Université Paris Cité
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

Master 1ère année, MMA, 2023-2024

Optimisation sous contraintes

Rattrapage du 25/06/2024

Durée 1h30. Une feuille recto-verso de notes est autorisée.

Soit $p \geqslant 1$ un réel. Soit $q \in [1, +\infty]$ défini par la relation $\frac{1}{p} + \frac{1}{q} = 1$ (notez que $q = +\infty$ si $p = 1$).

L'objectif de ce sujet est de redémontrer l'inégalité de Hölder

(1) $$\forall x, y \in \mathbb{R}^n, \qquad |\langle x, y\rangle| \leqslant \|x\|_p \|y\|_q ,$$

grâce à des outils d'optimisation.

On rappelle que le signe d'un nombre réel $t \in \mathbb{R}$ est défini par

$$\operatorname{sign}(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \\ 0, & t = 0 \end{cases},$$

et que $t = \operatorname{sign}(t)|t|$.

Exercice 1 (Cas $p = 1$) (5.5pt)

On souhaite donc ici démontrer que

$$\forall x, y \in \mathbb{R}^n, \qquad |\langle x, y\rangle| \leqslant \|x\|_1 \|y\|_\infty .$$

Soit $y \neq 0_{\mathbb{R}^n}$ et on considère le problème d'optimisation suivant

$(\mathcal{P}_1)$ $$\inf_{x \in S} \langle x, y\rangle ,$$

où $S = \{x \in \mathbb{R}^n \ / \ \|x\|_1 \leqslant 1\}$. On notera f la fonction objectif du problème $(\mathcal{P}_1)$.

1. Montrer que le problème $(\mathcal{P}_1)$ est bien posée.
2. On suppose dans cette question uniquement que $n = 2$.
   a) Représenter sur trois figures différentes les ensembles S, $T_S(-e_1)$ et $T_S(-e_1)^*$.
   b) On suppose que $y = (-1, 1)$. Est-ce que $-e_1$ est solution de $(\mathcal{P}_1)$? Justifier en vous aidant en partie de vos figures.
3. Soit $k_0 \in \{1, \dots, n\}$ tel que $|y_{k_0}| = \|y\|_\infty$. On admet que $-\operatorname{sign}(y_{k_0})e_{k_0}$ est solution de $(\mathcal{P}_1)$.
   a) En déduire que pour tout $x \in S$, $\langle x, y\rangle \geqslant -\|y\|_\infty$.
   b) En déduire que pour tout $x \in S$, $-\|y\|_\infty \leqslant \langle x, y\rangle \leqslant \|y\|_\infty$. *Indication : on pourra considérer $-x$ à la place de x.*
   c) Finalement montrer que pour tout $x \in \mathbb{R}^n$,
   $$|\langle x, y\rangle| \leqslant \|x\|_1 \|y\|_\infty .$$

<u>Correction</u>.

$5.5 = 0.75 + (1.5 + 1) + (0.75 + 0.75 + 0.75)$

1. La fonction objectif f est continue sur S la boule unité fermé pour la norme ℓ_1, en particulier S est compact. Donc f admet au moins un minimum global sur S i.e. le problème est bien posée (0.75pt).

2. a) (0.5pt) par figure.
   b) Si $-e_1$ est solution de $(\mathcal{P}_1)$ alors $\nabla f(-e_1) = y = (-1, 1) \in T_S(-e_1)^*$. Or on observe graphiquement que ce n'est pas le cas. Donc $-e_1$ n'est pas solution du problème (1pt).

3. a) Puisque $-\operatorname{sign}(y_{k_0})e_{k_0}$ est solution du problème, on a donc pour tout $x \in S$
   $$\langle x, y\rangle \geqslant \langle -\operatorname{sign}(y_{k_0})e_{k_0}, y\rangle = -\operatorname{sign}(y_{k_0})y_{k_0} = -|y_{k_0}| = -\|y\|_\infty . \ (0.75pt)$$

b) Soit $x \in S$, alors $-x \in S$ également car la boule unité d'une norme est symétrique. D'après la précédente question on obtient donc $\langle -x,\, y\rangle \geqslant -\|y\|_\infty$, d'où $\langle x,\, y\rangle \leqslant \|y\|_\infty$. Ainsi on a bien pour tout $x \in S$: $-\|y\|_\infty \leqslant \langle x,\, y\rangle \leqslant \|y\|_\infty$ (0.75pt).

c) D'après la précédente question, on a donc pour tout $x \in S$, $|\langle x,\, y\rangle| \leqslant \|y\|_\infty$. Soit $x \in \mathbb{R}^n$, $x \neq 0$, alors $\frac{1}{\|x\|_1} x \in S$. En appliquant l'inégalité précédente, on obtient bien le résultat attendu (0.75pt).

Exercice 2 (Cas $p, q \in]1, +\infty[$) (17pt)

On suppose dans cet exercice $p > 1$.

Partie I : questions préliminaires.

Soit $\kappa : \mathbb{R} \to \mathbb{R}$ la fonction définie par

$$\forall t \in \mathbb{R}, \quad \kappa(t) = \begin{cases} |t|^p, & t \neq 0, \\ 0, & t = 0. \end{cases}$$

1. Montrer que κ est dérivable sur $\mathbb{R}$ et que

$$\forall t \in \mathbb{R}, \quad \kappa'(t) = \begin{cases} p\,\mathrm{sign}(t)|t|^{p-1}, & t \neq 0, \\ 0, & t = 0. \end{cases}$$

2. En déduire que κ est $\mathcal{C}^1$ sur $\mathbb{R}$.

Dans la suite, quand cela sera nécessaire, on écrira pour simplifier $|t|^r$ pour tout $t \in \mathbb{R}$ et $r > 0$, au lieu de distinguer le cas $t = 0$ en supposant donc implicitement que $|t|^r = 0$ dans ce cas.

Partie II : différentiabilité de la p-norme.

Soit $n \in \mathbb{N}^*$. On rappelle que la p-norme d'un $x \in \mathbb{R}^n$ est définie par

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}},$$

et que c'est une norme sur $\mathbb{R}^n$.

3. Montrer que $g : x \in \mathbb{R}^n \mapsto \sum_{i=1}^{n} |x_i|^p$ est $\mathcal{C}^1$ sur $\mathbb{R}^n$.

4. Montrer que pour tout $x \in \mathbb{R}^n$, $\nabla g(x) = (\kappa'(x_i))_{1 \leqslant i \leqslant n}$.

Partie III : étude d'un problème d'optimisation.

On s'intéresse au problème d'optimisation suivant

$$(\mathcal{P}_2) \qquad \inf_{x \in S} \langle x,\, y\rangle\,,$$

où $y \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$ est un vecteur fixé et $S = \{x \in \mathbb{R}^n \,/\, g(x) \leqslant 1\}$. On notera f la fonction objectif du problème $(\mathcal{P}_2)$.

5. Donner une interprétation géométrique de S à partir de la p-norme.

6. Montrer que $(\mathcal{P}_2)$ est bien posé.

7. Montrer que S est convexe.

8. A ce stade de l'étude, que pouvez-vous dire de l'unicité de la solution du problème $(\mathcal{P}_2)$?

9. a) Écrire S sous la forme standard (la plus simple possible) étudiée en classe.

   b) Montrer que tous les points de S sont réguliers par rapport aux contraintes le définissant.

10. Soit $x^\star$ un minimum global de f sur S.

   a) Justifier la bonne applicabilité des conditions KKT en $x^\star$ pour le problème $(\mathcal{P}_2)$ et montrer qu'il existe $\lambda \geqslant 0$ tel que

$$\lambda(g(x^\star) - 1) = 0,$$
$$\forall i \in \{1, \ldots, n\}, \quad y_i = -\lambda \kappa'(x_i^\star).$$

b) En déduire que $\lambda > 0$ et pour tout $i \in \{1, \ldots, n\}$, $\operatorname{sign}(x_i^\star) = -\operatorname{sign}(y_i)$. Que vaut alors $g(x^\star)$? Interpréter géométriquement.

c) En calculant $\sum_{i=1}^n |y_i|^q$, montrer que $\lambda = \frac{1}{p}\|y\|_q$.

d) Pour tout $i \in \{1, \ldots, n\}$, en déduire $|x_i^\star|$ puis $x_i^\star$ en fonction de y. Que dire de l'ensemble des solutions de $(\mathcal{P}_2)$?

e) En s'inspirant de la fin du raisonnement de l'exercice 1, en déduire l'inégalité de Hölder (1).

Correction.
$17 = 1.5 + 1 + 0.75 + 0.5 + 0.75 + 0.75 + 0.5 + 0.5 + (0.5 + 1.75) + (1.5 + 1.5 + 1 + 2 + 2.5)$

1. Si $t > 0$, alors $\kappa(t) = t^p = e^{p\ln(t)}$ donc κ est dérivable sur $\mathbb{R}_+^*$ et $\kappa'(t) = \frac{p}{t}t^p = p\operatorname{sign}(t)|t|^{p-1}$.

Si $t < 0$, alors $\kappa(t) = (-t)^p = e^{p\ln(-t)}$ et donc κ est dérivable sur $\mathbb{R}_-^*$ et $\kappa'(t) = \frac{-p}{-t}|t|^p = \frac{p}{t}\operatorname{sign}(t)t|t|^{p-1} = p\operatorname{sign}(t)|t|^{p-1}$.

Puis pour tout $t \neq 0$, $\frac{\kappa(t)-\kappa(0)}{t-0} = \frac{|t|^p}{t} = \operatorname{sign}(t)|t|^{p-1} \underset{t\to 0}{\longrightarrow} 0$ comme $p > 1$. D'où κ est dérivable en 0 et $\kappa'(0) = 0$ (1.5pt).

2. Vu l'expression de κ', κ' est continue sur $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$ comme $p > 1$. Puis la limite à gauche et à droite de κ' en 0 est $0 = \kappa'(0)$, donc κ' est également continue en 0. D'où κ est $\mathcal{C}^1$ sur $\mathbb{R}$ (1pt).

> On aurait pu traiter ces deux questions plus rapidement en évoquant le théorème de la limite de la dérivée.

3. On a pour tout $x \in \mathbb{R}^n$, $g(x) = \sum_{i=1}^n \kappa(x_i) = \sum_{i=1}^n \kappa(\langle e_i, x\rangle)$. Or κ est $\mathcal{C}^1$ sur $\mathbb{R}$, donc par composée et somme de fonctions $\mathcal{C}^1$, g est $\mathcal{C}^1$ sur $\mathbb{R}^n$ (0.75pt).

4. On a pour tout $x \in \mathbb{R}^n$, $\nabla g(x) = (\partial_i g(x))_{1\leqslant i\leqslant n}$. Or $\partial_i g(x) = \kappa'(x_i)$, d'où la conclusion (0.5pt).

5. Soit $x \in \mathbb{R}^n$, on a $g(x) \leqslant 1$ si et seulement si $\|x\|_p \leqslant 1$. Donc $S = \{x \in \mathbb{R}^n \,/\, \|x\|_p \leqslant 1\}$, i.e. S est la boule unité fermé de $\mathbb{R}^n$ pour la p-norme (0.75pt).

6. La fonction objectif $f : x \in \mathbb{R}^n \mapsto \langle x, y\rangle$ est continue sur $\mathbb{R}^n$, l'ensemble S est compact en tant que boule unité fermé d'une norme sur $\mathbb{R}^n$. Ainsi f admet un minimum global sur S et donc $(\mathcal{P}_2)$ est bien posé (0.75pt).

7. On a vu que S est la boule unité (fermé) d'une norme sur $\mathbb{R}^n$. Automatiquement par l'inégalité triangulaire, on déduit que S est convexe (0.5pt).

8. On ne peut rien conclure en l'état concernant l'unicité de la solution de $(\mathcal{P}_2)$, car la fonction objectif f est seulement convexe sur le convexe S et non strictement convexe (0.5pt).

9. a) On a $S = \{x \in \mathbb{R}^n \,/\, f_1(x) \leqslant 0\}$ avec $f_1 : x \in \mathbb{R}^n \mapsto g(x) - 1$. La fonction f_1 est $\mathcal{C}^1$ sur $\mathbb{R}^n$ comme g l'est (0.5pt).

b) Soit $x \in S$. On a $\mathring{S} = \{x \in \mathbb{R}^n \,/\, f_1(x) < 0\}$, donc si $x \in \mathring{S}$ alors $I(x) = \varnothing$ et donc x est un point régulier. Sinon $x \in S \setminus \mathring{S} = \{x \in \mathbb{R}^n \,/\, f_1(x) = 0\}$, donc $I(x) = \{1\}$. Ainsi x est un point régulier si et seulement si $(\nabla f_1(x))$ est libre i.e. $\nabla f_1(x) \neq 0$. Or pour tout $t \in \mathbb{R}$, $\kappa'(t) \neq 0$ si et seulement si $t = 0$, d'où $0 = \nabla f_1(x) = \nabla g(x)$ ssi pour tout $i \in \{1, \ldots, n\}$, $x_i = 0$, i.e. $x = 0$. Ceci contredit le fait que $\|x\|_p = 1$. D'où on a bien $\nabla f_1(x) \neq 0$ et x est encore régulier.

En conclusion tous les point de S sont réguliers par rapport aux contraintes le définissant (1.75pt).

10. a) Comme tous les points de S sont réguliers, c'est aussi le cas de $x^\star$ donc les contraintes sont qualifiées en $x^\star$ et ainsi $x^\star$ satisfait les conditions KKT en tant que minimum global.

Par conséquent il existe $\lambda \geqslant 0$ (condition d'admissibilité duale) tel que

$$\lambda f_1(x^\star) = 0,$$
$$\nabla f(x^\star) + \lambda \nabla f_1(x^\star) = 0_{\mathbb{R}^n}.$$

la première équation étant la condition de complémentarité et la dernière la condition de stationnarité. La condition de stationnarité se réécrit

$$\begin{aligned} & y + \lambda \nabla g(x^\star) = 0, \\ \Leftrightarrow \quad & \forall i \in \{1, \ldots, n\}, \quad y_i = -\lambda \kappa'(x_i^\star). \ (1.5pt) \end{aligned}$$

b) On ne peut donc avoir $\lambda = 0$ car $y \neq 0_{\mathbb{R}^n}$. Ainsi $\lambda > 0$ et par condition de complémentarité, on déduit $f_1(x^\star) = 0$ i.e. $g(x^\star) = 1$ ($x^\star$ appartient à la frontière de S i.e. est sur la sphère unité pour la p-norme). On a de plus pour tout $i \in \{1, \dots, n\}$, $\operatorname{sign}(y_i) = \operatorname{sign}(-\lambda p \operatorname{sign}(x_i^\star)|x_i^\star|^{p-1}) = -\operatorname{sign}(x_i^\star)$ car $\lambda, p > 0$ et $|x_i^\star|^{p-1} \geqslant 0$ (1.5pt).

c) On a

$$\sum_{i=1}^n |y_i|^q = \sum_{i=1}^n \lambda^q p^q |x_i^\star|^{q(p-1)} = \lambda^q p^q \sum_{i=1}^n |x_i^\star|^p = \lambda^q p^q,$$

d'où $\lambda = \frac{1}{p}\|y\|_q$ (1pt).

d) Puis pour tout $i \in \{1, \dots, n\}$

$$|x_i^\star| = \left(\frac{1}{\lambda p}\lambda p |x_i^\star|^{p-1}\right)^{\frac{1}{p-1}} = \frac{1}{(\lambda p)^{\frac{q}{p}}}\left(\lambda p |x_i^\star|^{p-1}\right)^{\frac{q}{p}} = \frac{1}{(\lambda p)^{\frac{q}{p}}}|y_i|^{\frac{q}{p}} = \frac{1}{\|y\|_q^{\frac{q}{p}}}|y_i|^{\frac{q}{p}},$$

ainsi

$$x_i^\star = \operatorname{sign}(x_i^\star)|x_i| = -\operatorname{sign}(y_i)\frac{1}{\|y\|_q^{\frac{q}{p}}}|y_i|^{\frac{q}{p}}.$$

Ainsi $(\mathcal{P}_2)$ admet une unique solution (2pt).

e) On a donc pour tout $x \in S$

$$f(x) = \langle x,\, y\rangle \geqslant \langle x^\star,\, y\rangle = \sum_{i=1}^n -\operatorname{sign}(y_i)\frac{1}{\|y\|_q^{\frac{q}{p}}}|y_i|^{\frac{q}{p}} y_i = -\frac{1}{\|y\|_q^{\frac{q}{p}}}\sum_{i=1}^n |y_i|^{\frac{q}{p}+1} = -\frac{\|y\|_q^q}{\|y\|_q^{\frac{q}{p}}} = -\|y\|_q.$$

On a de même $f(-x) \geqslant -\|y\|_q$, i.e. $\langle x,\, y\rangle \leqslant \|y\|_q$, d'où $|\langle x,\, y\rangle| \leqslant \|y\|_q$.

Soit $x \in \mathbb{R}^n \setminus \{0\}$, alors $\frac{1}{\|x\|_p}x \in S$, donc $\left|\left\langle \frac{1}{\|x\|_p}x,\, y\right\rangle\right| \leqslant \|y\|_q$. Ainsi on a bien montré que pour $x \in \mathbb{R}^n$

$$|\langle x,\, y\rangle| \leqslant \|x\|_p \|y\|_q. \text{ (2.5pt)}$$