Université Paris Cité
UFR de Mathématiques et Informatique
L2 Mathématiques et Applications 2025-2026

M1 MMAS - Optimisation
Interro n°1

Durée 45mn. Aucun document n'est autorisé.

Exercice 1 (13.5pt)

Soit $n \in \mathbb{N}$, $n \geqslant 2$. On considère l'application $f : \mathbb{R}^n \to \mathbb{R}$ telle que

$$\forall x \in \mathbb{R}^n, \quad f(x) = -\sum_{1 \leqslant i < j \leqslant n} x_i x_j.$$

1. Montrer que f s'écrit sous la forme $f : x \in \mathbb{R}^n \mapsto \frac{1}{2}\langle Ax,\, x\rangle$ avec $A \in \mathcal{S}_n(\mathbb{R})$ à identifier.
2. Donner et justifier brièvement la régularité de f, puis exprimer pour tout $x \in \mathbb{R}^n$, $\nabla f(x)$ et $\nabla^2 f(x)$.
3. La fonction f est-elle convexe sur $\mathbb{R}^n$? *Indication : pour $h \in \mathbb{R}^n$, on pourra s'intéresser à la quantité* $(\sum_{i=1}^n h_i)^2$.
4. Le problème $\inf_{\mathbb{R}^n} f$ est-il bien posé ?

   On considère l'ensemble $D = \{x \in \mathbb{R}^n \;/\; \sum_{i=1}^n x_i = 1\}$.

5. Montrer que D est convexe.
6. Montrer que f est convexe sur D si et seulement si pour tout $h \in \mathbb{R}^n$ tel que $\sum_{i=1}^n h_i = 0$, on a $\langle Ah,\, h\rangle \geqslant 0$.
7. En déduire que f est convexe sur D.
8. Montrer que le problème $\inf_D f$ est bien posé.
9. Est-ce que f admet un unique minimum global sur D ?

<u>Correction</u>.
13.5 = 1 + 1.5 + 2.5 + 1 + 1 + 2.5 + 0.5 + 2 + 1.5

1. Soit $A = (a_{i,j})_{i,j}$ la matrice carré de taille $n \times n$ avec des 0 sur la diagonale et des -1 partout ailleurs. Alors pour tout $x \in \mathbb{R}^n$, on a $\frac{1}{2}\langle Ax,\, x\rangle = \frac{1}{2}\sum_{i=1}^n a_{i,i}^2 x_i^2 + \frac{1}{2}\sum_{i\neq j} a_{i,j} x_i x_j = -\sum_{i<j} x_i x_j = f(x)$ (1pt).
2. La fonction f est polynomiale donc elle est $\mathcal{C}^2$ sur $\mathbb{R}^n$ (même $\mathcal{C}^\infty$). Comme c'est une fonctionnelle quadratique, on sait que pour tout $x \in \mathbb{R}^n$, $\nabla f(x) = Ax$ et $\nabla^2 f(x) = A$ (1.5pt).
3. La fonction f est convexe sur $\mathbb{R}^n$ si et seulement si pour tout $x \in \mathbb{R}^n$ et tout $h \in \mathbb{R}^n$, on a $\left\langle \nabla^2 f(x)h,\, h\right\rangle \geqslant 0$, soit $\langle Ah,\, h\rangle \geqslant 0$ (1pt). Soit $h \in \mathbb{R}^n$, on a

$$\langle Ah,\, h\rangle = -\sum_{i\neq j} h_i h_j = \sum_{i=1}^n h_i^2 - \left(\sum_{i=1}^n h_i\right)^2, \text{ (0.75}pt\text{)}$$

   Posons $h = (1, \dots, 1) \in \mathbb{R}^n$, on a alors $\langle Ah,\, h\rangle = n - n^2 = n(1-n) < 0$ car $n \geqslant 2$ (0.5pt). Ainsi la CNS de convexité pour f sur $\mathbb{R}^n$: $\forall h \in \mathbb{R}^n, \langle Ah,\, h\rangle \geqslant 0$, n'est pas satisfaite. Par conséquent f n'est pas convexe sur $\mathbb{R}^n$ (0.25pt).

4. Soit $t \in \mathbb{R}$. On a $f(t(1,\dots,1)) = -\frac{n(n-1)}{2}t^2$, quantité qui tend vers $-\infty$ quand $t \to +\infty$. Donc $\inf_{\mathbb{R}^n} f = -\infty$, le problème n'est pas bien posé (1pt).
5. D est un hyperplan affine, donc est convexe. Mais justifions le tout de même.
   - Notons $x_0 = (1, 0, \dots, 0) \in D$. Alors on montre par double inclusion que $D = x_0 + V$ où $V = \{h \in \mathbb{R}^n \;/\; \sum_{i=1}^n h_i = 0\}$: si $x \in D$, alors $x - x_0 \in V$ et si $h \in V$ alors $x_0 + h \in D$. Or comme V est un sev de $\mathbb{R}^n$ (hyperplan), V est convexe et donc $D = x_0 + V$ est convexe.
   - Méthode alternative : on retourne directement à la définition. Soit $x, y \in D$, $t \in [0,1]$. Alors $(1-t)x + ty = ((1-t)x_1 + ty_1, \dots, (1-t)x_n + ty_n)$, donc $\sum_{i=1}^n((1-t)x_i + ty_i) = (1-t)\sum x_i + t\sum y_i = (1-t) + t = 1$. D'où $(1-t)x + ty \in D$ (1pt).

6. La fonction f est convexe sur D si et seulement si pour tous $x, y \in D$, on a $\left\langle \nabla^2 f(x)(y-x),\, y-x \right\rangle \geqslant 0$, soit $\langle A(y-x),\, y-x\rangle \geqslant 0$ (1pt).

Il s'agit donc de démontrer que la condition : pour tous $x, y \in D$, $\langle A(y-x),\, y-x\rangle \geqslant 0$, est équivalente à la condition : pour tout $h \in V$, $\langle Ah,\, h\rangle \geqslant 0$. Cette démonstration est essentiellement la même que celle au-dessus montrant que $D = x_0 + V$.

($\Leftarrow$) Soit $x, y \in D$. Alors $y - x \in V$, donc $\langle A(y-x),\, y-x\rangle \geqslant 0$.

($\Rightarrow$) Soit $h \in V$. Soit $x \in D$ et $y = x + h$. Alors $y \in D$ et donc $\langle Ah,\, h\rangle = \langle A(y-x),\, y-x\rangle \geqslant 0$ (1.5pt).

7. Soit $h \in V$. On a déjà vu que $\langle Ah,\, h\rangle = \sum_{i=1}^n h_i^2 - \left(\sum_{i=1}^n h_i\right)^2$, donc ici $\langle Ah,\, h\rangle = \sum_{i=1}^n h_i^2 \geqslant 0$. D'après la précédente question, f est donc bien convexe sur D (0.5pt).

8. L'ensemble D est fermé, non vide et f est continue sur D.

Montrons que f est coercive sur D. On a pour tout $x \in D$, $f(x) = \frac{1}{2}\left(\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2\right) = \frac{1}{2}\|x\|_2^2 - \frac{1}{2}$. Par conséquent si $(x_n)_{n\in\mathbb{N}}$ est une suite de D qui tend en norme 2 vers $+\infty$, on a $(f(x_n))_{n\in\mathbb{N}}$ qui tend vers $+\infty$. Donc par la caractérisation séquentielle de la coercivité, f est coercive sur D.

Ainsi le problème $\inf_D f$ est bien posé (2pt).

9. On a montré que pour tout $h \in V$, $\langle Ah,\, h\rangle = \sum_{i=1}^n h_i^2$. On a juste exploité que c'est une quantité positive pour obtenir via l'équivalence que pour tous $x, y \in D$, on a $\left\langle \nabla^2 f(x)(y-x),\, y-x \right\rangle \geqslant 0$, de sorte que f est convexe sur D. Or $\langle Ah,\, h\rangle = \sum_{i=1}^n h_i^2 > 0$, dès que $h \in V$ avec $h \neq 0$. En reprenant le raisonnement, on déduit alors que $\left\langle \nabla^2 f(x)(y-x),\, y-x \right\rangle > 0$ pour tous $x, y \in D$, $x \neq y$, ce qui implique que f est strictement convexe sur D. Par conséquent f admet exactement un minimum global sur D (puisqu'il en admettait au moins un) (1.5pt).

Exercice 2 (7.5pt)

Soit $n \in \mathbb{N}^*$. On considère la fonction

$$J_\lambda : x \in \mathbb{R}^n \mapsto \frac{1}{n}\sum_{i=1}^n f\left(y_i \langle x,\, x_i\rangle\right) + \lambda \|Lx\|_2^2,$$

où

- $\lambda > 0$,
- $L \in \mathcal{M}_n(\mathbb{R})$ est inversible,
- pour tout $i \in \{1, \dots, n\}$, $y_i \in \{-1, 1\}$ et $x_i \in \mathbb{R}^n$,
- $f : \mathbb{R} \to \mathbb{R}$ est une fonction $\mathcal{C}^2$ positive.

1. Justifier brièvement que J_λ est $\mathcal{C}^2$ sur $\mathbb{R}^n$ et calculer pour tous $x, h \in \mathbb{R}^n$:
   a) $dJ_\lambda(x)(h)$,
   b) $\nabla J_\lambda(x)$,
   c) $\nabla^2 J_\lambda(x)$.
2. Montrer que le problème $\inf_{\mathbb{R}^n} J_\lambda$ est bien posé.

Correction.
7.5 = (2.5 + 1 + 1.5) + 2.5

1. a) La fonction $x \mapsto \lambda \|Lx\|_2^2$ est $\mathcal{C}^2$ sur $\mathbb{R}^n$ car polynomiale. La fonction f est $\mathcal{C}^2$ sur $\mathbb{R}$ et pour tout $i \in \{1, \dots, n\}$, $x \mapsto y_i \langle x,\, x_i\rangle$ est $\mathcal{C}^2$ sur $\mathbb{R}^n$ car linéaire, donc par composée $x \mapsto f(y_i \langle x,\, x_i\rangle)$ est $\mathcal{C}^2$ sur $\mathbb{R}^n$. Finalement par sommations de fonctions $\mathcal{C}^2$ sur $\mathbb{R}^n$, on a J_λ qui est $\mathcal{C}^2$ sur $\mathbb{R}^n$ (0.5pt).

Soient $x, h \in \mathbb{R}^n$. On a

$$\begin{aligned} J_\lambda(x+h) &= \frac{1}{n}\sum_{i=1}^n f\left(y_i \langle x+h,\, x_i\rangle\right) + \lambda \|L(x+h)\|_2^2, \\ &= \frac{1}{n}\sum_{i=1}^n \Big(f\left(y_i \langle x,\, x_i\rangle\right) + f'\left(y_i \langle x,\, x_i\rangle\right) y_i \langle h,\, x_i\rangle + o(h)\Big) + \lambda \|Lx\|_2^2 + 2\lambda \left\langle L^T L x,\, h\right\rangle + o(h), \\ &= J_\lambda(x) + \frac{1}{n}\sum_{i=1}^n f'\left(y_i \langle x,\, x_i\rangle\right) y_i \langle h,\, x_i\rangle + 2\lambda \left\langle L^T L x,\, h\right\rangle + o(h). \end{aligned}$$

Ainsi $dJ_\lambda(x)(h) = \frac{1}{n}\sum_{i=1}^n f'\left(y_i \langle x,\, x_i\rangle\right) y_i \langle h,\, x_i\rangle + 2\lambda \left\langle L^T L x,\, h\right\rangle$ (2pt).

b) Puis $dJ_\lambda(x)(h) = \frac{1}{n}\sum_{i=1}^n f'\left(y_i\left\langle x,\, x_i\right\rangle\right) y_i \left\langle h,\, x_i\right\rangle + 2\lambda\left\langle L^T Lx,\, h\right\rangle = \left\langle \frac{1}{n}\sum_{i=1}^n f'\left(y_i\left\langle x,\, x_i\right\rangle\right) y_i x_i + 2\lambda L^T Lx,\, h\right\rangle$, donc $\nabla J_\lambda(x) = \frac{1}{n}\sum_{i=1}^n f'\left(y_i\left\langle x,\, x_i\right\rangle\right) y_i x_i + 2\lambda L^T Lx$ (1pt).

c) Enfin

$$\begin{aligned}
\nabla J_\lambda(x+h) &= \frac{1}{n}\sum_{i=1}^n f'\left(y_i\left\langle x+h,\, x_i\right\rangle\right) y_i x_i + 2\lambda L^T L(x+h),\\
&= \nabla J_\lambda(x) + \frac{1}{n}\sum_{i=1}^n f''\left(y_i\left\langle x,\, x_i\right\rangle\right) y_i \left\langle h,\, x_i\right\rangle y_i x_i + 2\lambda L^T Lh + o(h),\\
&= \nabla J_\lambda(x) + \frac{1}{n}\sum_{i=1}^n f''\left(y_i\left\langle x,\, x_i\right\rangle\right) \underbrace{y_i^2}_{=1}\, \underbrace{x_i^T h}_{\in\mathbb{R}}\, x_i + 2\lambda L^T Lh + o(h),\\
&= \nabla J_\lambda(x) + \frac{1}{n}\sum_{i=1}^n f''\left(y_i\left\langle x,\, x_i\right\rangle\right) \underbrace{x_i x_i^T}_{\in\mathcal{M}_n(\mathbb{R})}\, h + 2\lambda L^T Lh + o(h),\\
&= \nabla J_\lambda(x) + \left(\frac{1}{n}\sum_{i=1}^n f''\left(y_i\left\langle x,\, x_i\right\rangle\right) x_i x_i^T + 2\lambda L^T L\right) h + o(h).
\end{aligned}$$

Ainsi $\nabla^2 J_\lambda(x) = \frac{1}{n}\sum_{i=1}^n f''\left(y_i\left\langle x,\, x_i\right\rangle\right) x_i x_i^T + 2\lambda L^T L$ (1.5pt).

2. L'ensemble $\mathbb{R}^n$ est fermé non vide, la fonction J_λ est continue sur $\mathbb{R}^n$. Donc pour obtenir que le problème $\inf_{\mathbb{R}^n} J_\lambda$ est bien posé, il reste à prouver que J_λ est coercive sur $\mathbb{R}^n$.

Comme f est positive, on a donc pour tout $x \in \mathbb{R}^n$, $J_\lambda(x) \geqslant \lambda\left\langle L^T Lx,\, x\right\rangle$. La matrice $L^T L$ est symétrique et pour tout $h \in \mathbb{R}^n$, $\left\langle L^T Lh,\, h\right\rangle \geqslant 0$ et $\left\langle L^T Lh,\, h\right\rangle = \left\langle Lh,\, Lh\right\rangle = \|Lh\|_2^2$. Donc $\left\langle L^T Lh,\, h\right\rangle = 0$ si et seulement si $Lh = 0$. Comme L est inversible, on a donc $h = 0$. Ainsi $L^T L \in \mathbb{S}_n^{++}(\mathbb{R})$. Par conséquent il existe $m > 0$ tel que pour tout $x \in \mathbb{R}^n$, $\left\langle L^T Lx,\, x\right\rangle \geqslant m\,\|x\|_2^2$ et donc $J_\lambda(x) \geqslant \lambda m\,\|x\|_2^2$. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de $\mathbb{R}^n$ qui tend vers $+\infty$. Alors par la minoration précédente, on obtient que $J_\lambda(x_n) \underset{n\to+\infty}{\longrightarrow} +\infty$. Par la caractérisation séquentielle de la coercivité, J_λ est bien coercive sur $\mathbb{R}^n$.

D'où le problème est bien posée. En particulier J_λ admet un minimum global sur $\mathbb{R}^n$ (2.5pt).