x) \\
&= \langle df(x+l)h, df(x+l)k \rangle - \langle df(x)h, df(x)k \rangle \\
&= \langle df(x+l)h, df(x+l)k \rangle - \langle df(x+l)h, df(x)k \rangle \\
&\quad + \langle df(x+l)h, df(x)k \rangle - \langle df(x)h, df(x)k \rangle \\
&= \langle df(x+l)h, d^2f(x)(l,k) \rangle + \langle d^2f(x)(l,h), df(x)k \rangle + o(\parallel l \parallel)
\end{aligned}$$

ce qui donne

$$d\varphi(x)l = \langle df(x)h, d^2f(l,k) \rangle + \langle df(x)k, d^2f(x)(l,h) \rangle$$

(2) u est symétrique par rapport à k, l car d^2f est une forme bilinéaire symétrique si f est C^2.

u est antisymétrique en h, k car

$$0 = d\varphi(x)l = \langle df(x)h, d^2f(k,l) \rangle + \langle df(x)k, d^2f(x)(h,l) \rangle$$

(3) On a

$$\begin{aligned}
u(h,k,l) &= -u(k,h,l) = -u(k,l,h) \\
&= u(l,k,h) = u(l,h,k) \\
&= -u(h,l,k) = -u(h,k,l)
\end{aligned}$$

et $u(h,k,l) = -u(h,k,l)$ implique $u(h,k,l) = 0$.

Comme $df(x)$ est une isométrie, $df(x)$ est surjective. $d^2f(x)(k,l)$ est donc orthogonal à E ce qui implique que $d^2f(x) = 0$.

(4) Une telle fonction f est une isométrie affine : $d^2f = 0$ implique que df est constant et c'est une isométrie linéaire. Ceci implique $f = df$, à une constante (correspondant à une translation) près.

0.7 Exo 45

On a $\nabla f = 0$ sur un extremum.

(a) $\nabla f = (2x, -2y)$. Unique point critique $(0,0)$ qui est un point selle (ou un point col) : La Hessienne $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ est de signature $(1,1)$. Ce n'est donc pas un extremum.

(b) $\nabla f = (3x^2, 3y^2)$. Unique point critique $(0,0)$ et $f(x,0)$ est strictement croissant en $(0,0)$. Donc pas d'extremum.

(c) $\nabla f = (3x^2 - 3y, 3y^2 - 3x)$. On a

$$y(3x^2 - 3y) + 3y^2 - 3x = 3x(xy - 1) .$$

Premier point critique $x = y = 0$. Sinon, on a $y = \frac{1}{x}$ ce qui implique $3x^2 = 3/x$ donc $x = y = 1$ (dans $\mathbb{R}$). Hessienne:

$$H(x,y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix} .$$

$H(0,0)$ a déterminant -9. Donc $(0,0)$ est point col. $H(1,1)$ est défini positive en $(1,1)$. Donc minimum local $f(1,1) = -1$ en $(1,1)$. Ce minimum est seulement local car $f(x,0) \to -\infty$ pour $x \to -\infty$.

(d) $\nabla f = (2x - 2y, 2y - 2x)$ s'annule en $x = y$ qui est un minimum globale car $f(x,y) = (x-y)^2 + 1$. La Hessienne $\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$ est dégénérée.

0.8 Exo 46

(a) $\nabla f = (x + yz, 1 + xz, xy - 1)$. On a $1 + xz + xy - 1 = x(z + y)$ et $x = 0$ impossible pour un point critique donc $z = -y$ en tout point critique. On a donc $x = y^2$ et $y = 1/x$ ce qui donne $x^3 = 1$. Unique solution réelle: $x = y = 1, y = 1$ et $z = -1$. Hessienne

$$H(x,y,z) = \begin{pmatrix} 1 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix}, H(1,1,-1) = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Signes des déterminants des sous-matrices carrées formés par les premiéres lignes et colonnes sont $+, -, -$ ($J(1,1,-1)$ a déterminant -3) montrent que $J(1,1,-1)$ est de signature $++-$. Donc pas d'extremum global (ni même local).

(b) $\nabla f = (2x + y + 2, x + 2y + 3, 0)$. Points critiques $(-1/3, -4/3, z)$. Hessienne par rapport à x, y seulement :

$$H(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

est défini positive. Donc minima globaux sur la droite $(-1/3, -4/3, \mathbb{R})$.

(c) $\nabla f = (3x^2+3y^2+3y, 6xy+3x, 6z)$. Points critiques: $z = 0$. Deuxième coordonnée donne $x(2y+1) = 0$. Donc soit $x = 0$ et $y(y+1) = 0$ donne $(0,0,0)$ et $(0,-1,0)$. Soit $y = -1/2$ qui donne $(\pm 1/2, -1/2, 0)$. Hessienne

$$H(x,y,z) = \begin{pmatrix} 6x & 6y+3 & 0 \\ 6y+3 & 6x & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

a signature $++-$ en $(0,0,0), (0,-1,0), (1/2,-1/2,0)$ et a signature $+--$ en $(-1/2,-1/2,0)$, donc pas d'extremum local.

0.9 Exo 47

(a) $\nabla f = (2xy, x^2 + (\log y)^2 + 2\log y)$. Points critiques: $x = 0$ (car $y \in \mathbb{R}_+^*$) donc $(\log y)(2 + \log y) = 0$ ce qui donne $y = 1$ ou $y = e^{-2}$. Donc $(0,1)$ et $(0, e^{-2})$ sont les seuls points critiques. La Hessienne $\begin{pmatrix} 2y & 2x \\ 2x & (2+2\log y)/y \end{pmatrix}$ vaut $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ en $(0,1)$ qui est donc un minimum local et $\begin{pmatrix} 2e^{-2} & 0 \\ 0 & -2e^2 \end{pmatrix}$ en $(0, e^{-2})$ qui est un point selle (donc pas un extremum local).

(b) $\nabla f = (-\sin x \cos y, -\cos x \sin y)$. Points critiques $\pi\mathbb{Z}^2 \cup (\pi/2, \pi/2) + \pi\mathbb{Z}^2$. Hessienne: $\begin{pmatrix} -\cos x \cos y & \sin x \sin y \\ \sin x \sin y & -\cos x \cos y \end{pmatrix}$. Hessienne indéfini sur les points critiques $(\pi/2, \pi/2) + \pi\mathbb{Z}^2$ (ce sont des points cols). Hessienne défini positive sur $\pi(a,b) \in \pi\mathbb{Z}^2$ avec $a \not\equiv b \pmod 2$ (minima locaux) et défini négative sur $\pi(a,b)$ avec $a \equiv b \pmod 2$ (maxima locaux).

(c) En coordonnées polaires $f(x,y) = \rho^2(\rho^2 - 1)$ pour $\rho = \sqrt{x^2+y^2}$. Dérivée $4\rho^3 - 2\rho$ s'annulle en $\rho = 0$ et $\rho = 1/\sqrt{2}$. Dérivée deuxième $12\rho^2 - 2$ est négative en $\rho = 0$ (maximum local) et positive sur le cercle de rayon $\rho = 1/\sqrt{2}$ qui est un minimum global car $f(x,y) \to \infty$ pour $\rho \to \infty$.

(d) $\nabla f = (2x - a/(x^2y), 2y - a/(xy^2))$. Comme $x(2x - a/(x^2y)) - y(2y - a/(xy^2)) = 2(x^2 - y^2$, les points critiques satisfont $x^2 = y^2$ donc $y = \pm x$. Si $y = x$ alors $\pm(a/2)^{1/4}(1,1)$ pt critique pour $a > 0$. Sinon $\pm(-a/2)^{1/4}(1,-1)$ pt critique pour $a < 0$. Hessienne

$$\begin{pmatrix} 2 + 2a/(x^3y) & a/(x^2y^2) \\ a/(x^2y^2) & 2 + 2a/(xy^3) \end{pmatrix}$$

qui vaut $\begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}$ en $\pm((a/2)^{1/4}(1,1)$ (avec $a > 0$) et $\begin{pmatrix} 6 & -2 \\ -2 & 6 \end{pmatrix}$ en $\pm(-a/2)^{1/4}(1,-1)$ (avec $a < 0$). Donc minimum local dans tous les cas (c'est aussi un minimum global dans le quadrant concerné mais pas un minimum global sur $(\mathbb{R}_+^*)^2$).

0.10 Exo 48

(1) Sur F on a $x^2+y^2 \leq (x+y)^2$ avec égalité si $xy=0$. Comme $f(x,y) \geq 0$ sur F on a donc $\max_{(x,y)\in F} f(x,y) \leq \max_{z\in[0,+\infty]} z^2e^{-z}$ (en posant $z = x+y$) et $g(z) = z^2e^{-z}$ a des points critiques en $z=0$ et $z=2$. Comme $\lim_{z\to\infty} g(z) = 0 = g(0)$ on a donc $0 \leq f(x,y) \leq g(2) = f(2,0) = 4e^{-2}$ sur F.

L'inégalité $f(x,y) \leq 4e^{-2}$ équivaut à $x^2+y^2 \leq 4e^{x+y-2}$.

0.11 Exo 49

(1) 0 (l'exponentielle gagne!)

(2) On a

$$\max_{x,\|x\|=k} \langle a,x\rangle e^{-\|x\|^2/2} = ke^{-k^2/2}$$

et

$$\min_{x,\|x\|=k} \langle a,x\rangle e^{-\|x\|^2/2} = -ke^{-k^2/2}$$

et $h(x) = xe^{-x^2/2}$ a comme dérivée $h'(x) = (1-x^2)e^{-x^2/2}$ avec ± 1 comme points critiques. Donc minimum $-e^{-1/2}$ en $x=-1$ (correspondant à $-a$ pour la fonction f) et maximum $e^{-1/2}$ en $x=1$ (correspondant à a pour la fonction f).

(3) $df(x)h = e^{-\|x\|^2/2}\left(\langle a,h\rangle - \langle a,x\rangle\langle x,h\rangle\right)$.

(4) Points critiques sont $\pm a$. Valeurs extremales $\pm e^{-1/2}$.

(5) On peut suppose $a = (1,0,0,\dots,0)$ après changement de coordonnées isométrique. On a alors

$$f(x_1,\dots,x_n) = x_1 e^{-1/2\sum_{i=1}^n x_i^2}.$$

Comme $x_2, x_3, \dots$ jouent le même rôle, il suffit de considérer le cas $n=2$. Hessienne de $(x,y) \to xe^{-(x^2+y^2)/2}$ donnée par $e^{-x^2+y^2)/2}\begin{pmatrix} x(x^2-3) & y(x^2-1) \\ y(x^2-1) & x(y^2-1) \end{pmatrix}$ qui vaut $\begin{pmatrix} -4 & 0 \\ 0 & -1 \end{pmatrix}$ en $(1,0)$ (donc maximum local) et $\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$ en $(-1,0)$ (donc minimum local).

Sans coordonnées (et sans garantie concernant l'absence d'erreurs) :

$$\begin{aligned}
d^2(f)(x)(h,k) &= \frac{d}{ds}\frac{d}{dt}\langle a, x+sh+tk\rangle e^{-\|x+sh+tk\|^2/2}|_{s=t=0} \\
&= \frac{d}{dt}\Big(\langle a,h\rangle e^{-\|x+sh+tk\|^2/2} \\
&\qquad +\langle a, x+sh+tk\rangle e^{-\|x+sh+tk\|^2/2}(-\langle x+sh+tk, h\rangle)\Big)\,|_{s=t=0} \\
&= \frac{d}{dt}\Big(\langle a,h\rangle e^{-\|x+tk\|^2/2} - \langle a, x+tk\rangle\langle x+tk, h\rangle e^{-\|x+tk\|^2/2}\Big)\,|_{t=0} \\
&= \Big(\langle a,h\rangle e^{-\|x+tk\|^2/2}(-\langle x,k\rangle) \\
&\quad - \langle a,k\rangle\langle x+tk,h\rangle e^{-\|x+tk\|^2/2} - \langle a,x+tk\rangle\langle k,h\rangle e^{-\|x+tk\|^2} \\
&\quad -\langle a, x+tk\rangle\langle x+tk,h\rangle e^{-\|x+tk\|^2/2}(-\langle x,k\rangle)\Big)\,|_{t=0} \\
&= \frac{-\langle a,h\rangle\langle x,k\rangle - \langle a,k\rangle\langle x,h\rangle - \langle a,x\rangle\langle k,h\rangle + \langle a,x\rangle\langle x,h\rangle\langle x,k\rangle}{e^{\|x\|^2/2}}
\end{aligned}$$

En $x = a$ on trouve pour $k = h$:

$$-\langle a,h\rangle^2 - \langle h,h\rangle \leq 0$$

avec inégalité stricte si $h \neq 0$. C'est donc une forme quadratique définie négative (donc maximum local en $x = a$). Pour $x = -a$ et $k = h$ on trouve la forme quadratique opposée (donc minimum local).

0.12 Exo 50

(1) f est polynomial de degré 2.
$\nabla f = (\sum_{i=1}^n (ax_i + b - y_i)x_i, \sum_{i=1}^n (ax_i + b - y_i))$. Hessienne

$$H(a,b) = \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix}$$

(2) Comme il existe i, j avec $x_i \neq x_j$, on a $\sum_{i=1}^n x_i^2 > 0$ et comme $(x_1, \ldots, x_n)$ n'est alors pas colinéaire à $(1, 1, \ldots, 1)$ l'inégalité Cauchy-Schwarz montre que $H(a,b)$ est défini positive pour tout a, b. La fonction f est donc strictement convexe.

(3) Comme les x_i ne sont pas tous égaux, les parties linéaires en a et b des deux coefficients de ∇f ne sont pas proportionnelles. Il existe donc un unique point critique $(\overline{a}, \overline{b})$ (définie par $\nabla f(\overline{a}, \overline{b}) = (0,0)$).

(4) On a

$$\begin{aligned} f(\overline{a}+a,\overline{b}+b) &= \frac{1}{2}\sum_{i=1}^{n}(\overline{a}x_i+\overline{b}-y_i+ax_i+b)^2 \\ &= \frac{1}{2}\sum_{i=1}^{n}(\overline{a}x_i+\overline{b}-y_i)^2+\sum_{i=1}^{n}(\overline{a}x_i+\overline{b}-y_i)(ax_i+b)+\frac{1}{2}\sum_{i=1}^{n}(ax_i+b)^2 \\ &= f(\overline{a},\overline{b})+\nabla(\overline{a},\overline{b})(a,b)^t+\frac{1}{2}\sum_{i=1}^{n}(ax_i+b)^2 \end{aligned}$$

ce qui termine la preuve car $\nabla(\overline{a},\overline{b})=(0,0)$ par définition du point critique $(\overline{a},\overline{b})$.

Remarque : La droite des moindres carrés minimise la somme des carrés des “erreurs” mais n'est pas la droite qui minimise la somme des distances au carré : Si on prend les trois points $(\pm 1,0),(0,3)$ on trouve $a=0,b=1$ pour la droite des moindres carrés. La droite qui minimise les carrés de la distance est la droite verticale $x=0$.