

Mi-parcours du cours d'optimisation différentiable

Durée : 1 heure

Les documents ainsi que les calculatrices ne sont pas autorisés.

Exercice 0.1

Déterminer dans les cas suivants les points critiques et, en utilisant la condition du second ordre, déterminer leur nature (minimum local, maximum local ou indéterminé) :

- 1) $f : x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}_+^* \rightarrow \mathbb{R}, f(x_1, x_2) = x_1^2 + \sqrt{x_2}$,
- 2) $f : x = (x_1, x_2) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \rightarrow \mathbb{R}, f(x_1, x_2) = \sqrt{x_1 x_2}$,
- 3) $f : x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^2, f(x_1, x_2) = (x_1 x_2)^2$,
- 4) $f : x = (x_1, x_2, x_3) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^* \rightarrow \mathbb{R}, f(x_1, x_2, x_3) = \sqrt{x_1} + \sqrt{x_2} + \sqrt{x_3}$,
- 5) $f : x = (x_1, x_2, x_3) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^* \rightarrow \mathbb{R}, f(x_1, x_2, x_3) = \sqrt{x_1 x_2 x_3}$,
- 6) $f : x = (x_1, x_2, x_3) \in \mathbb{R}^3 \rightarrow \mathbb{R}, f(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3$.

Correction de l'exercice 0.1 Sur 8 points 1) [1 point] f n'a pas de point critique vu que pour tout $(x_1, x_2) \in \mathbb{R} \times \mathbb{R}_+^*$,

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 \\ \frac{1}{2\sqrt{x_2}} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

2) [1 point] On a pour tout $(x_1, x_2) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$

$$\nabla f(x_1, x_2) = \begin{pmatrix} \frac{\sqrt{x_2}}{2\sqrt{x_1}} \\ \frac{\sqrt{x_1}}{2\sqrt{x_2}} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Donc f n'a pas de point critique.

3) [2 points] Les points critiques de f sont les axes des ' x ' et ' y ', càd $\mathbb{R} \times \{0\} \cup \{0\} \cup \mathbb{R}$ car

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 x_2^2 \\ 2x_2 x_1^2 \end{pmatrix}.$$

La hessienne de f en (x_1, x_2) vaut

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 2x_2^2 & 4x_1 x_2 \\ 4x_1 x_2 & x_1^2 \end{pmatrix}$$

ainsi pour tout $x_1, x_2 \in \mathbb{R}$, on a

$$\nabla^2 f(x_1, 0) = \begin{pmatrix} 0 & 0 \\ 0 & x_1^2 \end{pmatrix} \text{ et } \nabla^2 f(0, x_2) = \begin{pmatrix} x_2^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Donc 0 est valeur propre de la Hessienne de f en tout point critique. Alors la condition du second ordre ne permet pas de conclure sur la nature des points critiques de f . Néanmoins, on peut voir que pour tout $x_1, x_2 \in \mathbb{R}$, on a $f(x_1, x_2) \geq 0 = f(x_1, 0) = f(0, x_2)$. Donc tous les points critiques de f sont des minima globaux de f .

4) [1 point] f n'a pas de point critique vu que pour tout $(x_1, x_2, x_3) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^*$,

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} \frac{1}{2\sqrt{x_1}} \\ \frac{1}{2\sqrt{x_2}} \\ \frac{1}{2\sqrt{x_3}} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

5) [1 point] f n'a pas de point critique vu que pour tout $(x_1, x_2, x_3) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$,

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} \frac{\sqrt{x_2 x_3}}{2\sqrt{x_1}} \\ \frac{\sqrt{x_1 x_3}}{2\sqrt{x_2}} \\ \frac{\sqrt{x_1 x_2}}{2\sqrt{x_3}} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

6) [2 point] Pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} x_2 + x_3 \\ x_1 + x_3 \\ x_1 + x_2 \end{pmatrix}$$

alors $(0, 0, 0)$ est l'unique point critique de f . La hessienne de f en (x_1, x_2, x_3) vaut

$$\nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} - I_3$$

où I_3 est la matrice identité de $\mathbb{R}^{3 \times 3}$. Comme 3 est valeur propre de $J := (1)_{3 \times 3}$ d'ordre 1 et 0 est valeur propre d'ordre 2 de J , on en déduit que 2 et -1 sont les valeurs propres de $\nabla^2 f(0, 0, 0)$ et donc la condition du second ordre ne permet pas de conclure sur la nature du point critique $(0, 0, 0)$.

Exercice 0.2 (Quotient de Rayleigh)

Soit $A \in \mathcal{S}_n^+$ une matrice symétrique positive de $\mathbb{R}^{n \times n}$. On considère la fonction (appelée quotient de Rayleigh) $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ définie pour tout $x \in \mathbb{R}^n \setminus \{0\}$ par

$$f(x) = \frac{\langle Ax, x \rangle}{\|x\|_2^2}$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire Euclidien de \mathbb{R}^n et $\|\cdot\|_2$ est la norme associée.

- 1) Montrer que f est de classe \mathcal{C}^∞ sur son ensemble de définition.
- 2) On note $\mathcal{S}_2^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ la sphère Euclidienne de \mathbb{R}^n . Montrer que

$$\inf_{x \neq 0} f(x) = \inf_{x \in \mathcal{S}_2^{n-1}} f(x) \text{ et } \sup_{x \neq 0} f(x) = \sup_{x \in \mathcal{S}_2^{n-1}} f(x).$$

En déduire que les problèmes

$$\inf_{x \neq 0} f(x) \text{ et } \sup_{x \neq 0} f(x) \tag{1}$$

admettent une solution.

- 3) Résoudre les deux problèmes ci-dessus. (On peut, par exemple, chercher les points critiques de f ou résoudre les problèmes sous contraintes $\min_{x \in \mathcal{S}_2^{n-1}} f(x)$ et $\max_{x \in \mathcal{S}_2^{n-1}} f(x)$ où \mathcal{S}_2^{n-1} est la sphère Euclidienne unité de \mathbb{R}^n).
- 4) Déterminer la matrice Hessienne de f en tout point de $\mathbb{R}^n \setminus \{0\}$. Expliciter la Hessienne de f en un point critique de f (on pourra faire un développement limité d'ordre de 2 de f en un point critique et utilisé que $(1+t)^{-1} = 1 - t + t^2 + o(t^2)$ quand $t \rightarrow 0$).
- 5) Montrer que la condition d'optimalité du second ordre ne permet pas de conclure sur la nature de tous les points critiques de f qui ne sont pas solution des deux problèmes (1).

Correction de l'exercice 0.2 [sur 12 points] 1) [1 point] f est un ratio de fonction polynomiale de degrés 2 dont le dénominateur ne s'annule pas sur son domaine de définition donc f est de classe \mathcal{C}^∞ .

2) [2 points] Pour tout $x \in \mathbb{R}^n$ tel que $x \neq 0$ et pour tout $\lambda \in \mathbb{R}_+^*$, on a $f(\lambda x) = f(x)$. Donc f est homogène, càd, f est constante sur toute les demi-droite $\{\lambda x : \lambda > 0\}$ pour $x \in \mathcal{S}_2^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$. On a donc

$$\inf_{x \neq 0} f(x) = \inf_{x \in \mathcal{S}_2^{n-1}} f(x) \text{ et } \sup_{x \neq 0} f(x) = \sup_{x \in \mathcal{S}_2^{n-1}} f(x).$$

Or f est continue et \mathcal{S}_2^{n-1} est compact (en tant que fermé borné) donc f atteint son infimum et son supremum sur \mathcal{S}_2^{n-1} . Si on note

$$x_* \in \operatorname{argmin}_{x \in \mathcal{S}_2^{n-1}} f(x) \text{ et } x^* \in \operatorname{argmax}_{x \in \mathcal{S}_2^{n-1}} f(x)$$

on a pour tout $x \in \mathcal{S}_2^{n-1}$, $f(x_*) \leq f(x) \leq f(x^*)$. Alors par homogénéité, pour tout $\lambda > 0$, $f(x_*) \leq f(\lambda x) \leq f(x^*)$. On en déduit donc que f atteint bien ses bornes sur $\mathbb{R}^n \setminus \{0\}$.

3) [4 points] On cherche les points critiques de f : soit $x \in \mathbb{R}^n \setminus \{0\}$. On a

$$\nabla f(x) = \frac{2Ax \|x\|_2^2 - \langle Ax, x \rangle 2x}{\|x\|_2^4}$$

car quand $h \rightarrow 0$

$$\langle A(x+h), x+h \rangle = \langle Ax, x \rangle + \langle Ah, x \rangle + \langle Ax, h \rangle + \langle Ah, h \rangle = \langle Ax, x \rangle + \langle 2Ax, h \rangle + o(\|h\|_2)$$

car $A^\top = A$ et

$$\|x+h\|_2^2 = \|x\|_2^2 + \langle 2x, h \rangle + o(\|h\|_2).$$

On obtient donc $\nabla f(x) = 0$ si et seulement si $Ax \|x\|_2^2 - \langle Ax, x \rangle x = 0$ càd $Ax = f(x)x$. Donc x doit être un vecteur propre de A . Réciproquement, si x est un vecteur propre de A associé à une valeur propre μ , on a nécessairement que $\mu \geq 0$ car A est positive (donc $\mu \|x\|_2^2 = \langle Ax, x \rangle \geq 0$) et $f(x) = \mu$ donc $Ax = f(x)x$ et donc x est un point critique de f . On en déduit donc que les points critique de f sont les points des sous-espace propre de f privé de 0.

De plus, si μ est une valeur propre de A et E_μ est son sous-espace propre associé alors f est constante égale à μ sur E_μ . Alors f atteint son maximum sur l'espace propre de A associé à la plus grande valeur propre et f atteint son minimum sur l'espace propre associé à la plus petite valeur propre de A .

On peut préciser ces valeurs car on sait déterminer les valeurs propres et sous-espaces propres d'une matrice symétrique positive étant donnée sa SVD : comme A est symétrique positive, en écrivant la SVD de $A = PDP^\top$ où $P \in \mathcal{O}(n)$ et $D = \text{diag}(\sigma_1, \dots, \sigma_n)$ où $\sigma_1 \geq \dots \geq \sigma_n$. Pour tout $i = 1, \dots, n$ on note p_i le i -ième vecteur ligne de P . Les sous-espace propres de A sont donnés par les espaces vectoriels engendrés par les p_i associés à la même valeur propre σ_i .

On en déduit que $\text{vect}(p_1, \dots, p_{i_1}) \setminus \{0\}$ où $i_1 = \max(i : \sigma_i = \sigma_1)$ est l'ensemble des solutions du problème $\max_{x \neq 0} f(x)$ et que $\text{vect}(p_{i_2}, \dots, p_n) \setminus \{0\}$ où $i_2 = \min(i : \sigma_i = \sigma_n)$ est l'ensemble des solutions du problème $\min_{x \neq 0} f(x)$.

4) [4 points] Soit $x \in \mathbb{R}^n \setminus \{0\}$ un point critique de f . On sait que $Ax = \sigma x$ où σ est une des valeurs propres de A . On a quand $h \rightarrow 0$,

$$\begin{aligned} f(x+h) &= \frac{\langle Ax, x \rangle + \langle 2Ax, h \rangle + \langle Ah, h \rangle}{\|x\|_2^2 \left(1 + \frac{\langle 2h, x \rangle}{\|x\|_2^2} + \frac{\|h\|_2^2}{\|x\|_2^2} \right)} \\ &= \frac{\langle Ax, x \rangle + \langle 2Ax, h \rangle + \langle Ah, h \rangle}{\|x\|_2^2} \left(1 - \frac{\langle 2h, x \rangle}{\|x\|_2^2} - \frac{\|h\|_2^2}{\|x\|_2^2} + \left(\frac{\langle 2h, x \rangle}{\|x\|_2^2} \right)^2 + o(\|h\|_2^2) \right) \\ &= f(x) + \left(\frac{\langle 2Ax, h \rangle}{\|x\|_2^2} - \frac{\langle Ax, x \rangle \langle 2h, x \rangle}{\|x\|_2^4} \right) \\ &\quad + \left(\frac{\langle Ah, h \rangle}{\|x\|_2^2} - \frac{\langle Ax, x \rangle \|h\|_2^2}{\|x\|_2^4} + \left(\frac{\langle 2h, x \rangle}{\|x\|_2^2} \right)^2 \frac{\langle Ax, x \rangle}{\|x\|_2^2} - \frac{\langle 2h, x \rangle \langle 2Ax, h \rangle}{\|x\|_2^2 \|x\|_2^2} \right) + o(\|h\|_2^2) \end{aligned}$$

où on a utilisé que $(1+t)^{-1} = 1-t+t^2+o(t^2)$ quand $t \rightarrow 0$. On retrouve le gradient de f en x en regardant les termes linéaires en h :

$$\nabla f(x) = \frac{\langle 2Ax, h \rangle}{\|x\|_2^2} - \frac{\langle Ax, x \rangle \langle 2h, x \rangle}{\|x\|_2^4}$$

et la Hessienne de f en x en regardant les termes quadratiques :

$$\frac{1}{2} h^\top \nabla^2 f(x) h = \frac{\langle Ah, h \rangle}{\|x\|_2^2} - \frac{\langle Ax, x \rangle \|h\|_2^2}{\|x\|_2^4} + \left(\frac{\langle 2h, x \rangle}{\|x\|_2^2} \right)^2 \frac{\langle Ax, x \rangle}{\|x\|_2^2} - \frac{\langle 2h, x \rangle \langle 2Ax, h \rangle}{\|x\|_2^2 \|x\|_2^2}$$

on en déduit que

$$(1/2) \nabla^2 f(x) = \frac{A}{\|x\|_2^2} - \frac{\langle Ax, x \rangle I_n}{\|x\|_2^4} + \frac{4xx^\top \langle Ax, x \rangle}{\|x\|_2^4 \|x\|_2^2} - \frac{4xx^\top A}{\|x\|_2^4}$$

où I_n est la matrice identité de $\mathbb{R}^{n \times n}$. Alors pour $Ax = \sigma x$, on a aussi $x^\top A = \sigma x^\top$, et on obtient,

$$\nabla^2 f(x) = \frac{2}{\|x\|_2^2} (A - f(x)I_n).$$

5) [1 point] Si x est un point critique non solution des problèmes 1 alors $f(x) = \sigma x$ pour $\sigma_1 > \sigma > \sigma_n$. On a alors que $A - \sigma I_n$ n'est ni strictement positive ni strictement négative car $\langle (A - \sigma I_n)p_1, p_1 \rangle > 0$ et $\langle (A - \sigma I_n)p_n, p_n \rangle < 0$. Ainsi la condition du second ordre ne permet pas de conclure dans ce cas.