Calcul différentiel et optimisation
Partiel du 19 mars 2024

Double licence IASO, L2, Université Paris-Dauphine-PSL, 2023-2024

Durée 2h, Calculatrice et documents non autorisés. Le barême est indicatif.

Dans toute la suite $\|\mathbf{x}\| := (x_1^2 + x_2 + \cdots + x_d^2)^{1/2}$ désigne la norme euclidienne de $\mathbb{R}^d$.

Exercice 1. (Questions de cours – 3 points)

a/ Énoncer la formule de Taylor à l'ordre 2.

b/ Énoncer le théorème de Schwarz.

c/ Donner la définition d'une fonction convexe quelconque sur un sous-ensemble de $\mathbb{R}^d$.

Solution

Voir cours. Ne pas oublier de donner les hypothèses d'application. Par exemple, dans la question c/ l'ensemble de définition doit être convexe.

Exercice 2. (Applications du cours – 6 points)

Soit $F : \mathbb{R}^d \to \mathbb{R}$ définie par $F(\mathbf{x}) = \frac{1}{4}\|\mathbf{x}\|^4$.

a/ Calculer le gradient de F.

b/ Montrer que la hessienne de F est $H_F(\mathbf{x}) = 2\mathbf{x}\mathbf{x}^T + \|\mathbf{x}\|^2 \mathbb{I}_d$.

c/ Montrer que F est convexe sur $\mathbb{R}^d$.

Soit $G : \mathbb{R}^d \to \mathbb{R}$ une fonction de classe C^2. Le **Laplacien** de G, notée ΔG, est la trace de la hessienne de G :

$$\Delta G := \frac{\partial^2 G}{\partial x_1^2} + \cdots + \frac{\partial^2 G}{\partial x_d^2}.$$

d/ Calculer le Laplacien de la fonction F définie précédemment.

e/ Calculer le Laplacien de la fonction $G : \mathbb{R}^2 \setminus \{\mathbf{0}\} \to \mathbb{R}$ définie par $G(x, y) := \ln(x^2 + y^2)$.

Solution

a/ Comme $F(\mathbf{x}) = \left(\sum_{i=1}^d x_i^2\right)^2$, F est un polynôme sur $\mathbb{R}^d$, donc une application de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. On trouve, en appliquant la formule de la chaîne, que, pour tout $1 \le i \le d$,

$$\frac{\partial F(\mathbf{x})}{\partial x_i} = \frac{1}{4} \times 2 \times (2x_i) \left(\sum_{i=1}^d x_i^2\right) = x_i \, \|\mathbf{x}\|^2, \tag{1}$$

puisque

$$\frac{\partial \|\mathbf{x}\|^2}{\partial x_i} = 2\, x_i. \tag{2}$$

Donc

$$\nabla F(\mathbf{x}) = \mathbf{x}\, \|\mathbf{x}\|^2. \tag{3}$$

b/ Soit $j \in \{1, \ldots, d\}$. On dérive partiellement l'expression (1) par rapport à x_j, en utilisant (2). On obtient

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_j x_i} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i x_j} = \delta_{i,j} \|\mathbf{x}\|^2 + 2\, x_i\, x_j, \tag{4}$$

la première égalité venant du théorème de Schwarz. On a utilisé ci-dessus la convention $\delta_{i,j} = 1$ si $i = j$, 0 sinon. Comme $(\mathbf{x}\mathbf{x}^T)_{ij} = x_i x_j$ pour tout $1 \le i, j \le d$, on a bien

$$H_F(\mathbf{x}) = 2\, \mathbf{x}\mathbf{x}^T + \|\mathbf{x}\|^2 \mathbb{I}_d. \tag{5}$$

Alternativement, on peut calculer $F(\mathbf{x} + \mathbf{h})$ pour tout $\mathbf{h} \in \mathbb{R}^d$. On obtient alors :

$$\begin{aligned} F(\mathbf{x}+\mathbf{h}) &= \frac{1}{4}\left(\|\mathbf{x}+\mathbf{h}\|^2\right)^2 = \frac{1}{4}\left(\|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{h}\rangle + \|\mathbf{h}\|^2\right)^2 \\ &= F(\mathbf{x}) + \|\mathbf{x}\|^2 \langle \mathbf{x}, \mathbf{h}\rangle + \left[\langle \mathbf{x}, \mathbf{h}\rangle^2 + \frac{1}{2}\|\mathbf{x}\|^2 \|\mathbf{h}\|^2\right] \qquad (6) \\ &\quad + \|\mathbf{h}\|^4 + \langle \mathbf{x}, \mathbf{h}\rangle \|\mathbf{h}\|^2. \qquad (7) \end{aligned}$$

Ensuite on identifie terme à terme avec la formule de Taylor à l'ordre 2 puisque F est de classe $\mathcal{C}^2$. Les termes dans (7) sont un $o(\|\mathbf{h}\|^2)$: c'est immédiat pour le terme $\|\mathbf{h}\|^4$, tandis que

$$\big|\langle \mathbf{x}, \mathbf{h}\rangle \|\mathbf{h}\|^2\big| \leq \|\mathbf{x}\|\,\|\mathbf{h}\|^3,$$

par l'inégalité de Cauchy–Schwarz. Le terme linéaire dans (6) est égal à $\langle \nabla F(\mathbf{x}), \mathbf{h}\rangle$, donc

$$\langle \nabla F(\mathbf{x}), \mathbf{h}\rangle = \|\mathbf{x}\|^2 \langle \mathbf{x}, \mathbf{h}\rangle = \langle \|\mathbf{x}\|^2 \mathbf{x}, \mathbf{h}\rangle.$$

Comme l'égalité est vraie pour tout vecteur $\mathbf{h}$ de $\mathbb{R}^d$, on retrouve bien l'expression (3) pour le gradient. Enfin, le terme d'ordre 2 dans (6) est égal à $\frac{1}{2}\langle H_F(\mathbf{x})\mathbf{h}, \mathbf{h}\rangle$, d'où la formule (5).

c/ Comme F est une fonction de classe C^∞ définie sur l'ensemble convexe $\mathbb{R}^d$, il suffit de montrer que $H_F(\mathbf{x})$ est positive en tout point. Or, quel que soit $\mathbf{h} \in \mathbb{R}^d$, on a d'après (5),

$$\langle H_F(\mathbf{x})\mathbf{h}, \mathbf{h}\rangle = 2\langle \mathbf{x}\mathbf{x}^T\mathbf{h}, \mathbf{h}\rangle + \|\mathbf{x}\|^2\|\mathbf{h}\|^2 = 2\langle \mathbf{x}, \mathbf{h}\rangle^2 + \|\mathbf{x}\|^2\|\mathbf{h}\|^2 \geq 0,$$

puisque $\mathbf{x}^T\mathbf{h} = \langle \mathbf{x}, \mathbf{h}\rangle \in \mathbb{R}$.

d/ D'après la formule (4), on a

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i^2} = \|\mathbf{x}\|^2 + 2\,x_i^2,$$

pour tout $1 \leq i \leq d$, donc

$$\Delta F(\mathbf{x}) = \mathrm{tr} H_F(\mathbf{x}) = \sum_{i=1}^d \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2} = (d+2)\,\|\mathbf{x}\|^2.$$

e/ On calcule les dérivées partielles premières et secondes de G à l'aide de la règle de la chaîne, et on trouve

$$\frac{\partial G(x,y)}{\partial x} = \frac{2\,x}{x^2+y^2}, \qquad \frac{\partial G(x,y)}{\partial y} = \frac{2\,y}{x^2+y^2},$$

puis

$$\frac{\partial^2 G(x,y)}{\partial x^2} = \frac{2}{x^2+y^2} - \frac{4x^2}{(x^2+y^2)^2} = \frac{2\,(y^2-x^2)}{(x^2+y^2)^2} \quad \text{et} \quad \frac{\partial^2 G(x,y)}{\partial y^2} = \frac{2\,(x^2-y^2)}{(x^2+y^2)^2}.$$

Donc $\Delta G = 0$ sur $\mathbb{R}^2 \setminus \{\mathbf{0}\}$.

Exercice 3. (Étude de points critiques – 4 points)

On définit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ par

$$f(x,y) := (x^2+y^2)\,\mathrm{e}^{x^2-y^2}.$$

Déterminer les points critiques de f et préciser leur nature.

Solution

a/ f est le produit d'un polynôme et de la composée d'une fonction $\mathcal{C}^2$ sur $\mathbb{R}$ par un polynôme.

b/ On calcule les dérivées partielles premières :

$$\frac{\partial f}{\partial x}(x,y) = 2x\,(1+x^2+y^2)\,e^{x^2-y^2}, \quad \frac{\partial f}{\partial x}(x,y) = 2y\,(1-x^2-y^2)\,e^{x^2-y^2}.$$

Donc

$$\nabla f(x,y) = \mathbf{0} \iff \begin{cases} 2x\,(1+x^2+y^2) & = 0 \\ 2y\,(1-x^2-y^2) & = 0 \end{cases} \iff \begin{cases} x & = 0 \\ y\,(1-y^2) & = 0 \end{cases}$$

puisque l'exponentielle et $1+x^2+y^2$ ne s'annulent jamais. On trouve donc trois points critiques :

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{et} \quad \mathbf{B} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Comme $f(x,y) \geq 0$ pour tout $(x,y) \in \mathbb{R}^2$ et que $f(0,0) = 0$, on peut déjà dire, sans faire de calcul supplémentaire, que f admet un minimim global (donc local) en $\mathbf{0}$.

La matrice hessienne en tout point est donnée par :

$$H_f(x,y) = e^{x^2-y^2} \begin{pmatrix} (2+4x^2)(1+x^2+y^2)+4x^2 & -4xy\,(x^2+y^2) \\ -4xy\,(x^2+y^2) & (2-4y^2)(1-x^2-y^2)-4y^2 \end{pmatrix}$$

(Il est inutile de simplifier puisque nous voulons juste calculer la matrice hessienne en trois points). Cette matrice est diagonale en tout point de la forme $(0, y)$. Il suffit donc de regarder les signes des deux éléments sur la diagonale (qui sont les valeurs propres).

Comme $H_f(\mathbf{0}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, on retrouve le fait que $\mathbf{0}$ est un minimum local.

Comme les matrices hessiennes en $\mathbf{A}$ et $\mathbf{B}$ satisfont $H_f(\mathbf{A}) = H_f(\mathbf{B}) = e^{-1} \begin{pmatrix} 4 & 0 \\ 0 & -4 \end{pmatrix}$ et admettent deux valeurs propres de signes contraires, on en déduit que $\mathbf{A}$ et $\mathbf{B}$ sont des points selles.

Exercice 4. (Étude d'un point critique – 5 points)

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$\forall (x, y) \in \mathbb{R}^2, \quad f(x, y) := \left(x^2 - y\right) \left(3x^2 - y\right).$$

a/ Justifier rapidement que f est de classe C^2 sur $\mathbb{R}^2$.

b/ Montrer que f n'est ni majorée ni minorée sur $\mathbb{R}^2$.

c/ Vérifier que f admet un seul point critique noté A. Que vaut la matrice hessienne en ce point ?

d/ Écrire la formule de Taylor à l'ordre 2 au point A.

e/ Représenter sur un même graphique les ensembles :

$$\mathcal{C}_1 = \left\{(\mathrm{x}, \mathrm{y}) \in \mathbb{R}^2, f(x, y) = 0\right\}, \quad \mathcal{C}_2 = \left\{(\mathrm{x}, \mathrm{y}) \in \mathbb{R}^2, f(x, y) < 0\right\}, \quad \mathcal{C}_3 = \left\{(\mathrm{x}, \mathrm{y}) \in \mathbb{R}^2, f(x, y) > 0\right\}.$$

Solution

a/ f est un polynôme.

b/ On remarque que $\lim_{x \to +\infty} f(x, 0) = \lim_{x \to +\infty} (3x^4) = +\infty$, donc f n'est pas majorée. On a aussi $\lim_{x \to +\infty} f(x, 2x^2) = \lim_{x \to +\infty} (-x^4) = -\infty$, donc f n'est pas minorée.

c/ En développant, on obtient $f(x, y) = 3x^4 - 4x^2 y + y^2$. Le gradient de f est donné par

$$\nabla f(x, y) = \begin{pmatrix} 12x^3 - 8xy \\ -4x^2 + 2y \end{pmatrix} = \begin{pmatrix} 4x(3x^2 - 2y) \\ 4x^2 + 2y \end{pmatrix}$$

La deuxième coordonnée du gradient est nulle ssi $y = -2x^2$. En reportant cette condition dans la première coordonnée, on constate qu'elle ne peut s'annuler que pour $x = 0$, donc $y = 0$ également. $\mathbf{0}$ est donc bien le seul point critique.

La matrice hessienne vaut $H_f(x, y) = \begin{pmatrix} 36x^2 - 8y & -8x \\ -8x & 2 \end{pmatrix}$, donc $H_f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$. C'est un point critique dégénéré.

d/ La formule de Taylor au voisinage de $\mathbf{0}$ s'écrit

$$f(x + h, y + k) = f(0, 0) + \left\langle \nabla f(0, 0), \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle + \frac{1}{2} \left\langle H_f(0, 0) \begin{pmatrix} h \\ k \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle + o(h^2 + k^2) = k^2 + o(h^2 + k^2).$$

(On la lit aussi directement sur l'expression de f puisque f est un polynôme.)

e/ On remarquera que $f(x, y) = 0$ ssi $y = x^2$ ou $y = 3x^2$. Donc l'ensemble $\mathcal{C}_1$ est la réunion de de ces paraboles. Entre ces paraboles, f est négative, et en dehors, f est positive.

Exercice 5. (Convexité des fonctions de Cobb-Douglas – 4 points)

Soient α, β deux nombres positifs ou nuls et soit $\mathcal{C} = (0, 1) \times (0, 1) \subset \mathbb{R}^2$. On définit la fonction $f : \mathcal{C} \to \mathbb{R}$ par

$$f(x, y) := -x^\alpha y^\beta.$$

a/ Justifier rapidement que f est de classe C^2 sur $\mathcal{C}$.

b/ Calculer le gradient et la hessienne de f en tout point de $\mathcal{C}$.

c/ Montrer que la fonction $(x, y) \mapsto -x^\alpha$ est convexe sur $\mathbb{R}_*^+ \times \mathbb{R}$ si et seulement si $0 \leq \alpha \leq 1$.

d/ Montrer que f est convexe sur $\mathcal{C}$ ssi α et β vérifient

$$0 \leq \alpha \leq 1, \quad 0 \leq \beta \leq 1 \quad \text{et} \quad \alpha + \beta \leq 1.$$

Solution

a/ Si $(x, y) \in \mathcal{C}$, $x > 0$ et $y > 0$, donc la fonction f peut s'écrire $f(x, y) = -e^{\alpha \ln x + \beta \ln y}$. La fonction $(x, y) \mapsto \alpha \ln x$ est de classe $\mathcal{C}^2$ sur $\mathcal{C}$ comme composée de deux fonctions de classe $\mathcal{C}^2$ (le logarithme sur $\mathbb{R}_*^+$et la fonction affine $(x, y) \mapsto x$ de $\mathcal{C}$ dans $(0, 1) \subset \mathbb{R}_*^+$). On raisonne à l'identique pour la fonction $(x, y) \mapsto \beta \ln y$. f est donc la composée de la fonction exponentielle et d'une fonction de classe $\mathcal{C}^2$ sur $\mathcal{C}$.

b/ On obtient

$$\nabla f(x, y) = \begin{pmatrix} -\alpha\, x^{\alpha-1}\, y^{\beta} \\ -\beta\, x^{\alpha}\, y^{\beta-1} \end{pmatrix} = -x^{\alpha-1} y^{\beta-1} \begin{pmatrix} \alpha\, y \\ \beta\, x \end{pmatrix}$$

et

$$H_f(x, y) = \begin{pmatrix} \alpha\,(1-\alpha)\, x^{\alpha-2} y^{\beta} & -\alpha\beta\, x^{\alpha-1} y^{\beta-1} \\ -\alpha\beta\, x^{\alpha-1} y^{\beta-1} & \beta(1-\beta)\, x^{\alpha}\, y^{\beta-2} \end{pmatrix} = x^{\alpha-2} y^{\beta-2} \begin{pmatrix} \alpha\,(1-\alpha)\, y^2 & -\alpha\beta\, xy \\ -\alpha\beta\, xy & \beta(1-\beta)\, x^2 \end{pmatrix}.$$

c/ On appelle g la fonction définie sur $\mathbb{R}_+^* \times \mathbb{R}$ par $g(x, y) = -x^{\alpha}$. La fonction g étant de classe $\mathcal{C}^2$ sur $\mathbb{R}_+^* \times \mathbb{R}$, il suffit de montrer que sa matrice hessienne est positive. Or,

$$H_g(x, y) = \begin{pmatrix} \alpha\,(1-\alpha)\, x^{\alpha-2} & 0 \\ 0 & 0 \end{pmatrix}. \tag{8}$$

C'est une matrice diagonale qui a deux valeurs propres positives ou nulles si et seulement si $\alpha\,(1-\alpha) \geq 0$ puisque $x > 0$. Comme $\alpha \geq 0$ par hypothèse, ceci est équivalent à $0 \leq \alpha \leq 1$.

d/ Il s'agit de trouver les conditions nécessaires et suffisantes pour que la matrice $H_f(x, y)$ définie par (8) soit positive en tout point de $\mathcal{C}$. Comme $x^{\alpha-2} y^{\beta-2} > 0$, il est équivalent de raisonner sur la matrice $\begin{pmatrix} \alpha\,(1-\alpha)\, y^2 & -\alpha\beta\, xy \\ -\alpha\beta\, xy & \beta(1-\beta)\, x^2 \end{pmatrix}$. On observe que

$$\det(H_f(x, y)) = x^2 y^2 \left(\alpha\beta(1-\alpha)(1-\beta) - \alpha^2\beta^2\right) = x^2 y^2 \left(\alpha\beta(1-\alpha-\beta)\right)$$

et

$$\operatorname{tr}(H_f(x, y)) = \alpha(1-\alpha)\, x^2 + \beta(1-\beta)\, y^2.$$

Si $\alpha = 0$ ou $\beta = 0$, le déterminant et la trace sont nuls ; la fonction f est bien convexe. On suppose désormais que $\alpha > 0$ et $\beta > 0$. Le déterminant est positif ou nul ssi $\alpha + \beta \leq 1$. En faisant tendre y vers 0 en gardant $x > 0$ fixé, on obtient la condition $\alpha(1-\alpha) \geq 0$, qui se résume à $\alpha \leq 1$, puisque $\alpha > 0$. En inversant les rôles joués par x et y, on obtient la dernière condition sur β.