Examen de mi-parcours du cours d'optimisation différentiable

Durée : 1 heure 30

Les documents ainsi que les calculatrices ne sont pas autorisés.

****

Exercice 0.1 (Calcul de divergence)

On considère l'application

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \longrightarrow & (x^4+y^2+z^2, x^2+y^4+z^2, x^2+y^2+z^4) := (f_i(x,y,z))_{i=1}^3 \end{array} \right.$$

La divergence de f est l'application

$$\mathrm{div} f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x,y,z) & \longrightarrow & \sum_{i=1}^3 \partial_i f_i(x,y,z) \end{array} \right.$$

Calculer la divergence de f sur $\mathbb{R}^3$.

****

Correction de l'exercice 0.1 `[2pts]` On a pour tout $(x,y,z) \in \mathbb{R}^3$,

$$\mathrm{div} f(x,y,z) = 4(x^3+y^3+z^3).$$

****

Exercice 0.2 (Calcul de rotationnel)

On considère l'application

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \longrightarrow & (x^4+y^2+z^2, x^2+y^4+z^2, x^2+y^2+z^4) := (f_i(x,y,z))_{i=1}^3 \end{array} \right.$$

Le rotationnel de f est l'application

$$\mathrm{rot} f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ u & \longrightarrow & \begin{pmatrix} \partial_2 f_3(u) - \partial_3 f_2(u) \\ \partial_3 f_1(u) - \partial_1 f_3(u) \\ \partial_1 f_2(u) - \partial_2 f_1(u) \end{pmatrix} \end{array} \right.$$

Calculer le rotationnel de f sur $\mathbb{R}^3$.

****

Correction de l'exercice 0.2 `[2pts]` On a

$$\mathrm{rot}f(x,y,z) = \begin{pmatrix} 2y-2z \\ 2z-2x \\ 2x-2y \end{pmatrix}$$

****

Exercice 0.3
Soit f la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longrightarrow & \left\{ \begin{array}{c} x^2 \sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) \textbf{ si } (x,y) \neq (0,0) \\ 0 \textbf{ sinon.} \end{array} \right. \end{array} \right.$$

Montrer que f est différentiable en $(0,0)$ et donner son gradient en $(0,0)$.

****

Correction de l'exercice 0.3 `[3pts]` On a quand $(h,k) \neq (0,0)$,

$$\frac{|f(h,k)-f(0,0)|}{\|(h,k)\|_2} \leq \frac{|h|^2}{\|(h,k)\|_2} \leq \|(h,k)\|_2 .$$

Alors quand $(h,k) \to (0,0)$

$$\frac{f(h,k)-f(0,0)}{\|(h,k)\|_2} \to 0.$$

Donc f est différentiable en $(0,0)$ et son gradient vaut $(0,0)$.

****

Exercice 0.4 (Calcul de matrice Jacobienne)
Soit $X \in \mathbb{R}^d$. On considère l'application

$$F : \left\{ \begin{array}{ccc} \mathbb{R}^d & \longrightarrow & \mathbb{R}^d \\ \theta & \longrightarrow & \frac{X-\theta}{\|X-\theta\|_2} \Psi\left(\frac{\|X-\theta\|_2}{\beta}\right) \end{array} \right.$$

où $\Psi : \mathbb{R} \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^1$ et $\|\cdot\|_2$ est la norme euclidienne sur $\mathbb{R}^d$. Soit $\theta \in \mathbb{R}^d$ tel que $\theta \neq X$. L'objectif de l'exercice est de calculer $J(F)(\theta)$ la matrice Jacobienne de F en θ.

a) **Donner le développement limité au premier ordre de $f : \theta \to \|X-\theta\|_2$ en θ ; en déduire son gradient $\nabla f(\theta)$.**

b) **Donner le développement limité au premier ordre de $g : \theta \to \frac{X-\theta}{\|X-\theta\|_2}$ en θ ; en déduire sa Jacobienne $J(g)(\theta)$.**

c) **Donner le développement limité au premier ordre de $h : \theta \to \Psi\left(\frac{\|X-\theta\|_2}{\beta}\right)$ en θ ; en déduire son gradient $\nabla h(\theta)$.**

d) **En déduire la matrice Jacobienne $J(F)(\theta)$ de F en θ.**

On pourra utiliser que pour tout réel α quand $x \to 0, (1+x)^\alpha = 1 + \alpha x + o(x)$ et que pour $F(\theta) = g(\theta)h(\theta)$ on a $J(F)(\theta) = J(g)(\theta)h(\theta) + g(\theta)\nabla h(\theta)^\top$.

****

Correction de l'exercice 0.4 **a)** `[2pts]` On a quand $x \to 0$,

$$f(\theta+x) = \left(\|X-\theta-x\|_2^2\right)^{1/2} = \left(f(\theta)^2 + \left\langle 2(\theta - X), x\right\rangle + o(x)\right)^{1/2} = f(\theta)\left(1 + \left\langle 2(\theta-X)/f(\theta)^2, x\right\rangle + o(x)\right)^{1/2}$$
$$= f(\theta) + \left\langle \frac{\theta - X}{f(\theta)}, x\right\rangle + o(\|x\|).$$

Ceci est le DL1 de f en θ d'où on déduit son gradient $\nabla f(\theta) = (\theta - X)/f(\theta) = (\theta - X)/\left\|\theta - X\right\|_2$.

b) `[2pts]` Quand $x \to 0$, on a

$$g(\theta+x) = \frac{X-\theta-x}{f(\theta+x)} = \frac{X-\theta-x}{f(\theta) + \left\langle \nabla f(\theta), x\right\rangle + o(\|x\|)} = \frac{X-\theta-x}{f(\theta)}\frac{1}{1+\left\langle \nabla f(\theta)/f(\theta), x\right\rangle + o(\|x\|)}$$
$$= \frac{X-\theta-x}{f(\theta)}\left(1 - \left\langle \nabla f(\theta)/f(\theta), x\right\rangle + o(\|x\|)\right) = g(\theta)\left(1 - \left\langle \nabla f(\theta)/f(\theta), x\right\rangle\right) - \frac{x}{f(\theta)} + o(\|x\|)$$
$$= g(\theta) - \left\langle \frac{g(\theta)\nabla f(\theta)}{f(\theta)}, x\right\rangle - \frac{x}{f(\theta)} + o(\|x\|) = g(\theta) + \left(-\frac{g(\theta)\nabla f(\theta)^\top}{f(\theta)} - \frac{I_d}{f(\theta)}\right)x + o(\|x\|).$$

Ceci est le DL1 de g en θ, on en déduit sa matrice Jacobienne

$$J(g)(\theta) = -\frac{g(\theta)\nabla f(\theta)^\top}{f(\theta)} - \frac{I_d}{f(\theta)} = \frac{(X-\theta)(X-\theta)^\top}{\|X-\theta\|_2^3} - \frac{I_d}{\|X-\theta\|_2}.$$

c) `[2pts]` Quand $x \to 0$, on a

$$h(\theta+x) = \Psi\left(\frac{f(\theta+x)}{\beta}\right) = \Psi\left(\frac{f(\theta) + \left\langle \nabla f(\theta), x\right\rangle + o(\|x\|)}{\beta}\right)$$
$$= \Psi\left(\frac{f(\theta)}{\beta}\right) + \Psi'\left(\frac{f(\theta)}{\beta}\right)\left\langle \nabla f(\theta)/\beta, x\right\rangle + o(\|x\|).$$

Ceci est le DL1 de h en θ. On en déduit que son gradient est

$$\nabla h(\theta) = \Psi'\left(\frac{f(\theta)}{\beta}\right)\frac{\nabla f(\theta)}{\beta} = \Psi'\left(\frac{\|X-\theta\|_2}{\beta}\right)\frac{\theta - X}{\beta\left\|X-\theta\right\|_2}.$$

d) `[2pts]` La matrice Jacobienne $J(F)(\theta)$ de F en θ est donnée par :

$$J(F)(\theta) = J(g)(\theta)h(\theta) + g(\theta)\nabla h(\theta)^\top$$
$$= \left[\frac{(X-\theta)(X-\theta)^\top}{\|X-\theta\|_2^3} - \frac{I_d}{\|X-\theta\|_2}\right]\Psi\left(\frac{\|X-\theta\|_2}{\beta}\right) + \frac{X-\theta}{\|X-\theta\|_2}\Psi'\left(\frac{\|X-\theta\|_2}{\beta}\right)\frac{(\theta-X)^\top}{\beta\left\|X-\theta\right\|_2}$$
$$= (X-\theta)(X-\theta)^\top\left[\Psi\left(\frac{\|X-\theta\|_2}{\beta}\right)\frac{1}{\|X-\theta\|_2^3} - \Psi'\left(\frac{\|X-\theta\|_2}{\beta}\right)\frac{1}{\beta\left\|X-\theta\right\|_2^2}\right] - \frac{I_d}{\|X-\theta\|_2}\Psi\left(\frac{\|X-\theta\|_2}{\beta}\right)$$

****

Exercice 0.5 (Optimisation sans contrainte)
Soit f la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x, y, z) & \longrightarrow & (x + z^2) \exp[x(y^2 + z^2 + 1)]. \end{array} \right.$$

1) Montrer que f admet un unique point critique. On le note a^*.

2) Ecrire la formule de Taylor-Young à l'ordre 2 de f en a^*.

3) f admet elle un minimum local en a^* ?

****

Correction de l'exercice 0.5 **1)** `[2pts]` Soit $a = (x, y, z) \in \mathbb{R}^3$. On voit que a est un point critique de f si et seulement si $\nabla f(a) = 0$ qui a lieu si et seulement si

$$\left\{ \begin{array}{cl} 1 + (x + z^2)(y^2 + z^2 + 1) & = 0 \\ xy(x + z^2) & = 0 \\ z(1 + x(x + z^2)) & = 0. \end{array} \right.$$

Si $z \neq 0$ alors la troisième équation implique que $x < 0$ et $x + z^2 > 0$. Or cette inégalité stricte rend impossible la première équation. Donc $z = 0$ et dans ce cas on voit que $y = 0$ et $x = -1$. Réciproquement, on voit que $a^* = (-1, 0, 0)$ satisfait bien le système du dessus et est donc l'unique point critique de f.

2) `[3pts]` La hessienne de f en a^* est

$$\nabla^2 f(a^*) = \frac{1}{e} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Comme f est de classe $\mathcal{C}^2$, on a d'après le formule de Taylor-Young au second ordre que pour $(h, k, l) \to 0$,

$$f(a^* + (h, k, l)) = f(a^*) + \frac{1}{2e} \left(h^2 + 2k^2 + 4l^2 \right) + o(\|(h, k, l)\|^2).$$

3) `[2pts]` Comme $\nabla^2 f(a^*)$ est diagonale, il est facile de voir que $\nabla^2 f(a^*) \succ 0$ (par exemple, ses valeurs propres sont $1/e, 2/e, 4/e$ donc toutes strictement positives ou en voyant que $(h, k, l)^\top \nabla^2 f(a^*)(h, k, l) = h^2 + 2k^2 + 4l^2 > 0$ pour $(h, k, l) \neq 0$). On en déduit que f admet un minimum local en a^*.