Nom :
Prénom :
No étudiant :

Université Paris Cité
UFR Mathématiques et Informatique
2023-2024

M1 MMA - Optimisation sous Contraintes - Interro n°1

Durée 35mn. Aucun document n'est autorisé.

Exercice 1 (11pt)

Soit $m, n \in \mathbb{N}^*$. Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $L \in \mathcal{M}_n(\mathbb{R})$. On suppose que L est symétrique définie positive.

On considère la relation d'ordre (partielle) sur $\mathbb{R}^m$ [1] suivante

$$\forall y, y' \in \mathbb{R}^m, \quad y \preceq y' \Leftrightarrow \forall i \in \{1, \ldots, m\},\, y_i \leqslant y'_i.$$

On fixe $y_0 \in \mathbb{R}^m$ tel que $y_0 \in \mathrm{Im}(A)$ et on s'intéresse au problème d'optimisation suivant

$$(\mathcal{P}) \qquad \inf_{x \in S} \|Lx\|_2^2 .$$

où $S = \{x \in \mathbb{R}^n \ / \ Ax \preceq y_0\}$.

1. Montrer que le problème $(\mathcal{P})$ est bien posé.
2. Montrer que S est convexe.
3. Le problème $(\mathcal{P})$ admet-il une unique solution ?
4. On suppose dans cette question que $n = 2$, $m = 3$, puis

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{et} \quad y_0 = (1, 1, 1).$$

L est toujours une matrice symétrique définie positive quelconque (mais ici de taille 2×2).

a) Représenter l'ensemble S ainsi que ses éventuels points extrêmes. *Aucune justification n'est attendue.*

b) Donner l'ensemble des solutions de $(\mathcal{P})$.

Correction.

11 = 4 + 2 + 1.5 + (2 + 1.5)

1. Notons $f : x \in \mathbb{R}^n \mapsto \|Lx\|_2^2$. Il suffit de montrer que S est non vide fermé, f continue et coercive sur S pour obtenir que $(\mathcal{P})$ soit bien posée (0.5pt). Vérifions chacun de ces points.

— Comme $y_0 \in \mathrm{Im}(A)$, il existe $x_0 \in \mathbb{R}^n$ tel que $Ax_0 = y_0$, en particulier $Ax_0 \preceq y_0$, i.e. $x_0 \in S$. D'où S est non vide (0.5pt).

— Posons pour tout $i \in \{1, \ldots, m\}$, $g_i : x \in \mathbb{R}^n \mapsto \langle e_i,\, Ax - y\rangle = (Ax)_i - y_i$, où e_i est le i-ème vecteur de la base canonique de $\mathbb{R}^n$. Alors comme g_i est linéaire, g_i est continue sur $\mathbb{R}^n$ et donc $S_i := g_i^{-1}(\mathbb{R}_-)$ est un fermé de $\mathbb{R}^n$. Or $S = \cap_{i=1}^m S_i$, donc S est un fermé de $\mathbb{R}^n$ comme intersection de fermés (1.5pt).

— f est continue sur $\mathbb{R}^n$ et comme L est symétrique définie positive, il existe $\lambda > 0$ tel que λ est la plus petite valeur propre de L et pour tout $x \in \mathbb{R}^n$, $f(x) = \|Lx\|_2^2 \geqslant \lambda^2 \|x\|_2^2$. D'où $\lim\limits_{\|x\|_2 \to +\infty} f(x) = +\infty$ i.e. f est coercive sur $\mathbb{R}^n$ donc en particulier sur S (1.5pt).

2. Soit $y, y', z, z' \in \mathbb{R}^m$ tels que $y \preceq y'$ et $z \preceq z'$. Soit $\alpha, \beta \in \mathbb{R}_+$. Alors pour tout $i \in \{1, \ldots, m\}$, on a $y_i \leqslant y'_i$ et $z_i \leqslant z'_i$, d'où $\alpha y_i + \beta z_i \leqslant \alpha y'_i + \beta z'_i$. Ainsi $\alpha y + \beta z \preceq \alpha y' + \beta z'$ [2].

Soit $x, x' \in S$ et $t \in [0, 1]$. Ainsi $Ax \preceq y_0$, $Ax' \preceq y_0$ et par le résultat précédent, on déduit que $(1-t)Ax + tAx' \preceq (1-t)y_0 + ty_0$, i.e. $A\left((1-t)x + tx'\right) \preceq y_0$, d'où $(1-t)x + tx' \in S$. S est bien convexe (2pt).

Autre méthode. Comme pour tout $i \in \{1, \ldots, m\}$, g_i est linéaire, on a g_i convexe. Ainsi S_i, qui est l'ensemble de sous-niveau 0 de g_i, est un convexe de $\mathbb{R}^n$. Or $S = \cap_{i=1}^m S_i$, donc S est un convexe de $\mathbb{R}^n$ comme intersection d'ensembles convexes.

3. La fonction f est $\mathcal{C}^2$ sur $\mathbb{R}^n$ et pour tout $x \in \mathbb{R}^n$, on a $\nabla^2 f(x) = 2L^2 \succ 0$ puisque L est définie positive. Donc f est strictement convexe sur $\mathbb{R}^n$. Et comme S est convexe, f est également strictement convexe sur S.

Comme f admet un minimum global sur S d'après 1., ce dernier est donc unique (1.5pt).

1. Cette relation d'ordre est appelée *ordre produit*.
2. On a montré que la relation d'ordre (partielle) $\preceq$ définie sur $\mathbb{R}^m$ est compatible avec la structure d'espace vectoriel de $\mathbb{R}^m$. On dit alors que $(\mathbb{R}^m, \preceq)$ est un espace vectoriel ordonné. Voir `https://fr.wikipedia.org/wiki/Espace_vectoriel_ordonné` pour plus d'informations.

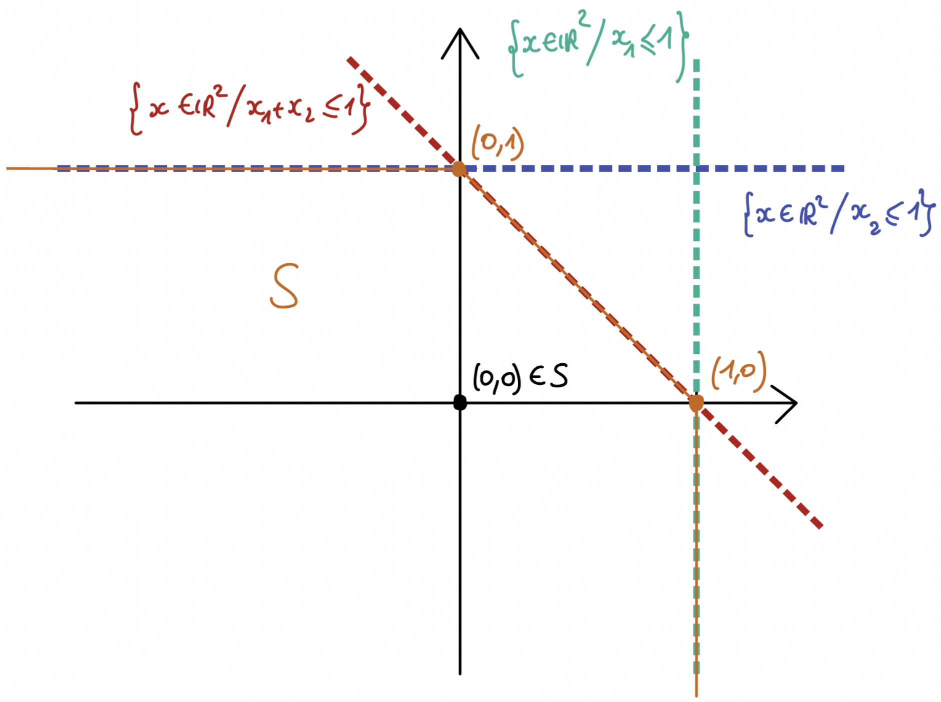

FIGURE 1. Représentation de l'ensemble S (en orange) qui est l'intersection des ensembles $\{x \in \mathbb{R}^2 \,/\, x_1 \leqslant 1\}$, $\{x \in \mathbb{R}^2 \,/\, x_2 \leqslant 1\}$ et $\{x \in \mathbb{R}^2 \,/\, x_1 + x_2 \leqslant 1\}$. Les points extrêmes de S sont : $(1, 0)$ et $(0, 1)$.

4. a) On a $x = (x_1, x_2) \in S$ si et seulement si $Ax \preceq (1, 1, 1)$ si et seulement si $x_1 \leqslant 1$, $x_2 \leqslant 1$ et $x_1 + x_2 \leqslant 1$. Voir la figure 1 pour une représentation de S (2pt).

b) L'ensemble S définie ici est de nouveau non vide car contient $(0, 0)$, puisque $A(0, 0) = (0, 0, 0) \preceq (1, 1, 1)$. Les autres hypothèses vérifiées en 1. sont toujours satisfaites, donc $(\mathcal{P})$ admet bien une solution par le même raisonnement qu'en première question. De plus l'unicité tient également toujours.

Or $f((0, 0)) = 0 \leqslant f(x)$ pour tout $x \in S$, d'où $(0, 0)$ est un minimum global de f sur $\mathbb{R}^n$. C'est donc l'unique solution de $(\mathcal{P})$. On a $\operatorname{argmin}_S f = \{(0, 0)\}$ (1.5pt).

Complément sur la coercivité de f.

On a pour tout $x \in \mathbb{R}^n$, $f(x) = \langle Lx,\, Lx\rangle = \left\langle L^T Lx,\, x\right\rangle = \left\langle L^2 x,\, x\right\rangle$, comme $L^T = L$ puisque L est symétrique. Sachant que L est symétrique réelle, L est diagonalisable sur $\mathbb{R}$ dans une base orthonormée de $\mathbb{R}^n$ notée $(v_1, \dots, v_n)$. Notons $(\lambda_1, \dots, \lambda_n)$ les valeurs propres réelles associées, qui sont donc strictement positives comme L est définie positive. Soit $x \in \mathbb{R}^n$, alors $x = \sum_{i=1}^n \mu_i v_i$ et donc $f(x) = \left\langle L^2 x,\, x\right\rangle = \left\langle L\left(\sum_{i=1}^n \lambda_i \mu_i v_i\right),\, \sum_{i=1}^n \mu_i v_i\right\rangle = \left\langle \sum_{i=1}^n \lambda_i^2 \mu_i v_i,\, \sum_{i=1}^n \mu_i v_i\right\rangle = \sum_{i=1}^n \lambda_i^2 \mu_i^2$ (la dernière égalité étant obtenue du fait que $(v_1, \dots, v_n)$ est orthonormée). En notant $\lambda \in \mathbb{R}$ la plus petite valeur propre de L, on a $\lambda > 0$ et pour tout $i \in \{1, \dots, m\}$, $\lambda_i \geqslant \lambda$, d'où par croissance de $t \in \mathbb{R} \mapsto t^2$ sur $\mathbb{R}_+$, on obtient pour tout $i \in \{1, \dots, m\}$, $\lambda_i^2 \geqslant \lambda^2$. Ainsi $f(x) \geqslant \lambda^2 \sum_{i=1}^n \mu_i^2 = \lambda^2 \|x\|_2^2$ (la dernière égalité étant obtenue par définition de la décomposition de x dans la base orthonormée $(v_1, \dots, v_n)$).

Complément sur le calcul de $\nabla^2 f$.

On vient de montrer que pour tout $x \in \mathbb{R}^n$, $f(x) = \left\langle L^2 x,\, x\right\rangle = \frac{1}{2}\left\langle 2L^2 x,\, x\right\rangle$. La matrice carré $2L^2$ étant symétrique, f est donc une fonctionnelle quadratique et on a vu alors que pour tout $x \in \mathbb{R}^n$, $\nabla^2 f(x) = 2L^2$.

Alternativement on peut refaire à la main le calcul. Par composée de fonctions $\mathcal{C}^2$ ($x \mapsto Lx$ l'est car linéaire, puis c'est aussi le cas de la norme euclidienne au carré), f est $\mathcal{C}^2$ sur $\mathbb{R}^n$[3]. Soit $x, h \in \mathbb{R}^n$, on a $f(x+h) = \langle L(x + h),\, L(x + h)\rangle = f(x) + 2\langle Lx,\, Lh\rangle + o(h) = f(x) + \left\langle 2L^2 x,\, h\right\rangle + o(h)$, donc par identification, $\nabla f(x) = 2L^2 x$. Puis $\nabla f(x + h) = 2L^2(x + h) = \nabla f(x) + 2L^2 h$, d'où par identification $\nabla^2 f(x) = 2L^2$.

3. Mentionner la régularité de f dès le début permet de s'affranchir de devoir justifier par le calcul du DL que f est bien différentiable, puis deux fois différentiable. Le calcul du DL est alors uniquement mené pour identifier le gradient et la hessienne de f